A Theoretical Framework for MAC-layer Design in Wireless Networks

vorgelegt von
Diplom-Ingenieurin
Angela Feistel
geb. Wiesjahn
aus Prenzlau

Von der Fakultät IV - Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
Dr.-Ing.

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Anja Feldmann
Berichter: Dr. habil. Sławomir Stańczak
Berichter: Prof. Dr. Dr. Holger Boche (TU München)
Berichter: Prof. Dr. Tansu Alpcan

Tag der wissenschaftlichen Aussprache: 21.03.2011

Berlin, 2011

D83

Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit einem Themengebiet an der Schnittstelle von Kommunikationstheorie, Vernetzung und Algorithmendesign: Wie sollen die vorhandenen Ressourcen Leistung, Raum und Richtung, Zeit und Frequenz in einem verteilten drahtlosen Netz mit einer allgemeinen Netztopologie den Nutzern zugewiesen bzw. von ihnen verwendet werden? Bekanntermaßen kann die Performanz eines drahtlosen Netzes erheblich verbessert werden, indem man die Vorteile einer intelligenten Ressourcenallokation und eines intelligenten Interferenzmanagements ausnutzt. Unter Verwendung eines mathematischen Frameworks untersuchen wir Probleme der Ressourcenallokation, die auf einer gewichteten Summe von Nutzenfunktionen oder der Unterstützung von QoS(Quality-of Service)-Anforderungen basieren bzw. beide Ansätze vereinen, und legen einen Schwerpunkt auf den Entwurf verteilter Algorithmen. Unsere Resultate sind auf drahtlose Netze mit verrauschten Kanälen und allgemeinen Leistungsbeschränkungen anwendbar und können mit einer beliebigen Routing- und/oder Scheduling-Strategie kombiniert werden. Neben dem Entwurf von Strategien zur Ressourcenallokation, der Ableitung ihrer Eigenschaften sowie geeigneter Algorithmen gewähren wir Einblicke in die Aspekte der praktischen Implementierung.

Im Hauptteil dieser Arbeit konzentrieren wir uns auf Modelle drahtloser Netze, in denen Interferenz als Rauschen behandelt wird. Dieser Teil leistet einen Beitrag dazu, die Mechanismen der Leistungskontrolle und des Beamformings sowie deren grundlegende Wechselbeziehungen und Abhängigkeiten besser zu verstehen. Zu Beginn identifizieren wir die optimale Leistungsallokation, welche eine Summe von gewichteten Nutzenfunktionen maximiert und gleichzeitig die durch SIR-Anforderungen repräsentierten QoS-Anforderungen der Nutzer unterstützt. Dazu analysieren wir einen primal-dual Algorithmus, welcher auf der dazugehörigen linearen Lagrangefunktion basiert, und zeigen, dass dieser Algorithmus effizient in einem verteilten Netz implementiert werden kann. Da jedoch die SIR-Anforderungen für einige Kanalzustände möglicherweise nicht unterstützbar sind und um mögliche best-effort Nutzer zu berücksichtigen, kombinieren wir das Problem der Nutzenfunktionen mit einer Barrier-Funktion, um das ursprüngliche Problem möglichst gut zu approximieren. Wir beweisen wichtige Eigenschaften dieses Ansatzes und stellen zur Lösung einen verteilten rekursiven Algorithmus mit globaler Konvergenz bereit.

Anschließend widmen wir uns den Abhängigkeiten der Leistungen und der Beamformer, indem wir das anspruchsvolle Problem der gemeinsamen Optimierung der Ressourcen

Sendeleistung und Empfangsbeamformer in einem verteilten Netz und für beide oben genannten Ansätze näher beleuchten. Leider ist nicht bekannt, für welche Klasse von Nutzenfunktionen das entsprechende gemeinsame Optimierungsproblem in ein konvexes Problem überführt werden kann, so dass eine globale Lösung in verteilten drahtlosen Netzen gefunden werden könnte. Deswegen stellen wir eine alternative verteilte algorithmische Lösung bereit, die zu einem stationären Punkt konvergiert. Wir untersuchen ihr Verhalten in realen drahtlosen Netzen, in denen Messungen und Schätzungen verrauscht sind, unter Verwendung der stochastischen Approximation. Da der zweite typische Optimierungsansatz, der auf QoS-Anforderungen basierende relative Max-Min-Ansatz, nicht differenzierbar ist, scheint dieses Konzept mit einer verteilten Lösung nicht vereinbar. Wir vermuten jedoch, dass dieses gemeinsame relative Max-Min-Problem für ein Netz mit allgemeinen Leistungsbeschränkungen optimal gelöst werden kann, u.z. unter Verwendung der gewichteten Summe der Nutzenfunktionen. Die Verbindung zwischen beiden Problemen wird dabei durch die Wahl der Gewichte hergestellt. Der abgeleitete Algorithmus ist zudem fast vollständig verteilt implementierbar.

Im letzten Teil der Arbeit untersuchen wir das gemeinsame Problem des Schedulings und der Leistungkontrolle in einem drahtlosen Netz mit orthogonalen Kanälen mit dem Ziel, die gewichtete Summenrate zu maximieren. Da die Schedulingvariablen typischerweise diskret sind, stellen wir mit Hilfe der Relaxation dieser Variablen eine obere Schranke für das gestellte Problem bereit. Anschließend leiten wir interessante Eigenschaften des ursprünglichen, diskreten Problems und eine heuristische Lösung ab.

Zum Abschluss der Arbeit zeigen wir vielversprechende weiterführende Themen und Problemstellungen auf.

Abstract

This thesis addresses a problem at the nexus of communication theory, networking and algorithm design: how to allocate radio resources to users in distributed wireless networks with a general network topology? We emphasize that the network performance can be significantly improved by taking advantage of an intelligent resource allocation and interference management. Based on a mathematical framework we study resource allocation problems which are utility-based, QoS(Quality-of Service)-based or a combination of both, with an emphasis on distributed algorithm design. Our results apply to networks with noisy channels and general power constraints and can be combined with any fixed routing and/or scheduling strategy. In addition to algorithmic solutions and their properties we give insights into practical implementation issues.

In the main part of this thesis we focus on models of wireless networks where interference is treated as noise. This is one step forward to a better understanding of the mechanisms power control and beamforming as well as their interdependencies. First we combine the QoS-based and utility-based power control approach by incorporating QoS support in terms of minimum SIR targets into the utility-based power control problem. We analyze a primal-dual algorithm that is used to find a stationary point of an associated linear Lagrangian function and show that this algorithm can be efficiently implemented in a distributed manner. Since the SIR requirements might be infeasible for some channel states and in order to incorporate possible best-effort users we combine the traditional utility function with a barrier function to closely approximate the utility-based power control problem with QoS support. We prove relevant properties of optimal solutions and propose a distributed recursive algorithm with global convergence.

After that, we address the challenging problem of jointly optimizing powers and receive beamformers in distributed wireless networks. Unfortunately it is not known for which class of utility functions such a joint utility-based problem can be converted into a convex one such that a global solution can be found in distributed wireless networks. Nevertheless we reveal a distributed algorithmic solution that converges to a stationary point and investigate its behaviour in real-world wireless networks using the framework of stochastic approximation. Considering on the other hand the QoS-based approach, the main challenge is to solve the max-min SIR-balancing problem in a distributed manner. In order to address this challenge, we solve the max-min SIR-balancing problem for a network with general power constraints by instead maximizing a weighted sum of utilities. The con-

nection between both problems is provided by the weights of the utilities. We conjecture that the algorithm derived converges to the optimal max-min SIR-balancing solution but is semi-distributed due to some control data exchange which is required to calculate the weights.

Finally we discuss a joint scheduling and power control problem in a wireless network with orthogonal channels. The main objective is to maximize a weighted sum rate. First, by relaxing the discrete scheduling variables and by solving the relaxed problem we provide an upper bound to the original discrete problem. Subsequently we analyze the original discrete problem and derive a heuristic solution.

We conclude the thesis by highlighting topics of future interest.

Acknowledgement

During the last years I received support from many people; this thesis would not have been possible without them. First of all I thank my advisors Dr. Sławomir Stańczak and Prof. Holger Boche for the possibility to work in such a stimulating research environment, for their guidance and support. My special thanks go to Dr. Stańczak and my room mate Shuying Shi for sharing countless ideas and discussions. I am very grateful to Prof. Tansu Alpcan, who dedicated his time and energy to act as second referee of this thesis. Thanks for valuable discussions at the end of the thesis.

My gratitude extends to all colleagues of the Fraunhofer German-Sino Lab for Mobile Communications, the Heinrich-Hertz-Institute and the Technical University of Berlin, who where always open for discussions. My special thanks here go to Shuying Shi, Michał Kaliszan and Mario Goldenbaum.

My most heartfelt gratitude goes to my family. I would like to thank my parents for their endless support and my parents in law for their endless encouragement. I thank my husband Stefan for his constant understanding and support, for his love and energy. Finally I thank my little daughter Nora, born in between these years, for putting things into perspective.

Contents

List of Figures

1 Introduction

1.1 Motivation

The origins of wireless communications date back to the 19th century. In 1888 Heinrich Hertz verified experimentally Maxwell's theory. The Italian engineer Guillermo Marconi was the first to practically and commercially use the wireless communication. Starting with successfull experiments in 1895 he established the first radiotelegraph transmissions in 1899 [Fre02]. The wireless communication network was born. In the 1920s radio broadcasting conquered rapidly the households. Since then, wireless communication has developed at an increasing speed culminating in the breakthrough of personal mobile wireless communication in the early 1990s in form of the second generation (GSM). A wireless access provides users with mobility, eases deployment where wired connections are not feasible, and may reduce costs.

Since the early 1990s, in addition to voice applications at anytime and anywhere, data services and applications increase in popularity with each year. This drives the evolution of new wireless networks and access technologies (WLAN, UMTS/HSPA, LTE, WiMAX) which are needed to fulfill the increasing demands on data rate of the wireless communication services. Besides, the fourth generation (LTE, WiMAX) is expected to be a cellular network with advanced attributes including the wireless backhaul in WiMAX networks or the possibility of relaying to save infrastructure costs. Last but not least, a very interesting and important wireless network is the wireless sensor network. Promising applications are envisaged in many areas such as industrial process monitoring and control, environment monitoring, home automation, healthcare applications and many more. Wireless networks like the backhaul of a WiMAX network, wireless LANs, networks with relays, sensor networks and other future wireless networks would benefit highly from a decentralized or distributed network organization that avoids a lot of control overhead.

All wireless networks have in common that they are characterized by limited frequency and power resources as well as time-varying fading channels due to both the wireless environment and the user mobility. Most importantly, the wireless spectrum is a shared medium, meaning that wireless links interfere with each other if they are not orthogonalized. Thus, resource allocation and interference management belong to the most important issues for implementing future wireless networks. The objective is to utilize the limited resources as efficiently as possible and at the same time to ensure the quality of service (QoS) in terms of data rate, bit error rate or delay for all customers. We emphasize that the QoS is determined by the signal-to-interference ratio (SIR) and thus can be represented by minimum SIR targets in many cases of practical interest [TH99, BG92]. The list of resources to be assigned includes frequency, power, time, space, codes, routes to name the most important ones. The resource allocation and interference management mechanisms that may control and optimize these resources include the media access control (MAC) in terms of power control, time and frequency scheduling; routing, multiple antenna techniques, and so on. In the main part of this thesis we assume that interference is treated as noise and focus on the mechanisms power control and adaptive beamforming with one data stream only. Besides we also discuss the mechanisms scheduling and power control under the simplifying assumption that the links are orthogonal to each other. Further, motivated by the fact that many future wireless networks require a decentralized network organization, in this thesis we devise power control, adaptive beamforming and scheduling strategies and algorithms that can be executed distributedly. In more detail, we focus on wireless networks that require or benefit from a decentralized or distributed network organization and optimization. A typical example for such a wireless network is the wireless backhaul of a WiMAX (IEEE 802.16) network depicted exemplarily in figure 1.1. Here the channels vary slowly and the algorithms are able to track the wireless channel. We emphasize, that by a decentralized or distributed network organization and optimization we usually mean that no central network controller is required. In this case the actions to control and optimize a wireless network are decoupled and executed seperately for each link.

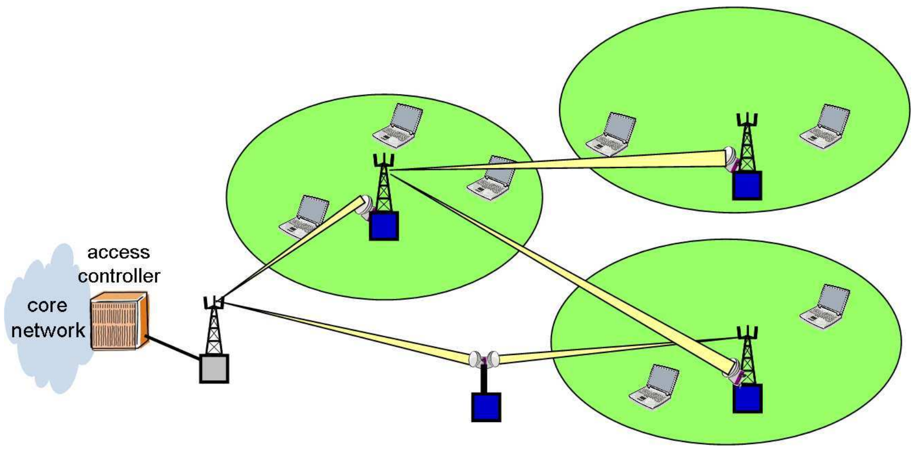

Figure 1.1: An example of a wireless backhaul of a WiMAX (IEEE 802.16) network. The base stations are connected with each other and the access point over wireless links.

Power control in wireless networks refers to an intelligent setting of the transmit power with respect to the available wireless resources. To be more precise we use the metaphor of a party where several people are talking simultaneously in different groups. If someone in one group does not hear the conversation well, the speaker will talk louder. However, at the same time the crosstalk, that is, the interference with other conversations increases. Thus other groups may have to increase their volume as well. This results in a competitive situation where all speakers try to compensate the interference by increasing their own volume. To avoid such possibly uncontrollable effects and to further ensure a certain signal quality in terms of SIR an intelligent power control should be used that incorporates these interdependencies and finds a suitable equilibrium exploiting the available power resources optimally. Note that due to the coupling of the links this requires a joint optimization of all links. Adaptive beamforming corresponds in the party metaphor to the case where a listener points his/her ears in the direction of the speaker (receive beamforming) or a speaker faces a listener directly (transmit beamforming). Further, assume that only one speaker is allowed to talk for a certain amount of time meaning that the speakers are ordered by talking one after the other. This provides an example for scheduling in different time intervals.

In the remainder of this chapter, beginning with section 1.2, we first summarize the related works on resource allocation and interference management in the area of power control, beamforming and scheduling as well as motivate the main topics and problems considered in this thesis. Then, in section 1.3 we summarize our main contributions and give an overview regarding the organization of the thesis. Finally, in section 1.4 we introduce the network and system model.

1.2 State of the Art and Related Work

The history of resource allocation and interference management dates back to the 1970s. Together with the shift to mass-market telephony the necessity occurred to manage the limited spectrum. More precisely, a frequency spectrum was assigned to the operator or owner of the wireless telephone system. Each operator had to organize the use of the spectrum assigned to his telephone system. Surely, to achieve a high profit the operator was interested to maximize the number of supported services or users. Thus, in order to organize the use of the spectrum profitably, the spectrum planning was born. In a wireless multi-cell network spectrum planning divides the spectrum in several bands and aims at reusing these bands in cells that are separated by a sufficient spatial distance, also called the frequency reuse distance. The problem of the limited spectrum was first understood as radio resource management [Zan97, ZK01]. However, to keep up with the increasing demands on wireless communications and with ongoing research on the optimization of wireless communication networks, radio resource management had to capture more dimensions. What started out with frequency planning became more and more an issue of managing available resources and at the same time tolerating interference in a controlled way by using mechanisms like power control, intelligent assignment of frequency and time, beamforming etc. That is why we refer to radio resource allocation and interference management in this thesis.

Studying and developing strategies for radio resource allocation and interference management has been the concern of many researchers and network providers over many years. There exists a great amount of literature on resource allocation from different points of view. Keep in mind that due to the interdependencies between the links a joint optimization is required. A network can be optimized with respect to certain design criteria such as feasibility of some QoS requirements in terms of minimum SIR targets, network efficiency, stability, fairness and so on. Thus an important problem in resource allocation is to find QoS points according to some of these criteria. In general, there exist two approaches which have attracted the main interest during the last years. The classical QoS-based approach aims at satisfying a certain QoS requirement with minimum power. To circumvent the feasibility problem a related approach is to solve the relative max-min or so-called max-min SIR-balancing problem. In contrast to this stands the utility-based resource allocation problem where the network operator aims at optimizing a weighted aggregate utility so as to maximize the overall network performance. Throughput, fairness, or a tradeoff between throughput and fairness can be achieved by choosing the utility function adequately. In figure 1.2 an examplary convex QoS region is depicted to illustrate which points can be achieved by the various approaches.

Classical QoS-based power control aims at allocating transmit powers to the users such that each user meets its SIR target [Wu00, BCP00, FM05]. Provided that the SIR requirements are feasible there exist iterative distributed algorithms that attain the target SIR [FM93, Yat95, BCP00]. However, to the best of our knowledge a distributed algorithm to detect the feasibility of the SIRs remains an open problem. Note, that maximizing the minimum (relative) SIR [Zan92a, Zan92b, Bam98, KG05] provides a closely related approach to the classical QoS-based power control problem.

In contrast to the classical QoS-based power control, the objective of utility-based power control is to optimize the overall network performance with respect to some aggregate utility function [GM00, SMG02, XSC03, JXB03, Chi04, Chi05, HBH05, HBH06, SWB07, SWB09, HRCW08, HSAB09b]. The utility-based approach has recently attracted a lot of attention, mainly because of a better throughput performance due to an efficient utilization of wireless resources. At the same time, the use of increasing and strictly concave utility functions ensures the desired degree of fairness [KMT98, MW00, SW05]. Some of the mentioned papers model the utility-based power control problem as a noncooperative game where users maximize their utilities [GM00, SMG02, XSC03]. In addition, these papers balance the throughput and fairness performance against power consumption. Distributed utility-based power control algorithms have been provided by [Chi05, HBH06, SWB07, HRCW08]. In [Chi05] the problem of joint power control and end-to-end congestion control is addressed where the power control part is a special case of the power control problem in

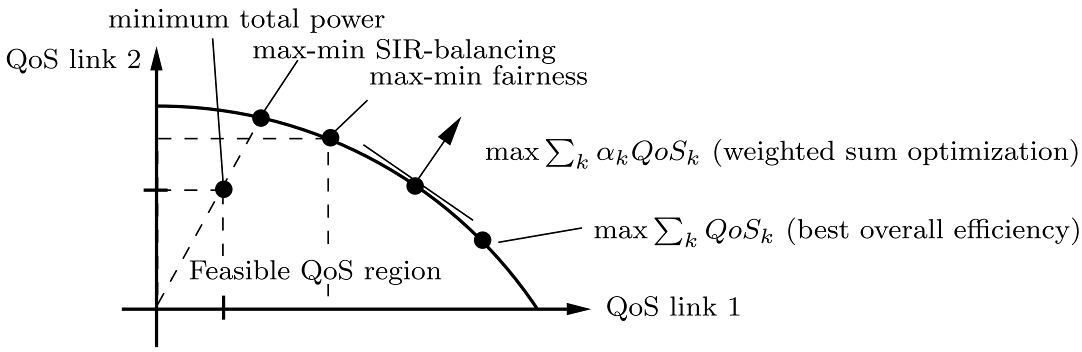

Figure 1.2: [SB06] Example of a feasible convex QoS region for 2 users. The feasible QoS region is the set of all QoS values that can be achieved by means of a certain resource allocation mechanism, e.g power control, with all links activated concurrently. The points illustrate the various approaches. In detail, the classical QoS-based approach supports, if feasible, the QoS requirements with minimum power. The closely related max-min SIR-balancing approach achieves the point at the boundary of the feasible QoS region where the line with the direction of the vector of QoS requirements intersects the boundary. The utility-based approach also achieves a point on the boundary. Here the weights determine which point is reached, because the weight vector is normal to the hyperplane that touches the boundary of the QoS region in one point. Thus by varying the weights any point on the boundary can be achieved even the max-min fair point or the max-min SIR-balancing point.

[SWB07]. The approach of [HBH06] is a game-theoretic one. Both [Chi05] and [HBH06] apply a flooding protocol to pass locally available quantities to other nodes. The authors of [HRCW08] interpret the utility-based power control problem as a joint optimization of powers and SIR assignment over the feasibility region. They propose a distributed power control and SIR assignment algorithm for the uplink in a multi-cell wireless network. In contrast, [SWB07] propose a distributed utility-based power control algorithm for wireless networks with a general network topology applying the notion of the adjoint network and thus avoiding to use a relatively expensive flooding protocol.

However, in view of many important applications, the main drawback of utility-based approaches is that, in general, QoS-requirements cannot be guaranteed even if the corresponding SIR targets are feasible. Thus, a utility-based approach with QoS support is surely desirable and highly beneficial because it kills two birds with one stone: the network provider maximizes the network efficiency and at the same time supports the needs of the customers. One possible approach to solve this problem can be found in [Chi06] where the utility function is restricted to the sum rate in the high SIR region. Another way is to enforce the SIR targets by projecting the users' transmit powers on a set of so-called valid transmit powers which provide the desired SIR performance to each user. As far as the implementation in a decentralized network is concerned, the main problem seems to be the projection operation. For instance, a very simple cyclic projection algorithm [FM05] where the users successively perform projections on individual sets corresponding to each SIR target would certainly require a lot of coordination in a network. In addition to the issue of distributed implementation, there is an inherent problem resulting from the fact of possible infeasibility. Assuming that admission control works, however, no valid power allocation may exist to meet the SIR targets due to channel and/or network dynamics. As a remedy, a resource allocation strategy should be also able to occasionally cope with the problem of infeasibility. These and further interesting and practically relevant questions, emerging from the topic of utility-based power control with QoS support, are the scope of chapter 2. Note that simultaneously and independently to our work the authors of [GMG08] imposed a maximum and a minimum SIR requirement in the utility-based power control problem and presented a Lagrangian gradient-based algorithm and its distributed implementation based on an expensive flooding protocol.

In addition to power control, beamforming provides another powerful mechanism to manage the interference in wireless networks. In order to ensure a high utilization of wireless resources, transmit powers and beamformers should be optimized jointly to exploit interdependencies between them. As a matter of fact, optimizing over the joint space of powers and beamformers is a much more intricate task than optimizing over powers only. So far almost all works on joint power control and beamforming have focused on the QoS-based approach by finding the optimum QoS point with minimum total power or by solving the related max-min SIR-balancing problem. More precisely, in [RTL98, SB04, SB06] the duality between uplink and downlink channels is exploited. Another strategy shows that the problem can be embedded in semidefinite and conic optimization programs [BO01, WES06]. Aside from that, in [BS06b] the authors consider the noiseless case and solve the max-min SIR-balancing problem by optimizing a specific aggregate utility function over the joint space of powers and beamformers. However, apart from [SWB09] the above set of publications considers only the deterministic case. First works incorporating imprecise knowledge of received waveforms include [HMV95, UY98, Var99]. Recently, stochastic algorithms for joint QoS-based power control and receive beamforming and their convergence analysis have been proposed by [LUE05, DV07]. Summarizingly, at the time we worked at this thesis the joint utility-based power control and receive beamforming problem has been still a problem uninvestigated and insights here especially for decentralized real-world wireless networks would provide a valuable extension to the pure utility-based power control framework and algorithm from [SWB07] even if a global optimum cannot be found. Besides we need to better understand the connection between the utility-based and the QoS-based resource allocation approach. This follows immediately from the fact that the max-min SIR-balancing objective function is not differentiable. As a consequence gradient-projection methods and its existing distributed implementation cannot be applied. A possible solution is to identify the connection between both problems and then to solve the equivalent utility-based resource allocation problem. Encouragingly, recent results [SKB10] have revealed the connection between the max-min SIR-balancing and the utility-based power control problem under general power constraints. Thus the question arises whether both works [BS06b] and [SKB10] can be combined to find a solution for general power constraints and a larger class of utility functions and thus enable us to solve the max-min SIR-balancing problem distributedly using an indirect way. Summarizingly, the topic resource allocation in multiantenna wireless networks, specifically the mentioned interdependencies between powers and beamformers will be discussed in chapter 3. Independently from and simultaneously to our work the authors of [LHC07] have proposed a distributed utility-based joint power control and receive beamforming algorithm for cellular uplinks applying the scheme of [HRCW08].

Scheduling is another powerful mechanism in the area of resource allocation and interference management. It facilitates orthogonalizing the users either a) by using a time-division multiplexed (TDM) system where a user is assigned to a single time slot (scheduling interval) or b) by using a combination of TDM and another multiplexing technique such as code division multiple access (CDMA) or orthogonal frequency division multiple access (OFDMA). In the latter case we have to decide both which users to schedule and how to divide the bandwidth in terms of signature sequences or subchannels among the scheduled users. Assuming that only one user is scheduled at the same time and for the same bandwidth resource, we arrive at a system with orthogonal channels. Several channel-aware scheduling algorithms exist for a system like a) that schedules a single user in each scheduling interval [AKR+01, SS01]. First works on bandwidth allocation in OFDMA/CDMA networks like in b) combined with power control include [WCLM99, KL00, KLKL01] among others. These mainly focus on minimizing the total transmit power given target bit rates for each user and give the impression that exploiting frequency diversity is quite a challenging task. Further related works consider the problem of maximizing the weighted sum capacity for the downlink channel from an information-theoretic point of view [Tse97, LG01] assuming multi-user coding/decoding. However, in the special case of maximizing the equal weight capacity the information-theoretic optimal approach is to transmit only one user in each time slot. These works and discussions motivate us to optimize powers and bandwidth allocation jointly in a wireless network with orthogonal channels and general power constraints such that the weighted sum rate is maximized. Note that the bandwidth allocation is usually constrained to a set of discrete or integer points. To sum it up, resource allocation in networks with orthogonal channels with an emphasis on sum-rate maximization will be discussed in chapter 4. Simultaneously to and independently from our work over the last years further algorithms have been developed that first dealt with the downlink problem, then extended the problem to the uplink and finally have been also generalized to networks with general power constraints [KV05, HSAB09b, HSAB09a, ASB10, HSBA10] (and references therein). These recent and newer publications partly take into account frequency diversity and several constraints on per user power and rate constraints.

1.3 Contributions and Thesis Overview

This thesis is dedicated to topics in the area of resource allocation and interference management in wireless networks of a general network topology that require a distributed or decentralized organization. More precisely, in chapters 2 and 3 we investigate the problems of power control as well as joint power control and beamforming assuming a wireless network where interference is treated as noise. The key contributions are two-fold. They do not restrict to the development of applicable distributed algorithms that can be combined with any routing and scheduling strategy. We also derive and prove practically interesting and significant properties of the approaches and algorithms.

Chapter 2: In order to deal with the drawback that QoS requirements cannot be guaranteed by the traditional utility-based approach, we incorporate QoS requirements in terms of minimum SIR targets into the utility-based power control problem. We first ensure the desired SIR targets by projecting the transmit powers on the valid power region. This approach is called hard QoS support, simply because the SIR targets are only met when being feasible. Because the projection operation on the feasible set of powers cannot be implemented in a distributed way or would require a prohibitive amount of data exchange we apply a primal-dual approach to find a stationary point of the associated Lagrangian. We show both that the derived corresponding primal-dual algorithm converges to the optimum and that it is amenable to efficient distributed implementation in case of feasibility. However, a challenge with hard QoS support occurs if it is impossible to enforce the desired SIR levels. For this reason we argue in favor of soft QoS support in which case a solution to the corresponding power control problem always exists even if the SIR targets are infeasible. We achieve soft QoS support by approxima-ting the max-min SIR-balancing solution which is a a straightforward result applying the argumentation of [MW00] to power-controlled wireless networks. The approximation of the max-min SIR balancing problem can be seen as a barrier function. We combine this barrier function with the traditional utility-based approach to include best-effort users and prove the following interesting properties of this approach:

(a) If the SIR targets are infeasible so that they cannot be met simultaneously, then the maximum relative error is minimized, provided that the concavity parameter of the utility function chosen is sufficiently large. If all SIR targets are equal, then the distance to the SIR targets is minimized with respect to the l_∞-norm (the maximum norm).

(b) The SIR targets are met provided that the concavity parameter of the utility function is sufficiently large and the targets are feasible. The irreducibility condition of the channel gain matrix is a sufficient condition to ensure that the achieved SIR has an upper bound and does not exceed the target SIR.

(c) After all SIR targets are satisfied remaining power resources are allocated to the users so as to optimize the aggregate utility function.

Finally we show that our extended gradient-projection algorithm can be efficiently implemented in a distributed manner using the scheme of [SWB07] and give some numerical examples to show the influence of noisy measurements.

Chapter 3: This chapter captures the possibility of jointly optimizing powers and beamformers. Such a joint optimization is much more challenging than a pure power control problem only. To the best of our knowledge it is not known which class of utility functions allows a convex formulation of the joint utility-based power control and receive beamforming problem so that a global solution can be found in distributed wireless networks. We analyze the structure of the problem and show that the joint problem can be formulated as a pure power control problem under inherent optimal receivers. More precisely, for any fixed power vector, if the optimal receive beamformer is chosen in the sense that it maximizes the SINR it solves at the same time the original utility-based problem. For the case of perfect synchronization where a closed-form solution of the optimal receiver exists we derive a gradient projection power control algorithm that cannot be implemented distributedly in an efficient way. Due to this and because synchronicity is not a property of decentralized wireless networks we decompose the problem into two coupled sub-problems and propose a decentralized alternating computation that converges to a stationary point. Simulations suggest that with increasing concavity of the utility function the algorithm converges to a global optimum for many initial points. But in fact we cannot even show that our proposed algorithm converges to a local maximum. Nevertheless, we provide a powerful distributed algorithm and investigate its behaviour in real-world wireless networks where noisy measurements and estimations occur.

The second part of chapter 3 considers the possibility of solving the max-min SIR-balancing problem. Because a gradient-projection algorithm cannot be applied in case of the non-differentiable relative max-min objective function we propose a different way to arrive at a possibly distributed solution. For this purpose we extend the work of [SB06] by solving the joint max-min SIR-balancing problem under general power constraints and for a larger class of utility functions. More precisely, we apply recent results of [SKB10] that reveal the connection between the max-min SIR-balancing and the utility-based power control problem under general power constraints. These results show how to choose the weights such that the utility-based power control problem achieves the max-min SIR-balancing solution. We analyze the properties of this connection and combine these results with receiver control to arrive at an algorithmic solution. We conjecture that the proposed algorithm converges to the optimal max-min SIR-balancing solution. Interestingly, we would be able to even distributedly solve the joint max-min SIR-balancing problem provided that the corresponding weights could be calculated in a distributed fashion. Unfortunately, the weights are not so easily calculated in a decentralized way, some coordination is still required.

Summarizingly, this chapter provides applicable distributed and semi-distributed algorithms to jointly optimize powers and receive beamformers. It provides also a distributed heuristic to additionally optimize the transmit filters and indicates by simulations that by this means an additional performance gain can be achieved.

Chapter 4: While chapters 2 and 3 are concerned with networks with interference, in this chapter we consider a wireless network with orthogonal channels to separate the users. Orthogonal channels are achieved by applying CDMA or OFDMA (with flat fading channels). As a consequence the problem of joint power control and link scheduling, such that a weighted sum of link rates is maximized, greatly simplifies. On the other hand we have scheduling variables that are usually discrete. We emphasize that in contrast to other simultaneous works we focus on networks with general power constraints. First, relaxing the discrete scheduling variables, we show that the problem is jointly convex and derive an optimal iterative solution which provides an upper bound to the original discrete problem. The optimal real solution can be simply rounded to result heuristically in a discrete solution. However, a more promising way is to analyze the problem structure first and then to derive a convenient heuristic based on the problem properties. Interestingly, we can show that the decision of how many subchannels or signature sequences are assigned to each node can be interpreted as a multiple-choice knapsack problem. Using the knowledge of the problem structure we derive a heuristic centralized algorithm for the problem and indicate a possible distributed implementation. Moreover, we emphasize that the approach of weighted throughput maximization offers throughput-optimality if the weights are chosen accordingly and that the choice of the weights highly influences the delay performance as well.

1.4 System Model

1.4.1 Notation

Throughout the text, $\mathbb{R}$ denotes the set of real numbers and $\mathbb{C}$ the set of complex numbers. $\mathbb{R}_+ \subset \mathbb{R}$ and $\mathbb{R}_{++} \subset \mathbb{R}$ are used to designate the set of non-negative and positive real numbers, respectively. The corresponding sets of all N-tuples will be denoted by $\mathbb{R}^N$, $\mathbb{C}^N$, $\mathbb{R}^N_+$ and $\mathbb{R}^N_{++}$, respectively. We use small-type bold letters $\mathbf{a} = (a_1, \dots, a_N)$ to represent a vector and capital bold letters $\mathbf{A} = (a_{k,l})$ to represent a matrix. Further $\mathbf{a}^T$ and $\mathbf{a}^H$ denote the transpose and conjugate transpose of vector $\mathbf{a}$, respectively.

1.4.2 Network Model

A wireless network is a collection of nodes being able to communicate with each other over a wireless communication link. Let $\mathcal{N} = \{1, \ldots, N\}$ be the set of nodes and let (n, m) with $m \neq n$ represent a wireless link from transmitter node $n \in \mathcal{N}$ to receiver node m. We assume that wireless links are bidirectional meaning that (m, n) exists if and only if there exists (n, m). The set of nodes and the set of wireless links constitute the network topology. We consider a wireless network with an established network topology, in which all links share a common wireless spectrum.

Now, let $K \geq 2$ users compete for access to the wireless links and let $\mathcal{K} = \{1, \ldots, K\}$ denote the index set of all users. The set of users originating in node $n \in \mathcal{N}$ is labeled by $\mathcal{K}_n \subseteq \mathcal{K}$ [1]. In the case that only one user is competing for access on a certain wireless link, the notion of users and active wireless links is the same. In contrast, if more than one user competes for access on a wireless link we have several so-called logical links on a wireless link, one for each user. We point out, that the notion of logical links is only used in this thesis in the case where we consider a multi-hop wireless network with several source-destination pairs and a fixed, established route. Then, several routes or paths may use the same wireless link, in which case we establish a number of logical links on each wireless link.

1.4.3 Signal-to-Interference-Ratio

The transmit powers of the users are denoted by $p_k \geq 0$ and collected in the vector $\mathbf{p} = (p_1, \ldots, p_K)$, called the power vector. We define $V_{k,l} \geq 0$ with $V_{k,k} > 0$ to be the path attenuation between transmitter $l \in \mathcal{K}$ and receiver $k \in \mathcal{K}$ (see figure 1.3) which depends on the bandwidth and spectrum allocation, the channel state of the wireless link[2], the receiver structure and other system parameters. Further, $\sigma_k^2 > 0$ denotes the Gaussian noise variance at the kth receiver.

The main figure of merit is the SIR at each receiver output given by

$$\mathrm{SIR}_k(\mathbf{p}) := \frac{V_{k,k}\, p_k}{\sum_{\substack{l=1\\ l\neq k}}^{K} p_l V_{k,l} + \sigma_k^2} = \frac{p_k}{\sum_{\substack{l=1\\ l\neq k}}^{K} p_l \frac{V_{k,l}}{V_{k,k}} + \frac{\sigma_k^2}{V_{k,k}}},\ k \in \mathcal{K}\,. \tag{1.1}$$

Moving $V_{k,k}$ into the denominator, we define $v_{k,l} = V_{k,l}/V_{k,k}$ if $l \neq k$. Here and hereafter, we assume no self-interference meaning that $v_{k,l} = 0$ if $l = k$. Further, we define $z_k = \sigma_k^2/V_{k,k} > 0$ to be the effective noise power at the receiver k. Based on this, we define $\mathbf{z} := (z_1, \ldots, z_K) > 0$ and $\mathbf{V} := (v_{k,l}) \in \mathbb{R}_+^{K\times K}$ with $\mathrm{trace}(\mathbf{V}) = 0$ to be the vector of effective noise variances and the matrix of power gains, respectively. The nonnegative

[1] Later, in subsection 1.4.4 defining the power constraints, this set is transformed into a Matrix $\mathbf{A}$ mapping the users and the power constraints

[2] The channel state of a wireless link depends on the characteristics of the radio channel, e.g. path loss.

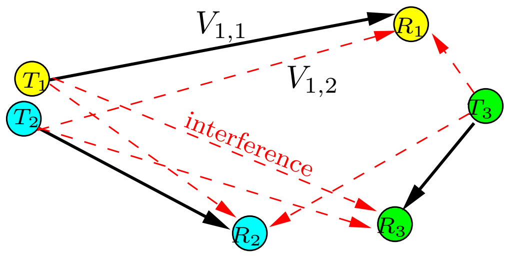

Figure 1.3: An example of a network with 6 nodes and 3 active links or users. Each link consists of a transmitter node T_1, T_2, T_3 and a receiver node R_1, R_2, R_3. By transmitter (receiver) $l \in \mathcal{K}$ we refer to the transmit (receive) node of link or user l. The path attenuation between different transmitters and receivers is illustrated as follows: The transmission links are depicted by black lines, the interference links are depicted by dashed red lines.

matrix $\mathbf{V}$ is called the gain matrix. Unless otherwise stated, it is assumed that $\mathbf{V}$ is irreducible, in which case the network is entirely coupled by interference. In practice, this is a reasonable assumption and means that there is no isolated subnetwork perceiving no interference from links outside the subnetwork.

Using this notation the SIR can be also written as follows

$$\mathrm{SIR}_k(\mathbf{p}) = \frac{p_k}{I_k(\mathbf{p})} = \frac{p_k}{(\mathbf{V}\mathbf{p} + \mathbf{z})_k}, k \in \mathcal{K}, \tag{1.2}$$

where

$$I_k(\mathbf{p}) = \sum_{l \in \mathcal{K}} v_{k,l} p_l + z_k \tag{1.3}$$

is the interference function.

1.4.4 Power Constraints

In wireless networks with a general network topology the transmit powers are subject to *general* power constraints

$$\mathrm{P} = \{\mathbf{p} \in \mathbb{R}_+^K : \mathbf{A}\mathbf{p} \le \mathbf{p}', \mathbf{A} \in \{0,1\}^{N \times K}\} \subset \mathbb{R}_+^K \tag{1.4}$$

for some given $\mathbf{p}' = (P_1, \dots, P_N) > 0$. We use $\mathcal{N} = \{1, \dots, N\}$ to denote the set of power constraints, one for each active transmit node, where N is the number of power constraints. The mapping between the users/active links and the power constraints is represented by $\mathbf{A} = (a_{n,k}) \in \{0,1\}^{N \times K}$: $a_{n,k} = 1$ if user k belongs to power constraint n or, in other words, if active link k originates in node n. In practice, we may have several links originating at node n such that we have a sum power constraint at node n. Note, there is at least one 1 in each column of $\mathbf{A}$. Thus P is a compact set.

With this model the subsequent two special cases can be represented as follows:

- *Individual* power constraints on each link: This scenario occurs in practice in case that each active node is the origin for exactly one link. We have $N = K$, $\mathbf{A} = \mathbf{I}$ and may define
$$\mathrm{P} := \{\mathbf{p} \in \mathbb{R}_+^K : \forall_{k\in\mathcal{K}} p_k \leq P_k\}. \tag{1.5}$$
A typical example is the uplink channel in a single-cell wireless network.

- *Sum* power constraint: We have only one node $N = 1$, where each active link originates. Thus $\mathbf{A}$ reduces to a vector $\mathbf{A} = \mathbf{1}^T$ and we have $\mathrm{P} := \{\mathbf{p} \in \mathbb{R}_+^K : \sum_{k\in\mathcal{K}} p_k \leq P\}$. A well-known example for such a network is the downlink of a single-cell wireless network.

If not otherwise stated, for brevity and simplicity in presentation, we assume individual power constraints throughout this thesis.

1.4.5 Linear Receiver

As emphasized, the value $\mathrm{SIR}_k(\mathbf{p})$ is not only a function of the power vector $\mathbf{p}$ but depends also on the choice of the transmit and the receive filters. Note that all transmitters but only the kth receiver influence the kth SIR.

Let $k \in \mathcal{K}$ be arbitrary but fixed. We assume that all users are equipped with $M \geq 1$ antennas and define $\mathbf{b}_l := \mathbf{b}_l^{(k)} = (b_{1,l}^{(k)}, \dots, b_{M,l}^{(k)}) \in \mathbb{C}^M$ to be the effective transmit vector of transmitter $l \in \mathcal{K}$ associated with receiver $k \in \mathcal{K}$. The effective transmit vector $\mathbf{b}_l$ is the product of the channel matrix between transmitter $l \in \mathcal{K}$ and receiver $k \in \mathcal{K}$ and its transmit beamformer. It determines the direction of the transmit signal. We give an example for a wireless network with M antennas at both the transmitter and the receiver side. Here, $\mathbf{H}^{(k,l)} \in \mathbb{C}^{M\times M}$ describes the channel from the transmitter of user l to the receiver of user k. Further, defining $\mathbf{u}_l$ with $||\mathbf{u}_l||_2 = 1$ to be the transmit beamformer at the transmitter-side of link l, we arrive at $\mathbf{b}_l^{(k)} = \mathbf{H}^{k,l}\mathbf{u}_l$ describing the effective transmit beamformer of transmitter l into the direction of receiver k.

If not otherwise stated the effective transmit vector is assumed to be arbitrary but fixed, which implies that the channels and transmit beamformers are fixed. In addition, we use $\mathbf{c}_k \in \mathbb{C}^M$ to denote the receive beamformer of user k. The receive beamformers of all users are collected in the receive beamforming matrix $\mathbf{C} = (\mathbf{c}_1, \dots, \mathbf{c}_K) \in \mathbb{C}^{M\times K}$. Furthermore, since the signal-to-interference ratio (SIR) is independent of the norm of the receive beamformers, we can assume that $\|\mathbf{c}_k\|_2 = 1$ for each $1 \leq k \leq K$, and hence $\mathrm{C} = \{\mathbf{C} = (\mathbf{c}_1, \dots, \mathbf{c}_K) \in \mathbb{C}^{M\times K} : \forall_k \|\mathbf{c}_k\|_2 = 1\}$ denotes the set of all beamforming matrices.

Using the above notation, assuming that the effective transmitters are fixed and considering the fact that all users are perfectly synchronized, the SIR at the output of an arbitrary linear receiver is given by

$$\begin{aligned}\mathrm{SIR}_k(\mathbf{p},\mathbf{c}_k) :&= \frac{p_k|\langle \mathbf{c}_k,\mathbf{b}_k\rangle|^2}{\sum_{\substack{l=1\\l\neq k}}^K p_l|\langle \mathbf{c}_k,\mathbf{b}_l\rangle|^2 + ||\mathbf{c}_k||_2^2\sigma_k^2} = \frac{p_k V_{k,k}(\mathbf{c}_k)}{\sum_{\substack{l=1\\l\neq k}}^K p_l V_{k,l}(\mathbf{c}_k)+\sigma_k^2}\\ &= \frac{p_k}{\sum_{\substack{l=1\\l\neq k}}^K p_l \frac{V_{k,l}(\mathbf{c}_k)}{V_{k,k}(\mathbf{c}_k)} + \frac{\sigma_k^2}{V_{k,k}(\mathbf{c}_k)}}\end{aligned} \tag{1.6}$$

where $V_{k,l}(\mathbf{c}_k) = |\langle \mathbf{c}_k,\mathbf{b}_l\rangle|^2$ is the attenuation of the power from the transmitter of user l to the receiver of user k and $\langle \mathbf{x},\mathbf{y}\rangle$ denotes the inner product of the vectors $\mathbf{x},\mathbf{y}$. As before the SIR can be also written as follows

$$\mathrm{SIR}_k(\mathbf{p},\mathbf{c}_k) = \frac{p_k}{I_k(\mathbf{p},\mathbf{c}_k)} = \frac{p_k}{(\mathbf{V}(\mathbf{C})\mathbf{p}+\mathbf{z}(\mathbf{C}))_k}, k\in\mathcal{K}, \tag{1.7}$$

where

$$I_k(\mathbf{p},\mathbf{c}_k) = \sum_{l\in\mathcal{K}} v_{k,l}(\mathbf{c}_k)p_l + z_k(\mathbf{c}_k) \tag{1.8}$$

is the interference function. Notice, under the assumption of a fixed receiver $\mathbf{c}_k \in \mathbb{C}^M$ we have $\mathrm{SIR}_k(\mathbf{p}) = \mathrm{SIR}_k(\mathbf{p},\mathbf{c}_k), k\in\mathcal{K}$.

Optimal Linear Receiver

By an optimum linear receiver usually a receiver is meant that maximizes the SIR of each user. More precisely, the optimum receive vector $\mathbf{c}_k$ is a vector that maximizes the SIR_k under the assumption that the power vector $\mathbf{p}$ and the effective transmit vectors $\mathbf{b}_l := \mathbf{b}_l^{(k)}, \forall l\in\mathcal{K}$ are fixed. Hence, rewriting (1.6) as

$$\mathrm{SIR}(\mathbf{p},\mathbf{c}_k) = \frac{p_k\mathbf{c}_k^H\mathbf{b}_k\mathbf{b}_k^H\mathbf{c}_k}{\mathbf{c}_k^H\mathbf{Z}_k(\mathbf{p})\mathbf{c}_k} \tag{1.9}$$

with

$$\mathbf{Z}_k(\mathbf{p}) = \sum_{l\neq k} p_l\mathbf{b}_l\mathbf{b}_l^H + \sigma^2\mathbf{I} \tag{1.10}$$

we have

$$\mathbf{c}_k^*(\mathbf{p}) = \arg\max_{\|\mathbf{c}_k\|_2=1} \frac{p_k\mathbf{c}_k^H\mathbf{b}_k\mathbf{b}_k^H\mathbf{c}_k}{\mathbf{c}_k^H\mathbf{Z}_k(\mathbf{p})\mathbf{c}_k}.$$

Since σ^2 is positive, the matrix $\mathbf{Z}_k(\mathbf{p})$ is Hermitian positive definite for any $\mathbf{p}\geq 0$. As a consequence, the inverse matrix of $\mathbf{Z}_k(\mathbf{p})$ exists regardless of the choice of receivers $\mathbf{C}$ and powers $\mathbf{p}\in \mathrm{P}$. Note that the SIR can be written in this compact form due to the assumption of perfect synchronization. An optimal receive beamformer $\mathbf{c}_k^*(\mathbf{p})$ can be easily found when the SIR_k is rewritten as a Rayleigh quotient to obtain [VAT99]

$$\mathbf{c}_k^*(\mathbf{p}) = a_k \mathbf{Z}_k^{-1}(\mathbf{p})\mathbf{b}_k \tag{1.11}$$

where $a_k > 0$ is a constant chosen such that $\|\mathbf{c}_k^*(\mathbf{p})\|_2 = 1$. Consequently, with an optimal beamformer, the SIR of user k is equal to

$$\mathrm{SIR}_k^*(\mathbf{p}) = p_k \mathbf{b}_k^H \mathbf{Z}_k^{-1}(\mathbf{p})\mathbf{b}_k. \tag{1.12}$$

1.4.6 Applied System Models

For a better understanding we shortly summarize the specific system models used in each of the following 3 chapters.

Chapter 2 considers a pure power control problem in a wireless network where interference is treated as noise. Thus, the defined SIR depends on the power vector only (see equation (1.2)). Further we derive solutions and distributed algorithms assuming individual power constraints for each user (see equation (1.5)). However, we point out that these solutions and distributed algorithms can be easily generalized to the case of general power constraints.

In chapter 3 we discuss a joint power control and beamforming problem in a wireless network where interference is treated as noise. Consequently, the SIR is a function of powers, transmit vectors and receive beamformers (see equation (1.6)). We emphasize that only one data stream is transmitted over each link, hence the beamformers indicate a signal direction. Note that the main focus is on the optimization of receive beamformers. That is why we introduced the optimal linear receiver (see equation (1.11)). Moreover, in the first part of chapter 3, for brevity and simplicity we again focus on networks with individual power constraints. Here as well as in chapter 2 the solutions can be easily adapted to the case of general power constraints. In the second part of chapter 3 we extend existing works to the case of general power constraints. Thus, here general power constraints are assumed (see equation (1.4)).

Finally in chapter 4 we consider a joint power control and scheduling problem in a network with orthogonal channels. As a consequence the SIR model defined in equation (1.2) is simplified, $V_{k,l} = 0$, $\forall_{k,l}$. Thus $\mathbf{V} := (v_{k,l})$ with $v_{k,l} = V_{k,l}/V_{k,k}$ is a zero matrix, $\mathbf{z}$ remains equivalent representing the vector of effective noise variances. Due to the fact that scheduling is a topic only in chapter 4, we directly introduce the scheduling variable in the SIR model at the beginning of chapter 4. Moreover, the wireless network of chapter 4 is investigated assuming a general network topology. Thus, again general power constraints (see equation (1.4)) are assumed.

2 Utility-based Power Control with QoS Support

During the last two decades, we have observed the evolution of a number of different wireless communication standards that support a wide range of services. These services have requirements on quality-of-service (QoS) parameters such as data rate, delay, bit error rate. The QoS is determined by the signal-to-interference ratio (SIR) and may be expressed in terms of some minimum SIR targets, which is a reasonable assumption in many cases of practical interest [TH99, BG92]. One of the most important mechanisms to manage interference in a wireless network is power control. This chapter addresses the question, how such QoS or SIR requirements may be supported using a utility-based power control scheme in a distributed wireless network.

Having a look at related works we know that classical QoS-based power control aims at allocating transmit powers to the users such that each user meets its SIR target [Wu00, BCP00, FM05]. Provided that the SIR requirements are feasible there exist iterative distributed algorithms that attain the target SIR [FM93, Yat95, BCP00]. However, to the best of our knowledge a distributed algorithm to detect the feasibility of the SIRs remains an open problem. Note, that maximizing the minimum SIR [Zan92a, Zan92b, Bam98, KG05] provides a closely related approach to the classical QoS-based power control problem.

In contrast to the classical QoS-based power control, the objective of utility-based power control is to optimize the overall network performance with respect to some aggregate utility function [GM00, SMG02, XSC03, JXB03, Chi04, Chi05, HBH05, HBH06, SWB07, SWB09, HRCW08, HSAB09b]. The utility-based approach has recently attracted a lot of attention, mainly because of a better throughput performance due to an efficient utilization of wireless resources. At the same time, the use of increasing and strictly concave utility functions ensures the desired degree of fairness [KMT98, MW00, SW05]. Some of the mentioned papers model the utility-based power control problem as a noncooperative game where users maximize their utilities (e.g. [GM00, SMG02, XSC03]). In addition, these papers balance the throughput and fairness performance against power consumption. Distributed utility-based power control algorithms have been provided by [Chi05, HBH06, SWB07, HRCW08]. In [Chi05] the problem of joint power control and end-to-end congestion control is addressed where the power control part is a special case of the power control problem in [SWB07]. The approach of [HBH06] is a game-theoretic one. Both [Chi05] and [HBH06] apply a flooding protocol to pass locally available quantities to other nodes. The authors of [HRCW08] interpret the utility-based power control problem as a joint optimization of powers and SIR assignment over the feasibility region. They propose a distributed power control and SIR assignment algorithm for the uplink in a multi-cell wireless network. In contrast, [SWB07] propose a distributed utility-based power control algorithm for general wireless networks applying the notion of the adjoint network and thus avoiding to use a relatively expensive flooding protocol.

However, in view of many important applications, the main drawback of utility-based approaches is that, in general, no QoS can be guaranteed even if the corresponding SIR targets are feasible. One possible approach to solve this problem is to enforce the SIR targets by projecting the users' transmit powers on a set of so-called valid transmit powers that provide the desired SIR performance to each user. This is called hard QoS support as the SIR targets are enforced whenever they are feasible. As far as the implementation in a decentralized network is concerned, the main problem seems to be the projection operation. For instance, a very simple cyclic projection algorithm [FM05] where the users successively perform projections on individual sets corresponding to each SIR target may require a lot of coordination in a network. Another way to manage the SIR requirements within a utility-based approach can be found in [Chi06, GMG08]. In [Chi06] the utility function is restricted to the sum rate in the high SIR region. The authors of [GMG08] impose a maximum and a minimum SIR requirement and present a Lagrangian gradient-based algorithm and its distributed implementation based on a flooding protocol. In this chapter we analyze a similar primal-dual algorithm to find a stationary point of the associated Lagrangian and show that this algorithm is amenable to distributed implementation.

In addition to the issue of distributed implementation, there is an inherent problem associated with hard QoS support resulting from the fact that no valid power allocation may exist to meet the SIR targets due to channel and/or network dynamics. For this reason, we argue in favor of soft QoS support in which case a solution to the corresponding power control problem always exists even if the SIR targets are not feasible. To be more precise, the proposed scheme attempts to approach the SIR targets closely, provided that the utility functions are chosen appropriately. Moreover, if the QoS requirements are feasible, once the SIR targets are met, remaining power resources may be allocated to all or some users so as to optimize the overall network performance expressed in terms of some aggregate utility function. As far as we know our scheme is the first to incorporate QoS constraints in this way.

Potential applications of the power control scheme presented in this chapter are envisaged for example in wireless mesh networks to control transmit powers of base stations (mesh routers). These base stations create a wireless backbone via multi-hop ad-hoc networking and have practically unlimited energy supply.

In this chapter, we first review the utility-based power control problem without QoS support in section 2.2 and show how a simple hop-by-hop congestion control protocol can be coupled with power control (section 2.2.1) such that end-to-end fairness is achieved. We emphasize that the objective of controlling the network congestion provides a motivation to investigate the utility-based power control problem. In the main part of this chapter we reformulate the conventional utility-based power control problem so as to take into account given minimum SIR targets of the users. Most importantly, the power control schemes and algorithms derived incorporate the demand on amenability to distributed implementation. Section 2.3 is devoted to the utility-based power control with hard QoS support. We analyze a primal-dual algorithm that is used to find a stationary point of an associated linear Lagrangian function. For a large class of utility functions identified in [MW00, SWB07], the algorithm is shown to exhibit a global convergence. Furthermore, it is shown that the use of an adjoint network introduced in [SWB07] allows for an efficient decentralized implementation of the algorithm. In section 2.4, we address the problem of utility-based power control with soft QoS support. This scheme circumvents the feasibility problem. First, for the sake of completeness it is mentioned that a widely-studied max-min SIR balancing solution [BS06a] can be approximated arbitrarily closely by a solution to a slightly modified utility-based power control problem, with a class of utility functions considered, for instance, in [MW00]. Then we combine this approach with the conventional utility-based power control problem in order to incorporate best-effort users. The main results (Proposition 2 with Corollary 2 and Proposition 3) state that under a solution to this problem, the following is true.

(a) If the SIR targets are infeasible so that they cannot be met simultaneously, then the maximum relative error is minimized, provided that the concavity parameter of the utility function chosen is sufficiently large. If all SIR targets are equal, then the distance to the SIR targets is minimized with respect to the l_∞-norm (the maximum norm).

(b) The SIR targets are met provided that some parameter is sufficiently large and the targets are feasible.

(c) After all SIR targets are satisfied remaining transmit powers are allocated to the users so as to optimize a given aggregate utility function.

Section 2.4.3 shows that the problem can be solved in a distributed manner, again applying the notion of the adjoint network from [SWB07]. Finally, in Section 2.5, we present some numerical results to better illustrate the proposed power control concepts. The results presented in this chapter have been published in [2, 9, 8, 11]. Parts of these works also appeared in [SWB09].

2.1 System Model, Assumptions and Definitions

We use the system model introduced in section 1.4. Considering a pure power control problem the SIR is given by equation (1.2) depending only on the power vector $\mathbf{p}$. For brevity, in this chapter we assume individual power constraints[1] as defined in equation (1.5).

Let $\omega_k \in \mathrm{Q} \subseteq \mathbb{R}$ represent the QoS requirement of user k. Interesting QoS parameters may be for example data rate, bit error rate or delay. Depending on the QoS parameter either larger values are desired (data rate) or smaller values imply a larger QoS (delay). Throughout this chapter, it is assumed that $\gamma : \mathrm{Q} \to \mathbb{R}_{++}, \gamma_k := \gamma(\omega_k)$ is the minimum SIR being necessary to provide ω_k to user k. We refer to $\gamma_1, \ldots, \gamma_K \geq 0$ as the SIR targets. If $\gamma_k = 0, k \in \mathcal{K}$, then user k has no QoS requirements and is called a best-effort user. Otherwise, we call it a QoS user. The QoS vector $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_K) \in \mathrm{Q}^K$ is said to be feasible if there is a power vector $\mathbf{p} \in \mathrm{P}$ such that $\forall_{k \in \mathcal{K}} \mathrm{SIR}_k(\mathbf{p}) \geq \gamma_k = \gamma(\omega_k)$. The set of all feasible QoS vectors is called the feasible QoS region.

We point out that since best-effort users have no QoS requirements, they could be denied access to the channel under some power control strategies such as those of [BCP00]. In the thesis, however, we consider power control strategies under which each user is allocated a positive transmit power.[2] As a result, we can focus on positive transmit powers $\mathbf{p} \in \mathrm{P}_+$ with $\mathrm{P}_+ := \mathrm{P} \cap \mathbb{R}_{++}$ and consider all the problems in the logarithmic power domain. To this end, let $\mathbf{s} := \log(\mathbf{p}), \mathbf{p} > 0$, denote the (logarithmic) power vector and $\mathrm{S} := \{\mathbf{s} \in \mathbb{R}^K : \forall_k e^{s_k} \leq P_k\} \neq \emptyset$. Notice that $\log(\mathbf{p})$ and $e^{\mathbf{s}}$ are taken component-wise and we may sometimes write $\exp(\mathbf{s})$ instead of $e^{\mathbf{s}}$.

Definition 1 *Given $\boldsymbol{\omega} \in \mathrm{Q}^K$, we say that $\mathbf{s} \in \mathbb{R}^K$ is a valid power vector if $\mathrm{SIR}_k(e^{\mathbf{s}}) \geq \gamma_k, k \in \mathcal{K}$. Any valid power vector $\mathbf{s} \in \mathbb{R}^K$ such that $\mathbf{s} \in \mathrm{S}$ is said to be feasible. The set*

$$\mathrm{S}(\boldsymbol{\omega}) := \{\mathbf{s} \in \mathrm{S} : \forall_{k \in \mathcal{K}} \mathrm{SIR}_k(e^{\mathbf{s}}) \geq \gamma_k\} \subset \mathrm{S}$$

is called the feasible power region as $\mathbf{s} \in \mathrm{S}$.

In words, the feasible power region is the set of those power vectors for which all SIR targets are met and the power constraints are fulfilled. By the above definition, we can conclude that

Observation 1 *$\boldsymbol{\omega} \in \mathrm{Q}^K$ is feasible if $\mathrm{S}(\boldsymbol{\omega}) \neq \emptyset$.*

Observation 2 *For any $\boldsymbol{\omega} \in \mathrm{Q}^K$, $\mathrm{S}(\boldsymbol{\omega})$ is a convex set.*

[1] We emphasize that the proposed solutions and distributed algorithms can be easily adapted to sum power or general power constraints.

[2] The only exception is Proposition 2 where one of the proposed strategies is shown to approximate a power vector with zero entries.

2.2 Utility-based Power Control without QoS Support

At the beginning of this section we briefly review the traditional utility-based power control problem (see [SWB07]) that deals with a network of best effort users only. In other words no QoS or minimum SIR requirements are incorporated in the optimization problem. This section provides the fundamentals for the main part of this chapter. We emphasize that although utility-based optimization problems are typically formulated as maximization problems, in the following we rewrite the utility-based power control problem as a minimization problem to be conform with standard results from the optimization theory. We introduce a class of monotonically decreasing strictly convex utility functions of the SIR for which the power control problem can be converted into a convex optimization problem. We summarize the gradient projection algorithm solving this convex problem and outline its amenability to distributed implementation by using the concept of the adjoint network. Giving some interesting additional aspects on cross-layer design, in subsection 2.2.1 we show how a simple rate control can be performed by coupling hop-by-hop congestion control with utility-based power control.

Explicitly, the traditional utility-based power control problem is

$$\mathbf{s}^* = \arg\max_{\mathbf{s}\in\mathrm{S}} G_e(\mathbf{s}) \tag{2.1}$$

where the maximum is assumed to exist and $G_e : \mathbb{R}^K \to \mathbb{R}$ is the aggregate utility function defined to be

$$G_e(\mathbf{s}) := \sum\nolimits_{k\in\mathcal{K}} w_k \Psi(\mathrm{SIR}_k(e^{\mathbf{s}}))\,. \tag{2.2}$$

Here and hereafter, $\mathbf{w} = (w_1,\dots,w_K) > 0$ is a given weight vector and $\Psi : \mathbb{R}_{++} \to \mathrm{Q}$ is the inverse function of γ so that $\gamma(\Psi(x)) = x, x > 0$. Consequently, $\Psi(\mathrm{SIR}_k(\mathbf{p}))$ is the QoS level of user $k \in \mathcal{K}$ under the power vector $\mathbf{p}$. Other typical interpretations of the function $\Psi(\mathrm{SIR}_k(\mathbf{p}))$ include the degree of user satisfaction with the received SIR or the revenue of the network operator.

In the thesis, it is assumed that Ψ satisfies the following conditions [SWB07, SWB09]:

(A.1) $\Psi : \mathbb{R}_{++} \to \mathrm{Q}$ is a twice continuously differentiable and strictly increasing (bijective) function where Q is an open interval on the real line such that $\Psi^{-1} : \mathrm{Q} \to \mathbb{R}_{++}$.

(A.2) $\lim_{x\to 0} \Psi(x) := -\infty, \lim_{x\to 0} \Psi'(x) = \lim_{x\to 0} \frac{d\Psi}{dx}(x) = +\infty$

(A.3) $\Psi_e(x) := \Psi(e^x)$ is concave on $\mathbb{R}$.

The following class of traditional utility functions $\Psi : \mathbb{R}_{++} \to \mathrm{Q}$ introduced by [MW00] satisfies (A.1)–(A.3):

$$\Psi(x) = \Psi_\alpha(x) := \begin{cases} \log x & \alpha = 1 \\ \frac{x^{1-\alpha}}{1-\alpha} & \alpha \geq 2 \end{cases} \quad x > 0\,. \tag{2.3}$$

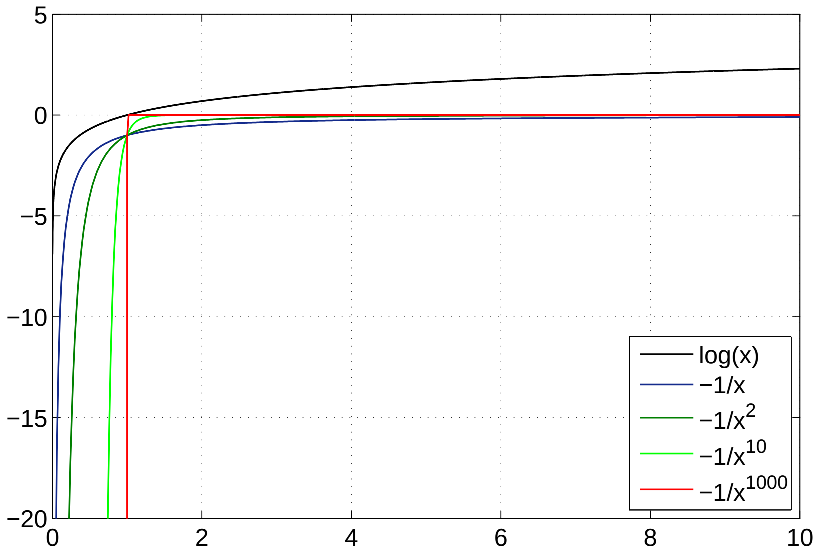

Figure 2.1: The class of utility functions (2.3) from [MW00].

We emphasize that another interesting class of utility functions is

$$\Psi(x) = \tilde{\Psi}_\alpha(x) := \begin{cases} \log(x) & \alpha = 1 \\ \log \frac{x}{1+x} & \alpha = 2 \\ \log \frac{x}{1+x} + \sum_{j=1}^{\alpha-2} \frac{1}{j(1+x)^j} & \alpha > 2 \end{cases} \quad x > 0\,. \tag{2.4}$$

We point out that condition (A.2) ensures that each link is assigned a nonzero SIR. More important, condition (A.3) permits to convert the utility optimization problem into a convex problem.

As mentioned at the beginning of this section, in order to be conform with the standard results from the optimization theory, we rewrite the utility-based power control problem as a minimization problem. Therefore we define a strictly decreasing SIR-QoS mapping $\psi : \mathbb{R}_{++} \to \mathrm{Q}$

$$\psi(x) := -\Psi(x), x \in \mathbb{R}_{++} \tag{2.5}$$

Now, $\psi : \mathbb{R}_{++} \to \mathrm{Q}$ is the inverse function of γ with $\gamma(\psi(x)) = x, x > 0$. Using

$$F_e(\mathbf{s}) := \sum\nolimits_{k\in\mathcal{K}} w_k \psi(\mathrm{SIR}_k(e^{\mathbf{s}}))\,. \tag{2.6}$$

the power control problem can be equivalently rewritten as

$$\mathbf{s}^* = \arg\min_{\mathbf{s}\in\mathrm{S}} F_e(\mathbf{s}), \tag{2.7}$$

From (2.5) the following properties of function ψ immediately result:

(A.4) $\psi : \mathbb{R}_{++} \to \mathrm{Q}$ is a twice continuously differentiable and strictly decreasing (bijective) function where Q is an open interval on the real line such that $\psi^{-1} : \mathrm{Q} \to \mathbb{R}_{++}$.

(A.5) $\lim_{x\to 0} \psi(x) := +\infty, \lim_{x\to 0} \psi'(x) = \lim_{x\to 0} \frac{d\psi}{dx}(x) = -\infty$.

(A.6) $\psi_e(x) := \psi(e^x)$ is convex on $\mathbb{R}$.

Condition (A.6) permits to convert the utility optimization problem into a convex problem. More precisely, the main reason for considering these classes of functions $\psi(x)$ is Observation 3.

Observation 3 ([SWB09]) *F_e is convex on $\mathbb{R}^K$, and thus, by convexity of* S*, the problem (2.7) is convex if (A.4)–(A.6) hold.*

This result evolved from the work on the geometry of the so-called feasible QoS region [BS04, BS05, CIM04]. In [BS04, BS05], it was shown that this region is a convex set if the SIR is a log-convex function of a QoS parameter of interest. As a consequence, the problem of optimizing the aggregate QoS of a network over the set of all feasible QoS levels is a convex problem. In [BWS04], it was shown that if the SIR is log-convex in the QoS, then the Karush-Kuhn-Tucker conditions for the corresponding power control problem are necessary and sufficient to characterize an optimal power allocation. Furthermore, if the log-convexity property holds, the power control problem can be converted into a convex optimization problem by the logarithmic transformation of the power vector.

The following recursive gradient projection algorithm with a constant, small enough step size $\delta > 0$ can be applied to solve the problem [BT89]:

$$\mathbf{s}(n+1) = \Pi_{\mathrm{S}}\big[\mathbf{s}(n) - \delta\nabla F_e(\mathbf{s}(n))\big], \quad \mathbf{s}(0) \in \mathrm{S},\ n \in \mathbb{N}_0 \tag{2.8}$$

where $\Pi_{\mathrm{S}}[\mathbf{x}]$ denotes the projection of $\mathbf{x} \in \mathbb{R}^K$ on S with respect to the Euclidean norm. The kth partial derivative $\nabla_k F_e(\mathbf{s})$ is equal to

$$\nabla_k F_e(\mathbf{s}) = \frac{\partial F_e}{\partial s_k}(\mathbf{s}) = e^{s_k}\Big(\phi_k(e^{\mathbf{s}}) - \sum_{l\in\mathcal{K}} v_{l,k}\mathrm{SIR}_l(e^{\mathbf{s}})\phi_l(e^{\mathbf{s}})\Big), \quad k \in \mathcal{K}, \tag{2.9}$$

with

$$\phi_k(e^{\mathbf{s}}) = \frac{w_k\psi'(\mathrm{SIR}_k(e^{\mathbf{s}}))}{I_k(e^{\mathbf{s}})}, \quad k \in \mathcal{K}. \tag{2.10}$$

The sequence $\{\mathbf{s}(n)\}$ generated by (2.8) converges to a stationary point that minimizes $F_e(\mathbf{s})$ over S.

Distributed power control algorithms for problem (2.7) that do not resort to the use of classical flooding protocols can be found in [SWB07, WSB08, SWB09]. The key ingredient in distributed implementation of these power control schemes is the use of an adjoint network to efficiently distribute some locally measurable quantities to other (logical) transmitters. More precisely, instead of each transmitter sending its message separately as in case of classical flooding protocols, some information is transmitted simultaneously over the adjoint network such that each transmitter can estimate its gradient component from the power of the received signal [SWB07, WSB08, SWB09]. For the sake of completeness and because of its importance to this thesis, the details of this concept are summarized in section 2.6 Appendix A.

Gradient projection algorithms are recursive in nature and mostly require an appropriate choice of step sizes. Since (adaptive) step size control is difficult to implement in distributed wireless networks, the step size sequence $\{\delta(n)\}$ is usually a non-increasing real-valued sequence satisfying one of the following conditions [KY03]:

(A.7) $\delta(n) = \delta, n \in \mathbb{N}_0$, for some sufficiently small $\delta > 0$.

(A.8) $\delta(n) > 0, \sum_{n=0}^{\infty} \delta(n) = \infty$ and $\lim_{n\to\infty} \delta(n) = 0$.

We observe that the second condition might be more reasonable if an algorithm is based on some noisy measurements. An appropriate choice of the step size is a key ingredient to the effectiveness of the algorithm [KY03]. Throughout the thesis, $\{\delta(n)\}$ is an appropriate step size sequence assumed to fulfill either (A.7) or (A.8).

2.2.1 Hop-by-Hop Congestion and Power Control

In this subsection we show how hop-by-hop congestion control can be combined with utility-based power control so as to achieve end-to-end fairness. Over the past years the problem of congestion control has received a lot of attention. Starting in the internet context the research mainly focused on the analysis and algorithm development of end-to-end schemes (e.g. TCP, Vegas) [Kel97, LPD02, LPW02, MW00]. Later these schemes have been adapted to wireless networks incorporating that in wireless networks the link capacities cannot be assumed to be fixed quantities. End-to-end window-based congestion control typically couples traditional transport layer protocols with physical or MAC layer protocols mainly to enhance the network performance in terms of some aggregate utility function. Chiang [Chi05] has proposed a cross-layer approach where power control is coupled with TCP Vegas in wireless multihop networks. Lin and Shroff [LS06] proposed a framework for joint rate control and scheduling and additionally investigated the impact of imperfect scheduling on the performance of the cross-layer congestion control. Such end-to-end schemes have some important disadvantages. The rates are adjusted with a period proportional to the end-to-end round-trip delays, which are usually large in multi-hop wireless networks and cause slow convergence and rate oscillations. The latter one may lead to large queues or data loss on congested links. In fact, because of slow convergence, dynamic wireless networks cannot determine correct rate allocations. In addition, the scalability argument does not hold in wireless networks where on the one hand the number of flows per node is much smaller than in the Internet and on the other hand per-flow queueing is typically used for reasons of scheduling.

Due to these facts, we argue in favor of MAC or link-layer fairness with hop-by-hop congestion control. Different MAC-layer strategies coupled with hop-by-hop congestion control were used to achieve network-wide end-to-end fairness [YS04, ST05]. In the following we couple a hop-by-hop congestion control similar to that of [ST05, TE92b] with power control.

We define $\mathcal{S}$ to be the set of S flows where $\mathcal{K}_s$ denotes the set of logical links belonging to flow $s, s \in \mathcal{S}$. Assume that each logical link $k \in \mathcal{K}$ is equipped with a single-user decoder achieving the Gaussian capacity so that the data rate $\nu_l^{(s)}(\mathbf{p})$ on link l of flow s is $\nu_l^{(s)}(\mathbf{p}) = \Phi(\mathrm{SIR}_l^{(s)}(\mathbf{p}))$ with $\Phi(x) = a\log(1+bx), x \geq 0$ for some positive constants a, b.[3] The notion of MAC-layer fairness is defined along similar lines as end-to-end fairness [MW00], except that instead of network flows, link-layer flows are considered. Our proposed MAC-layer fair policy with hop-by-hop congestion control works as follows: At the beginning of every frame, a (distributed) MAC controller chooses link rates $\boldsymbol{\nu}^* \in \mathbb{R}_+^K$ such that (assume that the maximum exists)

$$\boldsymbol{\nu}^* = \arg\max_{\boldsymbol{\nu} \in \mathrm{R}} \sum_{s \in \mathcal{S}} \sum_{l \in \mathcal{K}_s} U_l^{(s)}\big(\nu_l^{(s)}\big) \tag{2.11}$$

where $U_l^{(s)}$ is a continuously differentiable, increasing and strictly concave function, $\mathrm{R} := \{\boldsymbol{\nu}(\mathbf{p}) : \mathbf{p} \in \mathrm{P}\} \subset \mathbb{R}_+^K$ is the feasible rate region and $\boldsymbol{\nu}(\mathbf{p})$ is the vector of all data rates under the power vector $\mathbf{p}$. Since the noise vector $\mathbf{z}$ is a positive vector $\Phi : \mathbb{R}_+ \to \mathbb{R}_+$ is bijective, it may be seen from (1.2) that $\boldsymbol{\nu}(\mathbf{p})$ provides a continuous bijective map from R onto P. Hence, (2.11) can be formulated as a power control problem:

$$\mathbf{p}^* = \arg\max_{\mathbf{p} \in \mathrm{P}} \sum_{s \in \mathcal{S}} \sum_{l \in \mathcal{K}_s} U_l^{(s)}\big(\Phi(\mathrm{SIR}_l^{(s)}(\mathbf{p}))\big) \tag{2.12}$$

where the maximum exists. Clearly, given $\mathbf{p}^*$, the optimal data rate on link l of flow s is given by $\nu_l^{(s)*} = \Phi(\mathrm{SIR}_l^{(s)}(\mathbf{p}^*))$. We emphasize that this discussion on fair MAC-layer rate control and on end-to-end rate control provides in fact a motivation for considering the utility-based power control problem of section 2.2.

Now, taking (2.12) for the 2-user case it can be verified that (2.12) is in general not convex. Therefore, the following class of utility functions is considered:

$$U_l^{(s)}(x) = w_l^{(s)}\Psi(\Phi^{-1}(x)),\ x > 0, s \in \mathcal{S}, l \in \mathcal{K}_s \tag{2.13}$$

where $\Psi(x) =, x \in \mathbb{R}_{++}$, fulfills (A.1)–(A.3).

The class of utility functions is obtained by composing the traditional utility functions

[3] For brevity we assume $a = b = 1$ and the natural logarithmic function.

with $\Phi^{-1}(x) = e^x - 1, x > 0$. The effect of this is a linearization of the logarithmic rate-SIR curve. We see that the power control problem in (2.12) becomes

$$\mathbf{p}^* = \arg\max_{\mathbf{p}\in\mathrm{P}} G(\mathbf{p}) = \arg\max_{\mathbf{p}\in\mathrm{P}} \sum_{s\in\mathcal{S}} \sum_{l\in\mathcal{K}_s} w_l^{(s)} \Psi(\mathrm{SIR}_l^{(s)}(\mathbf{p}))\,. \tag{2.14}$$

Therefore, if $U_l^{(s)}$ has the form given by (2.13) and (2.3), then $\mathbf{p}^*$ in (2.14) can be regarded as being $(\mathbf{w},\alpha)$-fair with respect to the SIR [MW00]. Note that the proportional fair policy with $\alpha = 1$ is close to a throughput optimal one if $\mathrm{SIR}_l^{(s)}(\mathbf{p}) \gg 1$ for each $l \in \mathcal{K}_s, \forall_s$ and $\Phi(x) = \log(1+x)$. On the other hand, if $\mathrm{SIR}_l^{(s)}(\mathbf{p}) \ll 1$ for each $l \in \mathcal{K}_s, \forall_s$, then $\nu_l^{(s)}(\mathbf{p}) = \log(1 + \mathrm{SIR}_l^{(s)}(\mathbf{p})) \approx \mathrm{SIR}_l^{(s)}(\mathbf{p})$. So, at low SIR values, the modified utility functions are good approximations of the traditional ones in the sense that $\Psi_\alpha(\mathrm{SIR}_l^{(s)}(\mathbf{p})) \approx \Psi_\alpha(\nu_l^{(s)}(\mathbf{p}))$.

Now, the nonnegative weight w_l is used to couple the MAC layer with a congestion control protocol for each link and therefore depends on the current queue states at this link. In other words, the weights establish an interface between resource allocation and congestion control. One possible way to choose the weights provides the back-pressure policy [TE92a]. Let $u_l^{(s)}, 1 \le s \le S, 1 \le l \le |\mathcal{K}_s|$ denote the queue backlog at the source node of link l belonging to flow s. The weights $w_l^{(s)}, 1 \le s \le S, 1 \le l \le |\mathcal{K}_s|$ are updated according to:

$$w_l^{(s)} = \max\{q_l^{(s)} - q_{l+1}^{(s)}, 0\} \tag{2.15}$$

where $q_{|\mathcal{K}_s|+1}^{(s)} = 0$ denotes the queue backlog of the destination node of flow s. Another way may just use the current queue states at each link:

$$w_l^{(s)} = q_l^{(s)}. \tag{2.16}$$

In addition to the hop-by-hop congestion control, at source nodes, a simple window-based congestion control is used to prevent excessive queues at the nodes. The window size determines the maximum number of outstanding bits (or packets) a source can send to the network. In the following we assume that at the beginning of every frame, each source $s \in \mathcal{S}$ refills its window so that its size is W_s. As a consequence, each window size is fixed and equal to $W_s > 0, s \in \mathcal{S}$, implying that $q_1^{(s)} = W_s, \forall_s$.

Summarizingly, the joint hop-by-hop congestion control and power control is performed iteratively in an algorithm, in which the weights and transmit powers are updated alternately. The nth iteration step for both ways of choosing the weights (see for a) (2.15) and for b) (2.16)) is:

(i) Require: $q_1^{(s)} = W_s, \forall_s,\ w_l^{(s)}(0) > 0, \forall_s \forall_{l\in\mathcal{K}_s}, \mathbf{s}(0)$

(ii) $\mathbf{s}(n+1) = \Pi_{\mathrm{S}}\big(\mathbf{s}(n) + \delta\nabla G(\mathbf{s}(n))\big)$

(iii) a) $w_l^{(s)}(n+1) = \max\{q_l^{(s)}(n+1) - q_{l+1}^{(s)}(n+1), 0\}$ or

b) $w_l^{(s)}(n+1) = q_l^{(s)}(n+1)$

where

$$q_l^{(s)}(n+1) = q_l^{(s)}(n) + \min\big\{q_{l-1}^{(s)}(n), T\Phi\big(\mathrm{SIR}_{l-1}^{(s)}(\mathbf{p}(n+1))\big)\big\} \\ - \min\big\{q_l^{(s)}(n), T\Phi\big(\mathrm{SIR}_l^{(s)}(\mathbf{p}(n+1))\big)\big\}$$

(iv) $n+1 \to n$ and go to (ii)

where $\mathbf{p}(n) = e^{\mathbf{s}(n)}$ ($\mathbf{p}(n) > 0$ is positive), $\delta > 0$ is the step size, $\Pi(x)$ is the projection of $x \in \mathbb{R}^K$ on the modified set P, ∇G denotes the gradient of G defined in (2.14) and $T > 0$ denotes the frame length. We point out that step (iii) simplifies, if the queues are sufficiently long or T is sufficiently small. Then the min-operator can be neglected.

Since the window size at each source node $s \in \mathcal{S}$ is fixed and equal to W_s, we have $q_1^{(s)} = W_s$. Therefore, due to the back-pressure policy of (2.15), we see that all queue backlogs along the path associated with source $s \in \mathcal{S}$ are bounded. Now assume that the channel remains constant and that T is sufficiently small. The proposed joint power control and congestion control algorithm has been extensively tested in simulations and these experiments suggest that it converges to an equilibrium point[4]. The equilibrium is defined to be the weight and power vector pair $(\tilde{\mathbf{w}}, \tilde{\mathbf{p}})$, where $\tilde{\mathbf{w}} > 0$ and $\tilde{\mathbf{p}}$ is optimal in the sense of (2.14) with fixed $\tilde{\mathbf{w}}$. Since $\forall_{s\in\mathcal{S}} \forall_{k,l\in\mathcal{K}_s}$

$$w_l^{(s)}(n+1) = w_l^{(s)}(n) + T\Phi\big(\mathrm{SIR}_{l-1}^{(s)}(\mathbf{p}(n+1))\big) - 2T\Phi\big(\mathrm{SIR}_l^{(s)}(\mathbf{p}(n+1))\big) \\ + T\Phi\big(\mathrm{SIR}_{l+1}^{(s)}(\mathbf{p}(n+1))\big),$$

in an equilibrium we must have

$$\forall_{s\in\mathcal{S}} \forall_{k,l\in\mathcal{K}_s}\ \mathrm{SIR}_k^{(s)}(\mathbf{p}) = \mathrm{SIR}_l^{(s)}(\mathbf{p}) := \mathrm{SIR}_s(\tilde{\mathbf{p}}) \tag{2.17}$$

where the rates along the path of flow s are equal. Now suppose, that the problem converges to an equilibrium point defined by $(\tilde{\mathbf{w}}, \tilde{\mathbf{p}})$, then for the back-pressure policy it follows from (2.17) that

$$\sum_{s\in\mathcal{S}} \sum_{l\in\mathcal{K}_s} \tilde{w}_l^{(s)} \Psi(\mathrm{SIR}_s(\tilde{\mathbf{p}})) = \sum_{s=1}^{S} W_s \Psi(\mathrm{SIR}_s(\tilde{\mathbf{p}}))\,.$$

So, in an equilibrium point, a network-wide end-to-end fairness is provided, with the weights equal to the window sizes. Note that the window sizes influence the fairness between the flows. A large window size prioritizes its flow compared to other flows with small window sizes. Nevertheless, stronger fairness is achieved by increasing α instead of equalizing the window sizes. Max-min fairness can be achieved as $\alpha \to \infty$.

[4]If T is arbitrarily chosen the algorithm may alternate between different states.

Now the question is whether $(\tilde{\mathbf{w}}, \tilde{\mathbf{p}})$ converges to an equilibrium. As mentioned, supported by extensive simulations it seems that the joint congestion and power control algorithm converges to an equilibrium point for sufficiently small T. However, the proof of this problem is still open.

Corollary 1 *Suppose that A1-2 hold. Then, for sufficiently large window sizes, the joint congestion and power control algorithm converges to an equilibrium point specified by (2.17).*

Finally we present some numerical results of the proposed control scheme. We consider a simple network with 4 end-to-end flows and a string topology (Fig. 2.2). Figure 2.3 depicts the throughput and individual source rate performance as a function of α. In this simulation, all end-to-end flows have the same source window size W. The simulation confirms exemplarily that the link rates converge to a max-min fair rate allocation as $\alpha \to \infty$. Since there is no noise in the simulation and $\mathbf{V}$ is a primitive matrix, all link (and hence all end-to-end) rates are equal under the max-min fair policy. Furthermore, fairness improves with $\alpha \to \infty$ at the expense of the throughput performance. For $\alpha = 1$, the rate discrepancy is significant, with Flow 1 being allocated the lowest end-to-end rate.

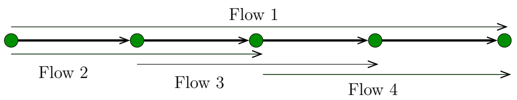

Figure 2.2: String topology with 4 end-to-end flows and an asymmetric scenario as the first flow goes through all wireless links.

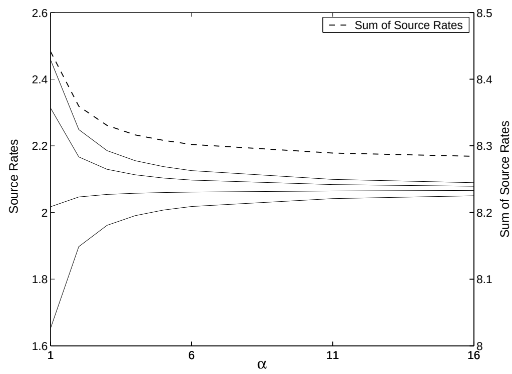

Figure 2.3: Individual flow rate and throughput performance as a function of α. $W_1 = \ldots = W_S$

Now let us look at the transient dynamics of the hop-by-hop scheme with $\Psi(x) = \log(x)$ ($\alpha = 1$). Starting from a steady state, the transient response to a change of the gain matrix $\mathbf{V}$ (see Fig.2.4) and to a change of the source window sizes (see 2.5) is examplarily depicted. In addition the transient response of an end-to-end scheme based on the congestion control protocol TCP Vegas is shown. In this case, the utility function for source $s \in \mathcal{S}$ is $\alpha_s d_s \log(x)$ where α_s is a constant and d_s is the propagation delay along the path originating from source s. Thus, the comparison is not quite fair because the end-to-end scheme adjusts the window sizes at source nodes to channel changes, whereas the source window sizes in the hop-by-hop scheme can remain constant. In other words, our protocol provides proportional fair rates with constant weights. The simulation however illustrates that the traditional end-to-end congestion control schemes may not be suitable for wireless networks where the propagation environment is highly dynamic and resources are strictly limited.

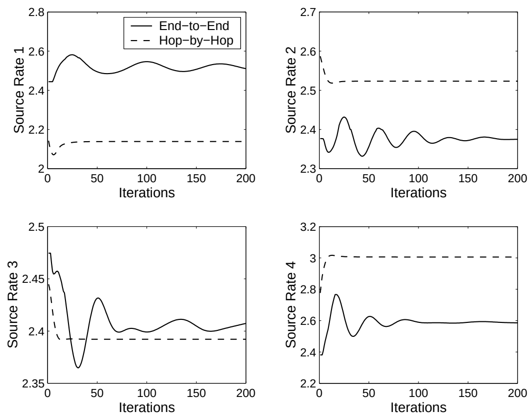

Figure 2.4: The transient behavior in response to a change of the gain matrix $\mathbf{V}$ for $\alpha = 1$.

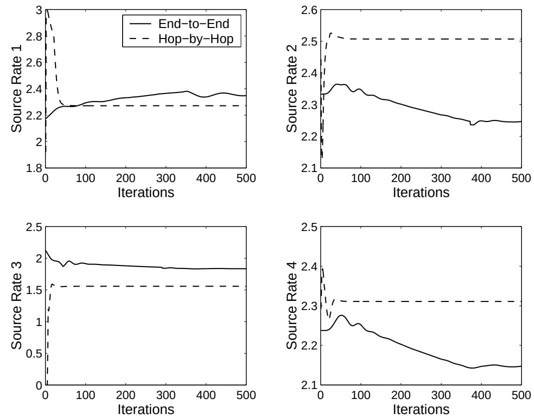

Figure 2.5: The transient behavior in response to a change of the source window sizes for $\alpha = 1$.

2.3 Hard QoS Support with a Distributed Primal-Dual Algorithm

In the following two sections 2.3 and 2.4 we present the main results of this chapter showing how QoS requirements expressed as SIR targets can be incorporated in the utility-based power control problem. We distinguish between hard and soft QoS support.

Power control with hard QoS support ensures that each link satisfies its QoS requirement, provided that the QoS vector $\boldsymbol{\omega}$ is feasible. Remind that this is equivalent to fulfilling $\mathrm{SIR}_k(\mathbf{p}) \geq \gamma_k, k \in \mathcal{K}$ with the SIR target $\gamma_k = \gamma(\omega_k) \geq 0$. We emphasize that although we have defined the function $\gamma : \mathrm{Q} \to \mathbb{R}_{++}$ to be the inverse function of the utility function $\psi : \mathbb{R}_{++} \to \mathrm{Q} \subseteq \mathbb{R}$, we may use any SIR target value $\gamma_k \geq 0$ by appropriately choosing $\boldsymbol{\omega}_k$.

The problem with hard QoS support writing as a minimization problem takes the following form

$$\mathbf{s}^*(\boldsymbol{\omega}) := \arg\min_{\mathbf{s} \in \mathrm{S}(\boldsymbol{\omega})} F_e(\mathbf{s}) \tag{2.18}$$

where $F_e : \mathbb{R}^K \to \mathbb{R}$ is defined by (2.6). Obviously, (2.18) has a solution if and only if

$\mathrm{S}(\boldsymbol{\omega}) \neq \emptyset$. In this section, we additionally assume that

(A.9) $\mathrm{int}(\mathrm{S}(\boldsymbol{\omega})) \neq \emptyset$ where $\mathrm{int}(\mathrm{A})$ is the interior of A.

Comparing (2.18) with (2.7) reveals that the only difference to the traditional utility-based power control is that the projection must be performed instead of S on the set $\mathrm{S}(\boldsymbol{\omega})$. Unfortunately the projection on $\mathrm{S}(\boldsymbol{\omega})$ is in general not amenable to distributed implementation as it may require excessive coordination between wireless nodes. The difficulty results from the fact that the SIR performance of any user generally depends on transmit powers of all other users.

To alleviate the projection problem, we propose a primal-dual algorithm that is based on a standard Lagrangian function and show that it can be implemented in a distributed manner using the scheme of [SWB07] summarized in the previous section. To this end, let

$$\mathcal{A} = \{k \in \mathcal{K} : \gamma_k > 0\} \tag{2.19}$$

be the set of QoS users and assume that the users are ordered: $\mathcal{A} = \{1, \ldots, J\}$ with $J = |\mathcal{A}|$. We consider an associated (linear) Lagrangian function $L : \mathbb{R}^K \times \mathbb{R}^J_+ \to \mathbb{R}$:

$$L(\mathbf{s}, \boldsymbol{\lambda}) = F_e(\mathbf{s}) + \sum\nolimits_{k\in\mathcal{A}} \lambda_k\, f_k(\mathbf{s}) \tag{2.20}$$

where $\mathbf{s} \in \mathrm{S}$, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_J) \in \mathbb{R}^J_+$ are the dual variables and $f_k : \mathbb{R}^K \to \mathbb{R}$ are the constraint functions given by

$$f_k(\mathbf{s}) := I_k(e^{\mathbf{s}})/e^{s_k} - 1/\gamma_k,\ k \in \mathcal{A}\,.$$

Since $f_k(\mathbf{s}) \leq 0, k \in \mathcal{A}$, $L(\mathbf{s}, \boldsymbol{\lambda}) \leq L(\mathbf{s}, \mathbf{0})$ for all $\mathbf{s} \in \mathrm{S}(\boldsymbol{\omega})$ and all $\boldsymbol{\lambda} \geq 0$. Hence for any given $\mathbf{s} \in \mathrm{S}(\boldsymbol{\omega}) \neq \emptyset$, $L(\mathbf{s}, \boldsymbol{\lambda})$ attains its maximum (which exists) at $\boldsymbol{\lambda} = \mathbf{0}$. From this and the unboundedness of $\boldsymbol{\lambda} \to L(\mathbf{s}, \boldsymbol{\lambda})$ on $\mathbb{R}^J_+$ for any set $\mathbf{s} \notin \mathrm{S}(\boldsymbol{\omega})$, it follows that $\min_{\mathbf{s}\in\mathrm{S}(\boldsymbol{\omega})} F_e(\mathbf{s}) = \min_{\mathbf{s}} \max_{\boldsymbol{\lambda}\geq\mathbf{0}} L(\mathbf{s}, \boldsymbol{\lambda})$ and

$$\mathbf{s}^*(\boldsymbol{\omega}) = \arg\min_{\mathbf{s}\in\mathrm{S}} \max_{\boldsymbol{\lambda}\geq 0} L(\mathbf{s}, \boldsymbol{\lambda})\,. \tag{2.21}$$

The corresponding dual problem to (2.18) is defined to be

$$\boldsymbol{\lambda}^*(\boldsymbol{\omega}) = \arg\max_{\boldsymbol{\lambda}\geq 0} \min_{\mathbf{s}\in\mathrm{S}} L(\mathbf{s}, \boldsymbol{\lambda})\,, \tag{2.22}$$

where the minimum (for any $\boldsymbol{\lambda} \geq 0$) and the maximum can be shown to exist. Now by Observation 2 and the results of [SWB09], we can conclude the following.

Lemma 1 *If (A.4)–(A.6) hold, then F_e is convex on $\mathbb{R}^K$, and hence, by convexity of $\mathrm{S}(\boldsymbol{\omega})$, the problem (2.18) is convex. Moreover, $f_k, k \in \mathcal{A}$, is convex as well so that the Lagrangian is a convex-concave function [Roc71].*

As a consequence of the lemma and (A.9), the Slater constraint qualification for the problem (2.18) is satisfied [Ber03]. So, by standard results of convex optimization theory, we have the following result.

Lemma 2 *The complementary slackness conditions are satisfied and strong duality holds so that:* $\max_{\boldsymbol{\lambda}\geq 0}\min_{\mathbf{s}\in \mathrm{S}} L(\mathbf{s},\boldsymbol{\lambda}) = \min_{\mathbf{s}\in \mathrm{S}}\max_{\boldsymbol{\lambda}\geq 0} L(\mathbf{s},\boldsymbol{\lambda})$.

These observations imply that the Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient conditions for optimality. Thus, we can find the minimum in (2.18) by solving the KKT conditions. This is equivalent to finding a stationary point of the Lagrangian function $L(\mathbf{s},\boldsymbol{\lambda})$ over $\mathrm{S}\times\mathbb{R}^J_+$, which is the saddle point $(\mathbf{s}^*(\boldsymbol{\omega}),\boldsymbol{\lambda}^*(\boldsymbol{\omega}))\in \mathrm{S}(\boldsymbol{\omega})\times\mathbb{R}^J_+$.

In order to find a saddle point of L with respect to $\mathrm{X} := \mathrm{S}\times\mathbb{R}^J_+$ we can apply a primal-dual algorithm of the following form:

$$\mathbf{x}(n+1) = \Pi_{\mathrm{X}}\Big[\mathbf{x}(n) - \delta\begin{pmatrix}\mathbf{I}_K & \mathbf{0}\\ \mathbf{0} & -\mathbf{I}_J\end{pmatrix}\nabla\tilde{L}(\mathbf{x}(n))\Big], n\in\mathbb{N}_0, \tag{2.23}$$

where $\mathbf{x} = (\mathbf{s},\boldsymbol{\lambda})\in\mathbb{R}^{K+J}$ is a stacked vector, $\tilde{L}:\mathbb{R}^{K+J}\to\mathbb{R}$ is an extension of $L(\mathbf{s},\boldsymbol{\lambda})$, $(\mathbf{s},\boldsymbol{\lambda})\in \mathrm{S}\times\mathbb{R}^J_+$ to the domain $\mathbb{R}^{K+J}$, $\delta > 0$ is a sufficiently small step size, $\mathbf{I}_m$ is the identity matrix of dimension m, $\nabla\tilde{L}(\mathbf{x})$ is the gradient of $\tilde{L}$ with respect to $\mathbf{x}\in\mathbb{R}^{K+J}$ and

$$\Pi_{\mathrm{X}}(\mathbf{x}) = \big(\min\{s_1,\log(P_1)\},\dots\min\{s_K,\log(P_K)\},\max\{0,\lambda_1\},\cdots,\max\{0,\lambda_J\}\big)$$

is the projection of $\mathbf{x}$ on X. Computing the partial derivatives of (2.20) shows that the algorithm (2.23) takes the form

$$\begin{cases} s_k(n+1) = \min\Big\{s_k(n) - \delta\Big[\phi_k(\mathbf{s}(n))e^{s_k(n)} - \dfrac{\mu_k(n)I_k(e^{\mathbf{s}(n)})}{e^{s_k(n)}} \\ \qquad + e^{s_k(n)}S_k\big(\mathbf{s}(n),\boldsymbol{\mu}(n)\big)\Big],\log(P_k)\Big\}, k\in\mathcal{K} \\ \lambda_k(n+1) = \max\Big\{0,\lambda_k(n)+\delta f_k\big(\mathbf{s}(n)\big)\Big\}, k\in\mathcal{A} \\ \mu_k(n+1) = \begin{cases}\lambda_k(n+1) & k\in\mathcal{A}\\ 0 & k\in\mathcal{K}\setminus\mathcal{A}\end{cases}\end{cases} \tag{2.24}$$

for each $k\in\mathcal{K}$ (note that $v_{k,k}=0$ for each $k\in\mathcal{K}$), where

$$\phi_k(\mathbf{s}) = w_k\psi'\big(\mathrm{SIR}_k(e^{\mathbf{s}})\big)/I_k(e^{\mathbf{s}}) \tag{2.25}$$

and

$$S_k(\mathbf{s},\boldsymbol{\mu}) = \sum\nolimits_{l\in\mathcal{K}} v_{l,k}\Big(\frac{\mu_l}{e^{s_l}}2.6 - \mathrm{SIR}_l(e^{\mathbf{s}})\phi_l(\mathbf{s})\Big) \\ = \sum\nolimits_{l\in\mathcal{K}} v_{l,k}\underbrace{\Big(\frac{\mu_l}{e^{s_l}} + \big|\mathrm{SIR}_l(e^{\mathbf{s}})\phi_l(\mathbf{s})\big|\Big)}_{m_l(\mathbf{s},\boldsymbol{\mu})}\,. \tag{2.26}$$

The last step follows from the fact that $\phi_k(\mathbf{s})$ is negative on $\mathbb{R}^K$ since ψ is strictly decreasing. Observe that the iterations are performed simultaneously from which it follows that the convergence to the saddle point that provides the optimal solution is not a straightforward result.

Proposition 1 *If $\mathbf{V} \geq 0$ is such that each column has at least one positive entry*[5]*, then the sequence $\{\mathbf{s}(n), \boldsymbol{\lambda}(n)\}_{n\in\mathbb{N}_0}$ generated by (2.23) converges to a saddle point $(\mathbf{s}^*(\boldsymbol{\omega}), \boldsymbol{\lambda}^*(\boldsymbol{\omega}))$ given by (2.21) and (2.22).*

Proof: By Lemma 1, Lemma 2 and [Roc71, pp. 125–126], it is sufficient to show that both $\mathbf{s} \mapsto L(\mathbf{s}, \boldsymbol{\lambda})$ is strictly convex for any $\boldsymbol{\lambda} \geq 0$ and $\boldsymbol{\lambda} \mapsto L(\mathbf{s}^*(\boldsymbol{\omega}), \boldsymbol{\lambda})$ is a strictqly concave function of $\boldsymbol{\lambda} \geq \mathbf{0}$. Then we can conclude, that a simultaneous application of gradient methods to minimize the Lagrangian over primal variables and to maximize it over dual variables converges to a saddle point $(\mathbf{s}^*(\boldsymbol{\omega}), \boldsymbol{\lambda}^*(\boldsymbol{\omega}))$.

Now we prove the strict convexity and strict concavity properties. If each column of $\mathbf{V}$ has at least one positive entry and ψ satisfies (A.4)– (A.6), then we know from [SWB07] that F_e is strongly convex on any bounded subset of $\mathbb{R}^K$ so that $\mathbf{s} \mapsto L(\mathbf{s}, \boldsymbol{\lambda})$ is strictly convex for any $\boldsymbol{\lambda} \geq 0$. Thus, we only need to prove the strict concavity property. To this end, let $\mu \in (0, 1)$ and $\hat{\boldsymbol{\lambda}}, \check{\boldsymbol{\lambda}} \geq 0$ with $\hat{\boldsymbol{\lambda}} \neq \check{\boldsymbol{\lambda}}$ be arbitrary, and let $L(\boldsymbol{\lambda}) := \min_{\mathbf{s}\in\mathrm{S}}\big(F_e(\mathbf{s}) + \sum_{l\in\mathcal{A}} \lambda_l f_l(\mathbf{s})\big)$. Due to the minimum operator, it is clear that

$$L(\boldsymbol{\lambda}(\mu)) = L((1-\mu)\hat{\boldsymbol{\lambda}} + \mu\check{\boldsymbol{\lambda}}) \geq (1-\mu)L(\hat{\boldsymbol{\lambda}}) + \mu L(\check{\boldsymbol{\lambda}})$$

with equality if and only if one of the following holds:

(i) There are $\hat{\mathbf{s}} \in \mathrm{S}$ and $\check{\mathbf{s}} \in \mathrm{S}$ with $\hat{\mathbf{s}} \neq \check{\mathbf{s}}$ such that $L(\hat{\boldsymbol{\lambda}}) = L(\hat{\mathbf{s}}, \hat{\boldsymbol{\lambda}}), L(\check{\boldsymbol{\lambda}}) = L(\check{\mathbf{s}}, \check{\boldsymbol{\lambda}})$ and $L(\boldsymbol{\lambda}(\mu)) = (1-\mu)L(\hat{\mathbf{s}}, \hat{\boldsymbol{\lambda}}) + \mu L(\check{\mathbf{s}}, \check{\boldsymbol{\lambda}})$, $\hat{\boldsymbol{\lambda}} \neq \check{\boldsymbol{\lambda}}$.

(ii) There exists $\mathbf{s}^* \in \mathrm{S}$ such that $L(\mathbf{s}^*, \hat{\boldsymbol{\lambda}}) = L(\hat{\boldsymbol{\lambda}})$ and $L(\mathbf{s}^*, \check{\boldsymbol{\lambda}}) = L(\check{\boldsymbol{\lambda}})$, $\hat{\boldsymbol{\lambda}} \neq \check{\boldsymbol{\lambda}}$.

Consequently, in order to prove the proposition, it is sufficient to show that neither (i) nor (ii) can hold. To disprove (i), assume that $\mathbf{s}(\mu) \in \mathrm{S}$ minimizes $L(\mathbf{s}, \boldsymbol{\lambda}(\mu))$ over $\mathbf{s} \in \mathrm{S}$ so that $L(\boldsymbol{\lambda}(\mu)) = L(\mathbf{s}(\mu), \boldsymbol{\lambda}(\mu))$. Then, by (i) and a linearity of $\boldsymbol{\lambda} \mapsto L(\mathbf{s}, \boldsymbol{\lambda})$ for any fixed $\mathbf{s}$,

[5] Note that this condition is weaker than irreducibility.

we have

$$
\begin{aligned}
(1-\mu)L(\hat{\mathbf{s}},\hat{\boldsymbol{\lambda}})+\mu L(\check{\mathbf{s}},\check{\boldsymbol{\lambda}}) &= L(\boldsymbol{\lambda}(\mu)) = L(\mathbf{s}(\mu),\boldsymbol{\lambda}(\mu)) \\
&= (1-\mu)L(\mathbf{s}(\mu),\hat{\boldsymbol{\lambda}})+\mu L(\mathbf{s}(\mu),\check{\boldsymbol{\lambda}}).
\end{aligned}
$$

Thus, (i) always implies (ii) so that it is sufficient to disprove (ii). By the KKT conditions and (ii), we have $\frac{\partial L}{\partial s_k}(\mathbf{s}^*,\hat{\boldsymbol{\lambda}}) = \frac{\partial L}{\partial s_k}(\mathbf{s}^*,\check{\boldsymbol{\lambda}}), k \in \mathcal{K}$, which implies that

$$
\frac{\hat{\lambda}_k}{e^{s_k^*}} - \frac{e^{s_k^*}}{I_k(e^{\mathbf{s}^*})}\sum_{l\in\mathcal{A}} v_{l,k}\frac{\hat{\lambda}_l}{e^{s_l^*}} = \frac{\check{\lambda}_k}{e^{s_k^*}} - \frac{e^{s_k^*}}{I_k(e^{\mathbf{s}^*})}\sum_{l\in\mathcal{A}} v_{l,k}\frac{\check{\lambda}_l}{e^{s_l^*}}
$$

for each $k \in \mathcal{A}$. Defining

$$
\hat{\mathbf{r}} = (\hat{\lambda}_k/e^{s_k^*})_{k\in\mathcal{A}} = (\hat{\lambda}_1/e^{s_1^*},\dots,\hat{\lambda}_K/e^{s_K^*}) \text{ and } \check{\mathbf{r}} = (\check{\lambda}_k/e^{s_k^*})_{k\in\mathcal{A}} = (\check{\lambda}_1/e^{s_1^*},\dots,\check{\lambda}_K/e^{s_K^*}),
$$

the equality above can be written as

$$
(\mathbf{I}-\tilde{\mathbf{D}}\tilde{\mathbf{V}}^T)\hat{\mathbf{r}} = (\mathbf{I}-\tilde{\mathbf{D}}\tilde{\mathbf{V}}^T)\check{\mathbf{r}} \tag{2.27}
$$

where we have $\tilde{\mathbf{D}} = \mathrm{diag}((\mathrm{SIR}_k(e^{\mathbf{s}^*}))_{k\in\mathcal{A}})$ and $\tilde{\mathbf{V}} = (v_{k,l})_{k\in\mathcal{A},l\in\mathcal{A}}$. First note that if $\mathcal{A} = \mathcal{K}$, then (2.27) is $(\mathbf{I}-\mathbf{D}\mathbf{V}^T)\hat{\mathbf{r}} = (\mathbf{I}-\mathbf{D}\mathbf{V}^T)\check{\mathbf{r}}$, with the matrix $\mathbf{D} = \mathrm{diag}((\mathrm{SIR}_k(e^{\mathbf{s}^*}))_{k\in\mathcal{K}})$. Since the diagonal entries of $\mathbf{D}$ are feasible SIRs, it is well-known [SWB09] that $\rho(\mathbf{D}\mathbf{V}) = \rho(\mathbf{D}\mathbf{V}^T) < 1$ must hold. Consequently, as $\tilde{\mathbf{D}}\tilde{\mathbf{V}}$ is a sub-matrix of $\mathbf{D}\mathbf{V}$, we have $\rho(\tilde{\mathbf{D}}\tilde{\mathbf{V}}^T) < 1$. So, the inverse of $\mathbf{I}-\tilde{\mathbf{D}}\tilde{\mathbf{V}}^T$ exists and thus, by (2.27), one has $\hat{\mathbf{r}} = \check{\mathbf{r}}$ or, equivalently, $\hat{\boldsymbol{\lambda}} = \check{\boldsymbol{\lambda}}$. This contradicts $\hat{\boldsymbol{\lambda}} \neq \check{\boldsymbol{\lambda}}$ and hence completes the proof. ■

We emphasize that the primal-dual algorithm does not provide monotonicity. However, the algorithm can be implemented in distributed wireless networks using the scheme based on the adjoint network [SWB07]. Except for S_k given by (2.26), all the other quantities (such as the weights w_k, the possibly different functions ψ_k and its derivatives) are either known locally at both the transmitter side and the receiver side of link $k \in \mathcal{K}$ or can be computed from local measurements (such as the SIR) and, if necessary, conveyed to the corresponding transmitter/receiver by means of a low-rate control channel. In contrast, S_k can be estimated from the received signal power in the adjoint network with the exchanged roles of transmitters/receivers and a channel inversion on each link as described in [SWB07]. The only difference to the scheme of [SWB07] is that each receiver, say receiver $l \in \mathcal{K}$, transmits in the nth iteration a sequence of independent zero-mean random symbols with the variance being equal to $m_l(\mathbf{s}(n),\boldsymbol{\mu}(n))$ defined in (2.26). The main steps of the distributed power control scheme are summarized below (see Algorithm 1). Finally, we point out that due to the projection on X in every step, it is ensured that the intermediate power vectors satisfy the power constraints, which is important due to Step

2 of Algorithm 1. A primal-dual algorithm based on a non-linear (modified) Lagrangian function can be found in [WSB08] and provides faster convergence than using a standard Lagrangian.

Algorithm 1 Distributed primal-dual algorithm

Require: $\mathbf{w} > 0, \epsilon > 0, n = 0, \mathbf{s}(0) \in \mathrm{S}$, $\boldsymbol{\omega}$ with $\mathrm{S}(\boldsymbol{\omega}) \neq \emptyset$, a sufficiently small step size $\delta > 0$.

Ensure: $\mathbf{s} \in \mathrm{S}$

1: **repeat**
2: Concurrent transmission at transmit powers $e^{s_k(n)}, k \in \mathcal{K}$, with receiver-side estimation of $I_k(e^{\mathbf{s}(n)})$ and $\mathrm{SIR}_k(e^{\mathbf{s}(n)})$.
3: Each transmitter-receiver pair exchanges the necessary estimates and variables including $I_k(e^{\mathbf{s}(n)}), s_k(n), k \in \mathcal{K}$ and $\lambda_k(n), k \in \mathcal{A}$.
4: Concurrent transmission in the adjoint network with transmitter-side estimation of the received power. The variance of the zero-mean input symbols is $m_k(\mathbf{s}(n), \mu_k(n))$ given by (2.26).
5: Transmitter-side computation of $s_k(n+1), k \in \mathcal{K}$ and $\lambda_k(n+1), k \in \mathcal{A}$ according to (2.24).
6: $n = n + 1$
7: **until** $|L(\mathbf{s}(n), \boldsymbol{\lambda}(n)) - L(\mathbf{s}(n-1), \boldsymbol{\lambda}(n-1))| < \epsilon$

2.4 Soft QoS Support

In addition to the non-monotonic behavior of the primal-dual algorithm, there is an inherent problem associated with hard QoS support resulting from the fact that no valid power allocation may exist to meet the SIR targets due to channel and/or network dynamics. Note that this feature pertains to any classical QoS-based power control strategy and may pose a significant challenge in distributed wireless networks as the necessary communication overhead for coping with the problem of in-feasibility can explode, requiring a lot of additional resources and thereby deteriorating the overall network performance significantly. For this reason, we argue in favor of soft QoS support in which case a solution to the corresponding power control problem always exists even if the SIR targets are not feasible.

One possibility of incorporating soft QoS into (2.1) is to dynamically choose the weights. There even exists a relation between the SIR targets and the choice of the weights in the utility function as is described in more detail in section 3.2.2 see Proposition 5. However, so far a fully distributed calculation of the weights does not exist. Thus, we have a closer look at the second possibility. Assuming that each user is a QoS user, meaning $\gamma_k > 0, \forall_{k \in \mathcal{K}}$, we may approximate a solution to the so-called max-min SIR balancing problem (see for instance [SB06] and references therein). It is known that $\boldsymbol{\omega}$ is feasible if and only if the SIR targets are met under a max-min SIR balancing power vector. Unfortunately, a max-min

SIR balancing problem is for general power constraints notoriously difficult to solve in a distributed manner [MW00]. However, on the positive side a max-min SIR balancing solution can be approximated arbitrarily closely by solving the following utility-based power control problem:

$$\bar{\mathbf{s}}(\alpha) := \arg\max_{\mathbf{s}\in\mathcal{S}} \sum\nolimits_{k\in\mathcal{K}} w_k \Psi_\alpha\big(g_k(\mathbf{s})\big) \tag{2.28}$$

where

$$g_k(\mathbf{s}) := \mathrm{SIR}_k\big(e^{\mathbf{s}}\big)/\gamma_k,\ \gamma_k > 0, k \in \mathcal{K} \tag{2.29}$$

and $\Psi_\alpha : \mathbb{R}_{++} \to \mathrm{Q}$ is given by 2.3. Since $\boldsymbol{\omega}$ is feasible if and only if the QoS requirements are satisfied under a max-min SIR balancing solution, the following can be easily deduced from [MW00, Lemma 3].

Observation 4 *Using (1.2) and assuming that $\gamma_k > 0$ for each $k \in \mathcal{K}$. Then, $\boldsymbol{\omega} \in \mathrm{Q}^K$ is feasible if and only if, for any $\epsilon > 0$, there exists $\alpha(\epsilon) \geq 2$ such that*

$$\forall_{\alpha\geq\alpha(\epsilon)}\,\forall_{k\in\mathcal{K}}\,\mathrm{SIR}_k\big(e^{\bar{\mathbf{s}}(\alpha)}\big)/\gamma_k \geq 1-\epsilon\,. \tag{2.30}$$

This observation says, that we can arbitrarily closely approximate a solution to the max-min SIR balancing problem by choosing the parameter α sufficiently large. The approximation becomes more accurate as α increases. Practically, the observation implies that if $\boldsymbol{\omega}$ is feasible, then each user meets its QoS requirement under power control (2.28) if $\alpha \geq 2$ is sufficiently large.

However, an important drawback of strategy (2.28) is certainly the inability of choosing $\gamma_k = 0$, which raises the question of how to incorporate best-effort users into the optimization. Also, once the SIR targets are met for a sufficiently large α, the users with relatively high SIR targets are preferred when allocating remaining powers, which may be inefficient in terms of throughput when compared to the strategies with hard QoS support considered in Section 2.3. As a possible solution, we consider a strategy that combines the power control strategy (2.28) with the "pure" utility-based approach (2.1). To this end, we define two sets $\mathcal{A} \subseteq \mathcal{K}$ and $\mathcal{B} \subseteq \mathcal{K}$ such that $\mathcal{A} \cup \mathcal{B} = \mathcal{K}$. The set $\mathcal{A}$ is defined by (2.19), while $\mathcal{B}$ contains indices of all users without any QoS requirements (best effort users) but may also contain indices of all other users. Thus, these sets are not necessarily disjoint so that $\mathcal{A}\cap\mathcal{B} \neq \emptyset$. We have however $\gamma_k = 0$ for each $k \in \mathcal{K}\backslash\mathcal{A}$. Without loss of generality, the users are assumed to be ordered so that $\mathcal{A} = \{1, \ldots, J\}, \mathcal{B} = \{M+1, \ldots, K\}, 0 \leq M \leq J \leq K$.

With these definitions in hand, we can formulate the power control problem of interest in this section as follows:

$$\tilde{\mathbf{s}}(\boldsymbol{\alpha}) := \tilde{\mathbf{s}}(\boldsymbol{\alpha}, \boldsymbol{\omega}) = \arg\max_{\mathbf{s}\in\mathcal{S}} \tilde{G}_{\boldsymbol{\alpha}}(\mathbf{s}) \tag{2.31}$$

where

$$\tilde{G}_{\boldsymbol{\alpha}}(\mathbf{s}) = \sum_{k\in\mathcal{A}} a_k \Psi_{\alpha_k}\big(g_k(\mathbf{s})\big) + \sum_{k\in\mathcal{B}} b_k \Psi\big(\mathrm{SIR}_k(e^{\mathbf{s}})\big) \tag{2.32}$$

and where $\mathbf{a} > 0$ and $\mathbf{b} > 0$ are given weight vectors, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_J)$ with $\alpha_k \geq 2, k \in \mathcal{A}$, and $g_k(\mathbf{s})$ is defined by (2.29), $\Psi_{\alpha_k}(x), x > 0, k \in \mathcal{A}$, is given by (2.3) with $\alpha = \alpha_k \geq 2$, and $\Psi(x), x > 0$, is any function that satisfies (A.1)–(A.3).

Since $\mathcal{A}$ contains indices of all QoS users, the first addend in (2.32) acts as a penalty function for not meeting the SIR targets. Thus, (2.31) can be seen as a barrier method with a barrier function approximating the indicator function of the feasible power region. The approximation improves with increasing $\alpha_k, \forall_k$. For a better visualization of the barrier function we refer to figure 2.1. Here it can be seen that the corner point of $\Psi_\alpha(x)$ for $\alpha \to \infty$ is at $x = 1$. For $x \geq 1$ we have $\Psi_{\alpha\to\infty}(x) \to 0$ and for $x < 1$ we have $\Psi_{\alpha\to\infty}(x) \to -\infty$. This means for a utility function with large α_k, as long as $g_k = \mathrm{SIR}_k/\gamma_k < 1$ a marginal increase of g_k leads to a high increase in utility. In constrast, if $g_k \geq 1$ the increase in utility is very small and vanishes for $\alpha_k \to \infty$ to 0. Consequently, for a sufficiently large α_k the user is urged to approximate the SIR-target γ_k. If $\alpha_k, \forall_k$ is sufficiently large, the choice of the weight vector $\mathbf{a}$ has negligible impact on the optimal power vector $\tilde{\mathbf{s}}(\boldsymbol{\alpha})$ so that one can often assume $\mathbf{a} = \mathbf{1}$. Notice that (2.31) with $\mathcal{A} = \emptyset, \mathcal{B} = \mathcal{K}$, is the utility-based power control problem (2.1), and if we choose $\mathcal{A} = \mathcal{K}, \mathcal{B} = \emptyset$, the problem reduces to (2.28) which approaches the max-min SIR-balancing solution.

In the following we prove some interesting problem properties of (2.31). Again, in order to stay consistent with optimization theory, we reformulate the power control problem (2.31) as an equivalent minimization problem:

$$\tilde{\mathbf{s}}(\boldsymbol{\alpha}) := \tilde{\mathbf{s}}(\boldsymbol{\alpha}, \boldsymbol{\omega}) = \arg\min_{\mathbf{s}\in\mathrm{S}} \tilde{F}_{\boldsymbol{\alpha}}(\mathbf{s}) \tag{2.33}$$

where

$$\tilde{F}_{\boldsymbol{\alpha}}(\mathbf{s}) = \sum_{k\in\mathcal{A}} a_k \psi_{\alpha_k}\big(g_k(\mathbf{s})\big) + \sum_{k\in\mathcal{B}} b_k \psi\big(\mathrm{SIR}_k(e^{\mathbf{s}})\big) \tag{2.34}$$

and where $\psi(x) = -\Psi(x)$ is any function that satisfies the corresponding conditions (A.4)–(A.6).

For simplicity, we prove our results for $\boldsymbol{\alpha} = (\alpha, \ldots, \alpha), \alpha \geq 2$, in which case (2.33) and (2.34) have the following form:

$$\tilde{\mathbf{s}}(\alpha) := \tilde{\mathbf{s}}(\alpha, \boldsymbol{\omega}) = \arg\min_{\mathbf{s}\in\mathrm{S}} \tilde{F}_{\alpha}(\mathbf{s}) \tag{2.35}$$

and

$$\tilde{F}_\alpha(\mathbf{s}) = \sum_{k\in\mathcal{A}} a_k \psi_\alpha\big(g_k(\mathbf{s})\big) + \sum_{k\in\mathcal{B}} b_k \psi\big(\mathrm{SIR}_k(e^{\mathbf{s}})\big). \tag{2.36}$$

It may be easily deduced from the proofs that the more general case can be reasoned in the same way (only with a more cumbersome notation) by defining the parameter $\alpha \geq 2$ in the following derivations to be

$$\alpha := \min_{k\in\mathcal{A}} \alpha_k\,. \tag{2.37}$$

Thus, each $\alpha_k, k \in \mathcal{A}$, can be made arbitrarily large by increasing α and $\alpha_k \to \infty$ for each $k \in \mathcal{A}$ as $\alpha \to \infty$. In view of distributed implementation, this is a very useful property and is in fact exploited in section 2.4.3.

Observation 5 *Each of the following is true: (i) The minimum in (2.35) exists. (ii) The problem is convex. (iii) $\nabla \tilde{F}_\alpha(\mathbf{s})$ is Lipschitz continuous on every bounded subset of $\mathbb{R}^K$.*

Proof: Due to [SWB07], both addends on the right-hand side of (2.36) attain their minima on S, are convex and their gradients satisfy (iii). So, (i)–(iii) must be also true for $\tilde{F}_\alpha$. ■

In the remainder of this section, we first discuss and prove some relevant properties of the proposed power control approach (2.35). Section 2.4.3 presents a distributed gradient algorithm for solving the power control problem.

2.4.1 Infeasible QoS Requirements

Let us first assume that the QoS vector $\boldsymbol{\omega}$ is infeasible, in which case we have $\mathrm{S}(\boldsymbol{\omega}) = \emptyset$. Our goal is to show that, as α tends to infinity, the vector $\exp(\tilde{\mathbf{s}}(\alpha))$ converges to a vector $\underline{\mathbf{p}} \in \mathrm{P}$ defined to be

$$\underline{\mathbf{p}} = \arg\max_{\mathbf{p}\in\mathrm{P}} \min_{k\in\mathcal{A}} \frac{\mathrm{SIR}_k(\mathbf{p})}{\gamma_k} = \arg\min_{\mathbf{p}\in\mathrm{P}} \max_{k\in\mathcal{A}} \psi_\alpha\Big(\frac{\mathrm{SIR}_k(\mathbf{p})}{\gamma_k}\Big). \tag{2.38}$$

In words, $\underline{\mathbf{p}}$ maximizes the worst relative SIR among the QoS users over P (see also the discussion after Proposition 2). We emphasize that $\underline{\mathbf{p}}$ is not necessarily unique and the second equality holds due to (i) in the following lemma.

Lemma 3 *Each of the following holds.*

(i) *We have $\underline{p}_k > 0, k \in \mathcal{A}$.*

(ii) *If the sub-matrix $\bar{\mathbf{V}} := (v_{k,l})_{k,l\in\mathcal{A}} = (v_{k,l})_{1\leq k,l\leq J}$ of $\mathbf{V}$ is irreducible, then the sub-vector $(\underline{p}_k)_{k\in\mathcal{A}}$ of $\underline{\mathbf{p}}$ is unique and*

$$\mathrm{SIR}_k(\underline{\mathbf{p}})/\gamma_k = \beta, \quad k \in \mathcal{A}, \textit{for some } \beta > 0. \tag{2.39}$$

If, in addition, $v_{k,l} > 0$ *for some* $k \in \mathcal{A}$ *and* $l \in \mathcal{K} \setminus \mathcal{A}$*, then* $\underline{p}_l = 0$.

Proof: (*i*) As $\min_{k\in\mathcal{A}}(\mathrm{SIR}_k(\mathbf{p})/\gamma_k)$ is positive for any positive power vector $\mathbf{p} \in \mathrm{P}$ and equals zero whenever $p_k = 0$ for some $k \in \mathcal{A}$, it is obvious that $\underline{p}_k > 0$ for each $k \in \mathcal{A}$. (*ii*) The uniqueness property and (2.39) follow from [SKB10]. Finally, if we had $\underline{p}_l > 0$ for some $l \in \mathcal{K} \setminus \mathcal{A}$ with $v_{k,l} > 0, k \in \mathcal{A}$, then, by (2.39), we could always increase $\min_{k\in\mathcal{A}}(\mathrm{SIR}_k(\underline{\mathbf{p}})/\gamma_k)$ by letting $\underline{p}_l = 0$. Thus, by (2.38), we must have $\underline{p}_l = 0$. ■
Quite importantly, the lemma states that $\underline{\mathbf{p}}$ defined by (2.38) may have zero components, in which case $\underline{\mathbf{p}}$ has no counterpart in S. This problem is circumvented in the proof of the following proposition by considering a sequence of positive power vectors that approach a given nonnegative power vector.

Proposition 2 *Assume that*

$$0 < \overline{g} := \sup_{\mathbf{s}\in\mathrm{S}} \min_{k\in\mathcal{A}} g_k(\mathbf{s}) = \max_{\mathbf{p}\in\mathrm{P}} \min_{k\in\mathcal{A}} \frac{\mathrm{SIR}_k(\mathbf{p})}{\gamma_k} < 1 \tag{2.40}$$

and

$$\lim_{x\to\infty} x\, a^x\, \psi(1/x) = 0, \; a \in (0,1)\,. \tag{2.41}$$

Then, for any $\epsilon > 0$ *and* $\mathbf{V} \geq 0$*, there exists* $\alpha(\epsilon, \mathbf{V}) \geq 2$ *such that*

$$1 \leq \frac{\overline{g}}{\min_{k\in\mathcal{A}} g_k(\tilde{\mathbf{s}}(\alpha))} \leq 1 + \epsilon \tag{2.42}$$

for all $\alpha \geq \alpha(\epsilon, \mathbf{V}) \geq 2$.

Proof: Let $\mathbf{s} \in \mathrm{S}$ be arbitrary and $\alpha' = \alpha - 1 \in \mathbb{N}$. We define $\mathcal{A}(\mathbf{s}) := \big\{k \in \mathcal{A} : g_k(\mathbf{s}) = \min_l g_l(\mathbf{s})\big\}$ and $\mathcal{A}^c(\mathbf{s}) = \mathcal{A} \setminus \mathcal{A}(\mathbf{s})$. Without loss of generality, we can assume that $\sum_{k\in\mathcal{A}(\mathbf{s})} a_k = 1$ for all $\alpha \geq 2$. So, by strict monotonicity of ψ_α and the fact that $a_k\psi_\alpha(g_k(\mathbf{s})) > 0$ for every $k \in \mathcal{A}, \mathbf{s} \in \mathrm{S}$, and $\alpha \geq 2$, we have (for any $\mathbf{s} \in \mathrm{S}$ and $\alpha' \in \mathbb{N}$)

$$R_\alpha(\mathbf{s}) = \frac{\tilde{F}_\alpha(\mathbf{s})}{\psi_\alpha(g_{k_0}(\mathbf{s}))} = \Big[1 + \sum_{k\in\mathcal{A}^c(\mathbf{s})} a_k \Big(\frac{g_{k_0}(\mathbf{s})}{g_k(\mathbf{s})}\Big)^{\alpha'} + \alpha'(g_{k_0}(\mathbf{s}))^{\alpha'} \sum_{k\in\mathcal{B}} b_k \psi(\mathrm{SIR}_k(e^{\mathbf{s}}))\Big]$$

where $k_0 \in \mathcal{A}(\mathbf{s})$ is arbitrary. Note that $\psi(x) \to \infty$ as $x \to 0$ and $g_{k_0}(\mathbf{s})/g_k(\mathbf{s}) < 1, \mathbf{s} \in \mathrm{S}$, for each $k \in \mathcal{A}^c(\mathbf{s})$. Moreover, by (2.40), we have $\sup_{\mathbf{s}\in\mathrm{S}} g_{k_0}(\mathbf{s}) \in (0,1)$ so that considering (2.41) shows each of the following.

(i) $\liminf_{\alpha\to\infty} R_\alpha(\tilde{\mathbf{s}}(\alpha)) \geq 1$, where $\tilde{\mathbf{s}}(\alpha)$ is given by (2.33).

(ii) $\limsup_{\alpha\to\infty} R_\alpha(\mathbf{s}(\alpha)) = \liminf_{\alpha\to\infty} R_\alpha(\mathbf{s}(\alpha)) = 1$ where $\mathbf{s}(\alpha) = \log \mathbf{p}(\alpha), \alpha \geq 2$, and

$\mathbf{p}(\alpha) = (p_1(\alpha), \dots, p_K(\alpha)) > 0$ is defined as follows:

$$p_k(\alpha) = \begin{cases} \frac{\alpha}{1+\alpha}\underline{p}_k & \underline{p}_k > 0 \\ \frac{1}{1+\alpha} & \underline{p}_k = 0 \end{cases}$$

with $\underline{\mathbf{p}}$ given by (2.38).

Note that since P is a convex set, there is $\alpha_0 \geq 2$ such that $\mathbf{s}(\alpha) \in \mathrm{S}$ for all $\alpha \geq \alpha_0$. Furthermore, we have $\lim_{\alpha\to\infty} \exp(\mathbf{s}(\alpha)) = \underline{\mathbf{p}}$. Thus, we have

$$\begin{aligned} 1 &\overset{(ii)}{=} \lim_{\alpha\to\infty} R_\alpha(\mathbf{s}(\alpha)) \overset{(ii)}{=} \limsup_{\alpha\to\infty} R_\alpha(\mathbf{s}(\alpha)) \\ &\overset{(a)}{\geq} \limsup_{\alpha\to\infty} \frac{\tilde{F}_\alpha(\tilde{\mathbf{s}}(\alpha))}{\max_{k\in\mathcal{A}} \psi_\alpha(g_k(\mathbf{s}(\alpha)))} \\ &= \limsup_{\alpha\to\infty} \frac{R_\alpha(\tilde{\mathbf{s}}(\alpha)) \max_{k\in\mathcal{A}} \psi_\alpha(g_k(\tilde{\mathbf{s}}(\alpha)))}{\max_{k\in\mathcal{A}} \psi_\alpha(g_k(\mathbf{s}(\alpha)))} \\ &\overset{(i)}{\geq} \limsup_{\alpha\to\infty} \frac{\max_{k\in\mathcal{A}} \psi_\alpha(g_k(\tilde{\mathbf{s}}(\alpha)))}{\max_{k\in\mathcal{A}} \psi_\alpha(g_k(\mathbf{s}(\alpha)))} \\ &\overset{(b)}{=} \limsup_{\alpha\to\infty} \frac{\max_{k\in\mathcal{A}} \psi_\alpha(g_k(\tilde{\mathbf{s}}(\alpha)))}{\inf_{\mathbf{s}\in\mathrm{S}} \max_{k\in\mathcal{A}} \psi_\alpha(g_k(\mathbf{s}))} \\ &= \limsup_{\alpha\to\infty} \frac{\max_{k\in\mathcal{A}} \psi_\alpha(g_k(\tilde{\mathbf{s}}(\alpha)))}{\psi_\alpha(\bar{g})} \geq 1 \end{aligned}$$

where (a) is due to the fact that $\tilde{\mathbf{s}}(\alpha)$ minimizes $\tilde{F}_\alpha(\mathbf{s})$ over S and (b) follows from (2.38) and the definition of $\mathbf{s}(\alpha)$. Thus, as $\max_{k\in\mathcal{A}} \psi_\alpha(g_k(\tilde{\mathbf{s}}(\alpha))) \geq \psi_\alpha(\bar{g})$ for all $\alpha \geq 2$, we have

$$\begin{aligned} \limsup_{\alpha\to\infty} \frac{\psi_\alpha(\min_{k\in\mathcal{A}} g_k(\tilde{\mathbf{s}}(\alpha)))}{\psi_\alpha(\bar{g})} &= \liminf_{\alpha\to\infty} \frac{\psi_\alpha(\min_{k\in\mathcal{A}} g_k(\tilde{\mathbf{s}}(\alpha)))}{\psi_\alpha(\bar{g})} \\ &= \lim_{\alpha\to\infty} \Big(\frac{\bar{g}}{\min_{k\in\mathcal{A}} g_k(\tilde{\mathbf{s}}(\alpha))}\Big)^{\alpha-1} = 1. \end{aligned}$$

This implies (2.42), and thus completes the proof. ■

We point out that the assumption (2.40) is slightly stronger than the assumption $\mathrm{S}(\boldsymbol{\omega}) = \emptyset$, as the latter one is equivalent to $\sup_{\mathbf{s}\in\mathrm{S}} \min_{k\in\mathcal{A}} g_k(\mathbf{s}) \leq 1$. Further, the assumption (2.41) holds for any function belonging to the function class (2.3).

The proposition implies that $\exp(\tilde{\mathbf{s}}(\alpha))$ with $\tilde{\mathbf{s}}(\alpha)$ defined by (2.35) converges to a vector $\underline{\mathbf{p}}$ as $\alpha \to \infty$, provided that (2.40) and (2.41) are satisfied. We observe that $\underline{\mathbf{p}}$ is not unique but if $\bar{\mathbf{V}}$ is irreducible, then, by Lemma 3, $\underline{\mathbf{p}}$ is unique up to the transmit powers of the users in the set $\mathcal{K} \setminus \mathcal{A}$. These powers however have no impact on the objective function in (2.38). For brevity, let us assume in the remaining discussion that

(A.10) for each $l \in \mathcal{K} \setminus \mathcal{A}$, there is $k \in \mathcal{A}$ such that $v_{k,l} > 0$. By Lemma 3, this implies that $\underline{p}_k = 0$ for each $k \in \mathcal{K} \setminus \mathcal{A}$ (all best-effort users are inactive).

It is important to emphasize that irreducibility of $\mathbf{V}$ does not imply that $\bar{\mathbf{V}}$ is irreducible, and vice versa.

Proceeding essentially as in [SKB10] shows that the point, where the half-line starting at the zero point in the direction of the vector $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_K)$ intersects the boundary of the feasible SIR region,[6] is achieved by some $\underline{\mathbf{p}}$. By Lemma 3 and (A.10), this point (and also $\underline{\mathbf{p}}$) is unique if the sub-matrix $\bar{\mathbf{V}}$ is irreducible, in which case the intersection point is called the max-min fair point. Thus, under the assumption of irreducibility of $\bar{\mathbf{V}}$, Proposition 2 shows that, as α tends to infinity, the QoS users achieve the max-min fair point under the power control strategy (2.35). This is illustrated in Figure 2.6.

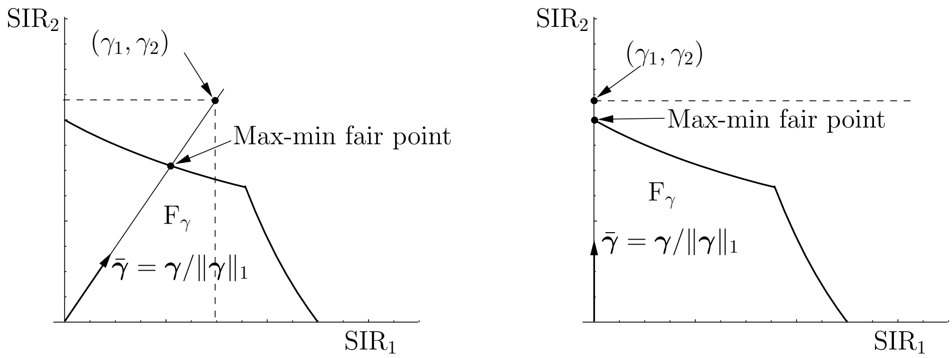

Figure 2.6: An illustration of Proposition 2 for the case of two users with (A.10) and an irreducible matrix $\bar{\mathbf{V}}$. F_γ is the feasible SIR region ($\gamma(x) = x, x \geq 0$) and the infeasible SIR targets are represented by the dashed lines. $\bar{\boldsymbol{\gamma}} = \boldsymbol{\gamma}/\|\boldsymbol{\gamma}\|_1$. Power control strategy (2.35) asymptotically achieves the max-min fair point as $\alpha \to \infty$. In the left picture, both users are QoS users, whereas in the right picture the first user is a best effort user. Thus, by (A.10), the first user is allocated no transmit power in the max-min fair point.

It is clear from the figure that the max-min fair point is not necessarily the nearest point to the point $\boldsymbol{\gamma}$ in the sense of the Euclidean distance (norm). In other words, if the SIR targets $\boldsymbol{\gamma}$ are infeasible, then the max-min fair point does not need to be the projection of $\boldsymbol{\gamma}$ on the feasible SIR region with respect to the Euclidean norm. It is however the projection of $\boldsymbol{\gamma}$ on the feasible SIR region with respect to a (weighted) maximum norm (l_∞-norm). To see this, consider the following corollary.

Corollary 2 *Let $\bar{\mathbf{V}}$ be irreducible, and let (A.10) be satisfied. Define $\boldsymbol{\Gamma} = \mathrm{diag}(\boldsymbol{\gamma})$ and*

$$\|\boldsymbol{\Gamma}^{-1}\mathbf{x}\|_\infty = \begin{cases} \max_{k\in\mathcal{A}}(x_k/\gamma_k) & \forall_{k\in\mathcal{K}\setminus\mathcal{A}}\, x_k = 0 \\ \infty & \exists_{k\in\mathcal{K}\setminus\mathcal{A}} x_k > 0\,. \end{cases}$$

[6]Note that the feasible SIR region is equal to F_γ if $\gamma(x) = x, x \geq 0$.

Suppose that (2.40) holds. Then,

$$\forall_{\mathbf{p}\in\mathrm{P}}\ \|\mathbf{\Gamma}^{-1}\mathbf{e}(\mathbf{p}')\|_\infty \le \|\mathbf{\Gamma}^{-1}\mathbf{e}(\mathbf{p})\|_\infty \tag{2.43}$$

holds with equality if and only if $\mathbf{p}' = \underline{\mathbf{p}}$, *where* $\mathbf{e}(\mathbf{p}) = (e_1(\mathbf{p}), \dots, e_K(\mathbf{p}))$ *is used to denote the vector of the absolute errors (distances):* $e_k(\mathbf{p}) = |\gamma_k - \mathrm{SIR}_k(\mathbf{p})|, k \in \mathcal{K}$.

Proof: By (2.40) and (A.10), we have

$$\begin{aligned}
\|\mathbf{\Gamma}^{-1}\mathbf{e}(\underline{\mathbf{p}})\|_\infty =& \max_{k\in\mathcal{A}} |1 - \mathrm{SIR}_k(\underline{\mathbf{p}})/\gamma_k| = (1 - \min_{k\in\mathcal{A}}(\mathrm{SIR}_k(\underline{\mathbf{p}})/\gamma_k)) \\
=& (1 - \max_{\mathbf{p}\in\mathrm{P}} \min_{k\in\mathcal{A}}(\mathrm{SIR}_k(\mathbf{p})/\gamma_k)) \\
\le& (1 - \min_{k\in\mathcal{A}}(\mathrm{SIR}_k(\mathbf{p})/\gamma_k)) \\
=& \max_{k\in\mathcal{A}} |1 - (\mathrm{SIR}_k(\mathbf{p})/\gamma_k)| \le \|\mathbf{\Gamma}^{-1}\mathbf{e}(\mathbf{p})\|_\infty
\end{aligned}$$

for all $\mathbf{p} \in \mathrm{P}$. So, the corollary follows since $\underline{\mathbf{p}}$ is unique by (*ii*) of Lemma 3 and (A.10). ■

Thus, in the special case of equal SIR targets $\mathbf{\Gamma} = \gamma\mathbf{I}$, the max-min fair point is the projection of the vector $\boldsymbol{\gamma}$ on the feasible SIR region with respect to the (standard) l_∞-norm. If we interpret $e_k(\mathbf{p})/\gamma_k, k \in \mathcal{K}$ (with $e_k(\mathbf{p})/0 = +\infty$ if $e_k(\mathbf{p}) > 0$ and zero if $e_k(\mathbf{p}) = 0$), as relative errors (distances), then the max-min fair point is the nearest point to $\boldsymbol{\gamma}$ in the sense of the maximum relative distance.

2.4.2 Feasible QoS Requirements

Now let us turn our attention to the case of feasible QoS requirements so that we have $\mathrm{S}(\boldsymbol{\omega}) \neq \emptyset$ (notice that (2.45) trivially holds if $\mathrm{S}(\boldsymbol{\omega}) = \emptyset$).

The following simple lemma is used to prove the main result in case of feasible QoS requirements (see Proposition 3).

Lemma 4 *Consider (2.35). If* $\mathrm{S}(\boldsymbol{\omega}) \neq \emptyset$, *then, for any* $\mathbf{s} \in \mathrm{S}(\boldsymbol{\omega})$, *there are constants* $c_1 > -\infty$ *and* $c_2 = c_2(\mathbf{s}) < +\infty$ *such that* $c_1 \le \tilde{F}_\alpha\big(\tilde{\mathbf{s}}(\alpha)\big) \le \tilde{F}_\alpha(\mathbf{s}) \le c_2$ *for all* $\alpha \ge 2$. *So, if* $\mathrm{S}(\boldsymbol{\omega}) \neq \emptyset$, *there are* $0 < c_3 \le c_4 < \infty$ *such that*

$$\forall_{\alpha\ge 2}\ c_3 \le \min_{k\in\mathcal{K}} \mathrm{SIR}_k(e^{\tilde{\mathbf{s}}(\alpha)}) \le \max_{k\in\mathcal{K}} \mathrm{SIR}_k(e^{\tilde{\mathbf{s}}(\alpha)}) \le c_4 \tag{2.44}$$

and hence the entries of $\tilde{\mathbf{s}}(\alpha) \in \mathrm{S}$ *are bounded for all* $\alpha \ge 2$.

Proof: The lower bounds follow from Observation 5(i). The upper bounds hold as $\tilde{F}_\alpha(\mathbf{s}(\alpha)) \le \tilde{F}_\alpha(\mathbf{s})$ for all $\mathbf{s} \in \mathrm{S}$ and $\tilde{F}_\alpha(\mathbf{s})$ is bounded above for any $\mathbf{s} \in \mathrm{S}(\boldsymbol{\omega})$. ■

Now we are in a position to prove the main result, which reveals important features of an optimal solution to (2.35).

Proposition 3 *Suppose that $\mathcal{A}\backslash\mathcal{B} \neq \emptyset$ and $\mathcal{B} \neq \emptyset$. Then, for any $\epsilon > 0$ and an irreducible matrix $\mathbf{V} \geq 0$, there exists $\alpha(\epsilon, \mathbf{V}) \geq 1$ such that*

$$\max_{k\in\mathcal{A}\backslash\mathcal{B}} \mathrm{SIR}_k(e^{\tilde{\mathbf{s}}(\alpha)})/\gamma_k \leq 1+\epsilon \tag{2.45}$$

for all $\alpha \geq \alpha(\epsilon, \mathbf{V}) \geq 1$. Moreover, if $\mathrm{S}(\boldsymbol{\omega}) \neq \emptyset$, then, for any nonnegative but not necessarily irreducible matrix $\mathbf{V}$,

$$1-\epsilon \leq \min_{k\in\mathcal{A}} \mathrm{SIR}_k(e^{\tilde{\mathbf{s}}(\alpha)})/\gamma_k\,. \tag{2.46}$$

Proof: Let $\gamma_k > 0, k \in \mathcal{A}$, and $\gamma_k = 0, k \in \mathcal{B} \setminus \mathcal{A} \neq \emptyset$. Without loss of generality, the weight vectors $\mathbf{a}$ and $\mathbf{b}$ are assumed to be $\mathbf{a} = \mathbf{b} = \mathbf{1}$. Furthermore, assume that $\alpha \geq 2$. First, we prove (2.46) by contradiction. Thus, assume that there are $\epsilon_1 \in (0,1)$ and $\mathcal{A}_0 \subseteq \mathcal{A}$ such that $0 < a := \max_{k\in\mathcal{A}_0} \mathrm{SIR}_k(e^{\tilde{\mathbf{s}}(\alpha)})/\gamma_k < 1-\epsilon_1$ for all $\alpha \geq 2$. So, using

$$\phi(\tilde{\mathbf{s}}(\alpha)) = \sum_{k\in\mathcal{A}\backslash\mathcal{A}_0} \psi_\alpha\Big(\frac{\mathrm{SIR}_k(e^{\tilde{\mathbf{s}}(\alpha)})}{\gamma_k}\Big) + \sum_{k\in\mathcal{B}} \psi\big(\mathrm{SIR}_k(e^{\tilde{\mathbf{s}}(\alpha)})\big)$$

one obtains

$$\tilde{F}_\alpha(\tilde{\mathbf{s}}(\alpha)) \geq |\mathcal{A}_0|\psi_\alpha(a) + \phi(\tilde{\mathbf{s}}(\alpha)) > |\mathcal{A}_0|\psi_\alpha(1-\epsilon_1) + \phi(\tilde{\mathbf{s}}(\alpha))$$

where the inequalities hold since ψ_α is strictly decreasing and $\phi(\tilde{\mathbf{s}}(\alpha))$ is bounded for all $\alpha \geq 2$ (by Lemma 4). Now observe that, as $\alpha \to \infty$, we have $\psi_\alpha(x) \to +\infty$ for any $x \in (0,1)$, and $\psi_\alpha(x) \to 0$ for any $x \geq 1$. Moreover, $\tilde{F}_\alpha(\mathbf{s})$ is bounded for any $\mathbf{s} \in \mathrm{S}(\boldsymbol{\omega}) \neq \emptyset$. Thus it follows

$$\tilde{F}_\alpha(\tilde{\mathbf{s}}(\alpha)) > \psi_\alpha(1-\epsilon_1) + \phi(\tilde{\mathbf{s}}(\alpha)) \geq \tilde{F}_\alpha(\mathbf{s})$$

for all $\alpha \geq \alpha(\epsilon_1, \mathbf{s})$ with a sufficiently large $\alpha(\epsilon_1, \mathbf{s})$. This, however, contradicts the fact that $\tilde{\mathbf{s}}(\alpha)$ minimizes $\tilde{F}_\alpha(\mathbf{s})$ over S, and thus proves (2.46).

In order to prove (2.45), we first remark that this bound trivially holds if $\mathrm{S}(\boldsymbol{\omega}) = \emptyset$. So, in what follows, assume that $\mathrm{S}(\boldsymbol{\omega}) \neq \emptyset$. By Observation 5, the minimum exists and the Karush-Kuhn-Tucker conditions are necessary and sufficient for $\tilde{\mathbf{s}}(\alpha)$ to be a global minimizer of $\tilde{F}_\alpha(\mathbf{s})$. So, for any fixed $\alpha \geq 2$, we have

$$\begin{aligned} \mathbf{u}(\alpha) &= \big(\mathbf{I} - \mathbf{V}^T\mathbf{D}(\alpha)\big)^{-1}\mathbf{t}(\alpha) \\ e^{\tilde{\mathbf{s}}(\alpha)} &= \big(\mathbf{I} - \mathbf{D}(\alpha)\mathbf{V}\big)^{-1}\mathbf{D}(\alpha)\mathbf{z} \in \mathrm{P}_+ \subset \mathrm{P}\,. \end{aligned} \tag{2.47}$$

The notation in (2.47) is defined as follows

(a) $\mathbf{D}(\alpha) := \mathbf{D}\big(\tilde{\mathbf{s}}(\alpha)\big) = \mathrm{diag}(d_1(\alpha), \dots, d_K(\alpha))$ where $d_k(\alpha) = \mathrm{SIR}_k\big(e^{\tilde{\mathbf{s}}(\alpha)}\big),\ k \in \mathcal{K}$. Observe that (2.44) implies that $d_k(\alpha) > 0$ for each $k \in \mathcal{K}$ regardless of the choice of

$\alpha \geq 2$, and hence $\mathbf{D}(\alpha)$ is positive definite for all $\alpha \geq 2$.

(b) $\mathbf{u}(\alpha) := \mathbf{u}\big(\tilde{\mathbf{s}}(\alpha)\big) = (u_1(\alpha), \dots, u_K(\alpha)) > 0$ with

$$u_k(\alpha) = \begin{cases} \phi_{\alpha,k}(\tilde{\mathbf{s}}(\alpha)) & k \in \mathcal{A} \setminus \mathcal{B} \\ \phi_{\alpha,k}(\tilde{\mathbf{s}}(\alpha)) - \theta_k(\tilde{\mathbf{s}}(\alpha)) & k \in \mathcal{A} \cap \mathcal{B} \\ -\theta_k(\tilde{\mathbf{s}}(\alpha)) & k \in \mathcal{B} \setminus \mathcal{A}\,, \end{cases} \tag{2.48}$$

where $\phi_{\alpha,k}(\tilde{\mathbf{s}}(\alpha)) = e^{-\tilde{s}_k(\alpha)}\Big(\frac{\gamma_k}{\mathrm{SIR}_k\left(e^{\tilde{\mathbf{s}}(\alpha)}\right)}\Big)^{\alpha-1} > 0$ and $\theta_k(\mathbf{s}) = \psi'(\mathrm{SIR}_k(e^{\mathbf{s}}))/I_k(e^{\mathbf{s}}) < 0$.

(c) $\mathbf{t}(\alpha) := \mathbf{t}\big(\tilde{\mathbf{s}}(\alpha)\big) = (t_1(\alpha), \dots, t_K(\alpha)) \geq 0$ depends on the Lagrangian multipliers resulting from the power constraints $\mathbf{s} \in \mathrm{S}$ (inequality constraints). Note that we must have $\mathbf{t}(\alpha) \neq \mathbf{0}$ since otherwise there would be no positive solution $\mathbf{u}(\alpha) > 0$ to (2.47), regardless of the choice of $\mathbf{V}$. By Theorem 1 (see 2.6 Appendix B), we see that there exists a positive vector $\mathbf{u}(\alpha)$ satisfying (2.47) iff $\rho(\mathbf{D}(\alpha)\mathbf{V}) = \rho(\mathbf{V}^T\mathbf{D}(\alpha)) < 1$, which is satisfied due to the second equality in (2.47) and the existence of $e^{\tilde{\mathbf{s}}(\alpha)} > 0$.

Since $\mathbf{D}(\alpha)$ is positive definite for all $\alpha \geq 2$ and $\mathbf{V}$ is irreducible, the matrices $\mathbf{D}(\alpha)\mathbf{V}$ and $\mathbf{V}^T\mathbf{D}(\alpha), \alpha \geq 2$, are irreducible as well. Now let $\mathbf{y}(\alpha)$ and $\mathbf{x}(\alpha)$ with $\mathbf{y}(\alpha)^T\mathbf{x}(\alpha) = 1$ be left and right positive eigenvectors of $\mathbf{V}^T\mathbf{D}(\alpha)$ associated with $\rho(\mathbf{V}^T\mathbf{D}(\alpha)) > 0$, respectively. Then, by Theorem 2 (see 2.6 Appendix B) and the fact that $\rho(\mathbf{V}^T\mathbf{D}(\alpha)) = \rho(\mathbf{D}(\alpha)\mathbf{V})$, the first equality of (2.47) can be written as

$$\mathbf{u}(\alpha) = \frac{1}{1 - \rho(\mathbf{D}(\alpha)\mathbf{V})}\mathbf{Z}(\alpha)\mathbf{t}(\alpha) + \mathbf{R}(\alpha)\mathbf{t}(\alpha) \tag{2.49}$$

where $\mathbf{Z}(\alpha) = \mathbf{x}(\alpha)\mathbf{y}(\alpha)^T$ is positive (and bounded in any norm due to $\mathbf{y}(\alpha)^T\mathbf{x}(\alpha) = 1$) and $\mathbf{R}(\alpha) \in \mathbb{R}^{K \times K}$ follows from (2.57). Since, for all $\alpha \geq 2$, $\rho(\mathbf{V}^T\mathbf{D}(\alpha)) < 1$ and $\tilde{\mathbf{p}}(\alpha) = e^{\tilde{\mathbf{s}}(\alpha)} \in \mathrm{P}_+$ belongs to a bounded subset of $\mathbb{R}^K_{++}$, it follows from the second equality in (2.47) and Theorem 3 (see 2.6 Appendix B) that there exists a constant $c_1 \in (0,1)$ independent of α such that $1 \leq 1/(1 - \rho(\mathbf{D}(\alpha)\mathbf{V})) \leq 1/c_1 < +\infty$ for all $\alpha \geq 2$. From this and (2.49), we then have $\|\mathbf{u}(\alpha)\|_1 \leq \|\mathbf{Z}(\alpha)\mathbf{t}(\alpha)\|_1/c_1 + \|\mathbf{R}(\alpha)\mathbf{t}(\alpha)\|_1$. By (2.44) and (2.48), the left-hand side is bounded (consider $u_k(\alpha)$ for any $k \in \mathcal{B} \neq \emptyset$). So, by the discussion above and the fact that induced matrix norms are compatible with their underlying vector norms, one obtains

$$\|\mathbf{t}(\alpha)\|_1 \geq \frac{\|\mathbf{u}(\alpha)\|_1}{\|\mathbf{Z}(\alpha)\|_1/c_1 + \|\mathbf{R}(\alpha)\|_1} \geq c_2 > 0\,.$$

Thus, $\|\mathbf{t}(\alpha)\|_1$ is bounded away from zero for all $\alpha \geq 2$. On the other hand, considering the Neumann series $\mathbf{u}(\alpha) = (\mathbf{I} - \mathbf{V}^T\mathbf{D}(\alpha))^{-1}\mathbf{t}(\alpha) = \sum_{j=0}^{\infty}(\mathbf{V}^T\mathbf{D}(\alpha))^j\mathbf{t}(\alpha)$, we obtain

$$u_k(\alpha) = \sum\nolimits_{j=0}^{\infty} \sum_{l \in \mathcal{K}} a_{k,l}^{(j)}(\alpha) t_l(\alpha),\ k \in \mathcal{K} \tag{2.50}$$

where $a_{k,l}^{(j)}(\alpha) = (\mathbf{A}^j(\alpha))_{k,l}$ and $\mathbf{A}(\alpha) = (a_{k,l}(\alpha)) = (\mathbf{V}^T\mathbf{D}(\alpha))$. Now assume that there exist $k_0 \in \mathcal{A} \setminus \mathcal{B}$ and $\epsilon_2 > 0$ such that $\text{SIR}_{k_0}(e^{\tilde{\mathbf{s}}(\alpha)})/\gamma_{k_0} > 1 + \epsilon_2$ for all $\alpha \geq 2$, which contradicts (2.45). This is equivalent to saying that

$$\forall_{\alpha \geq 2}\, 0 < b(\alpha) = \gamma_{k_0}/\text{SIR}_{k_0}(e^{\tilde{\mathbf{s}}(\alpha)}) < 1 - \epsilon_3$$

for $k_0 \in \mathcal{A} \setminus \mathcal{B}$ and some $\epsilon_3 = \epsilon_2/(1+\epsilon_2) \in (0, 1)$. Consequently, if α goes to infinity, (2.48) implies that the left-hand side of (2.50) with $k = k_0 \in \mathcal{A} \setminus \mathcal{B}$ tends to zero. However, the right-hand side of (2.50) is bounded away from zero by some constant for all $\alpha \geq 2$, and hence (2.45) must hold. To see that the right-hand side of (2.50) cannot be arbitrarily close to zero, let $t_{l_0}(\alpha) = \max_l t_l(\alpha) \geq \|\mathbf{t}(\alpha)\|_1/K \geq c_3 = c_2/K > 0$ and note that since $\mathbf{V}^T\mathbf{D}(\alpha)$ is irreducible, Lemma 5 (see 2.6 Appendix B) implies that there is a finite $j = j(k, l) \geq 0$ for each pair $(k, l), 1 \leq l, k \leq K$, such that $a_{k,l}^{(j)}(\alpha) > 0$. Moreover, since the SIRs cannot be arbitrarily small in the minimum due to Lemma 4, there exists some $j \geq 0$ and $c_4 > 0$ such that $a_{k_0,l_0}^{(j)}(\alpha) \geq c_4$, from which and (2.50) we have $u_{k_0}(\alpha) \geq c_3 \cdot c_4 > 0$ for all $\alpha \geq 2$. ■

According to the proposition, all users in $\mathcal{A}$ (QoS users) satisfy their SIR targets under an optimal power vector $\tilde{\mathbf{s}}(\alpha)$, provided that α is sufficiently large and the SIR targets are feasible. Additionally, for each user in $\mathcal{A} \setminus \mathcal{B}$, a potential overshoot of its SIR target vanishes as α tends to infinity. Thus, for large values of α, these users meet their SIR targets with equality whenever the targets are feasible. In contrast to that, each QoS user in $\mathcal{A} \cap \mathcal{B}$ may overshoot its SIR target. Indeed, if α is large enough and all SIR targets are met, the remaining power resources are allocated so as to minimize the second addend on the right-hand side of (2.36), which also accounts for best-effort users in $\mathcal{B} \setminus \mathcal{A}$. Thus, if the objective is to maximize the throughput over these remaining resources, a reasonable choice for $\psi(x) = -\Psi(x)$ is $\psi(x) = -\log(x), x > 0$, because this function seems to best approximate the (negative) rate function $\log(1 + x), x > 0$, among the functions fulfilling (A.4)–(A.6).

Figure 2.7 illustrates Proposition 3. We have two users so that the first user is a best-effort user ($\mathcal{B} \setminus \mathcal{A} = \mathcal{B} = \{1\}$) and the second user is a pure QoS user ($\mathcal{A} \setminus \mathcal{B} = \mathcal{A} = \{2\}$). Point 1 corresponds to the max-min SIR performance, while point 2 corresponds to the traditional utility-based power control (2.7) with $\psi(x) = \log(x), x > 0$. In this point, the SIR target of user 2 is not achieved although it is feasible. Now, under the power control strategy with soft QoS support (2.35), it follows from (2.46) that user 2 achieves its SIR target if $\alpha = \alpha_2$ is sufficiently large. Moreover, by (2.45) and (2.46), $\text{SIR}_2(e^{\tilde{\mathbf{s}}(\alpha)})$ converges

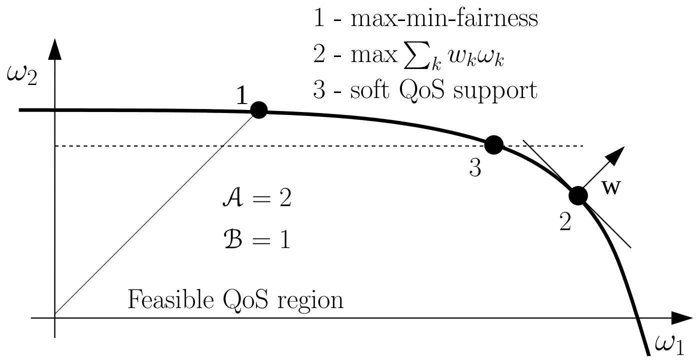

Figure 2.7: An illustration of power control strategy (2.33) with one QoS user and one best-effort user.

to γ_2 as $\alpha = \alpha_2 \to \infty$. In other words, for sufficiently large α, the operating point of the scheme is point 3 in Figure 2.7. We emphasize that this is true even if point 2 is somewhere left of 3, which is due to (2.45) and $2 \in \mathcal{A} \setminus \mathcal{B}$ (no overshoot of user 2). In contrast, if $2 \in \mathcal{A} \cap \mathcal{B}$, then, for sufficiently large α, the operating point can be in general any boundary point on the left side of point 3 depending on the choice of the weight vector $\mathbf{b}$ in (2.34).

We point out that irreducibility of $\mathbf{V}$ is a key ingredient in the proof of (2.45) being a sufficient condition for (2.45). In particular, if $\mathbf{V}$ is reducible, then, due to the assumption of individual power constraints, the network may decompose into smaller isolated subnetworks so that each subnetwork is described by an irreducible gain matrix that is a sub-matrix of $\mathbf{V}$. This means that each subnetwork is entirely coupled by interference but the subnetworks are mutually independent of each other. In such a case, it may be easily seen that (2.45) holds if each subnetwork has at least one user whose index belongs to $\mathcal{A} \setminus \mathcal{B}$ and at least one user whose index belongs to $\mathcal{B}$.

2.4.3 A Distributed Algorithm

Now we return to the general problem statement (2.33) where user $k \in \mathcal{A}$ is assigned the utility function $\psi_{\alpha_k}(x)$, $x > 0$, given by (2.3) with $\alpha = \alpha_k \geq 2$. As aforementioned, Proposition 3 straightforwardly extends to this case with $\alpha \geq 2$ defined by (2.37).

The gradient-projection algorithm for the problem (2.33) is

$$s_k(n+1) = \min\Big\{s_k(n) - \delta\nabla_k\tilde{F}_{\boldsymbol{\alpha}}(\mathbf{s}(n)), \log P_k\Big\} \tag{2.51}$$

for each $k \in \mathcal{K}$ with $\mathbf{s}(0) \in \mathrm{S}$. The kth entry of the gradient vector $\nabla_k\tilde{F}_{\boldsymbol{\alpha}}(\mathbf{s})$ is given by

$$\nabla_k\tilde{F}_{\boldsymbol{\alpha}}(\mathbf{s}) = e^{s_k}\big(\phi_k(\mathbf{s}) + S_k(\mathbf{s})\big) \tag{2.52}$$

where (note that $v_{k,k} = 0$)

$$S_k(\mathbf{s}) = -\sum_{l\in\mathcal{K}} v_{l,k}\mathrm{SIR}_l(e^{\mathbf{s}})\phi_l(\mathbf{s}) = \sum_{l\in\mathcal{K}} v_{l,k}|\mathrm{SIR}_l(e^{\mathbf{s}})\phi_l(\mathbf{s})| = \sum_{l\in\mathcal{K}} v_{l,k}m_l(\mathbf{s}) \tag{2.53}$$

since $\phi_l(\mathbf{s})$ is negative and

$$\phi_k(\mathbf{s}) = \begin{cases} \eta_k(\mathbf{s}) & k \in \mathcal{A}\setminus\mathcal{B} \\ \eta_k(\mathbf{s}) + \nu_k(\mathbf{s}) & k \in \mathcal{A}\cap\mathcal{B} \\ \nu_k(\mathbf{s}) & k \in \mathcal{B}\setminus\mathcal{A}. \end{cases}$$

Here we have $\eta_k(\mathbf{s}) = a_k\psi'_{\alpha_k}(g_k(\mathbf{s}))/(\gamma_k I_k(e^{\mathbf{s}}))$ and $\nu_k(\mathbf{s}) = b_k\psi'(\mathrm{SIR}_k(e^{\mathbf{s}}))/I_k(e^{\mathbf{s}})$. Now standard results from convex optimization theory together with Observation 5 lead to the following observation.

Observation 6 *A sequence* $\{\mathbf{s}(n)\}$ *generated by (2.51) converges to* $\tilde{\mathbf{s}}(\boldsymbol{\alpha}) \in \mathrm{S}$ *given by (2.33).*

It is worth pointing out that the gradient-projection algorithm exhibits monotoni-city, which is not guaranteed by the primal-dual algorithm. This may be advantageous in practice where often only few iterations can be performed. The algorithm can be implemented in a distributed manner by means of an adjoint network that is used for a distributed estimation of $S_k(\mathbf{s})$ given by (2.53). In fact, a distributed version of the algorithm is similar to Algorithm 1 of Section 2.3 except that there is no exchange of dual variables, $\tilde{F}_{\boldsymbol{\alpha}}$ defined by (2.34) is used for the termination and $m_k(\mathbf{s})$ is defined in (2.53). In addition, each transmitter, say transmitter $k \in \mathcal{A}$, must inform the corresponding receiver about α_k. For relatively large values of α_k, the step size may have to be chosen relatively small due to rapid variations of the gradient, resulting in a low convergence rate. One possible remedy is to let the QoS users start iterating with small values of $\alpha_k, k \in \mathcal{A}$, and then to increase α_k gradually as the algorithm proceeds until either some predefined maximum value is achieved or the SIR targets are met. As we already mentioned before, each QoS user can increase its $\alpha_k, k \in \mathcal{A}$, separately so that no information exchange to the other users is required.

2.5 Numerical Results

First, we consider a network with $K = 8$ users and a randomly chosen irreducible gain matrix $\mathbf{V} \geq 0$. The weight vector and the SIR targets are chosen to be $\mathbf{w} = \mathbf{1}$ and $\gamma_k = 5, k \in \mathcal{A}$. Each user operates at SNR = 40dB.

Figure 2.8 shows the impact of the parameter $\alpha \geq 1$ on the achievable balanced SIRs (SIR_k/γ_k) and the total throughput for different power control strategies.

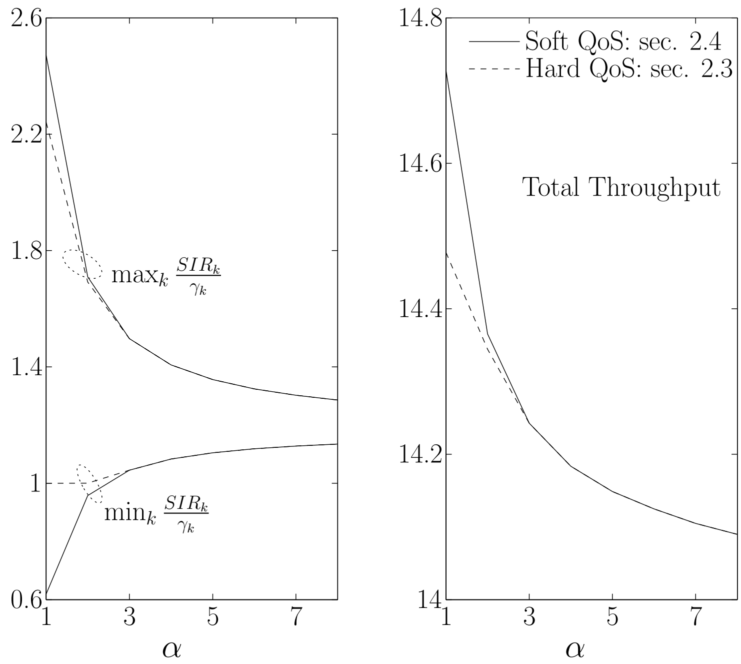

Figure 2.8: The minimum and maximum of SIR_k/γ_k (left picture) and the total throughput (right picture) as functions of α.

The simulation confirms Observation 4 ensuring that the SIR targets are met if they are feasible and α is sufficiently large. This stands in contrast to the case of hard QoS support where the SIR targets are achieved for any choice of $\alpha \geq 1$. We point out that the discrepancy between the rates achieved by different users may be significant for small values of α. As α increases, the fairness performance improves at the expense of the throughput performance. Figure 2.9 depicts exemplarily the convergence behavior of the primal-dual algorithm of Section 2.3 under the assumption that only noisy estimates of m_k and SIR_k (see Algorithm 1) are known [KY03]. The diminishing step size is chosen to be $\delta(n) = C/(n+1)^{0.5}$. For clarity reasons, the figure shows the convergence behavior only for 4 users, whose balanced SIR values in the optimum are depicted by dashed lines.

In order to illustrate Proposition 3 (see Figure 2.10), let us consider a network with $K = 5$ users so that 2 users are pure QoS users ($\mathcal{A} \setminus \mathcal{B}$), 2 users are both best-effort and QoS users ($\mathcal{A} \cap \mathcal{B}$) and 1 user is a best-effort user only ($\mathcal{B} \setminus \mathcal{A}$). As α increases, the SIR of the QoS users in $\mathcal{A} \setminus \mathcal{B}$ approach their SIR targets according to (2.45) and (2.46). Once the SIR targets are met for some sufficiently large α, the available remaining power resources are allocated to the users in $\mathcal{A} \cap \mathcal{B}$ and $\mathcal{B} \setminus \mathcal{A}$ so as to maximize the aggregate utility function, which is the second summand in (2.36). Figure 2.11 shows exemplarily the convergence behavior of the algorithm in Section 2.4.3. In this simulation, in order to

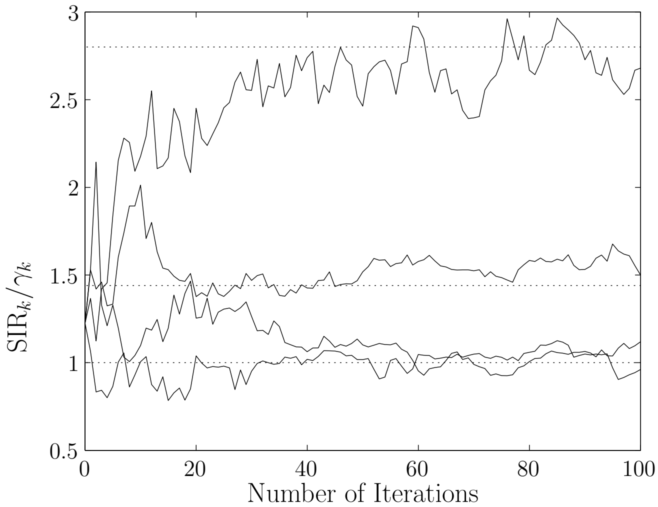

Figure 2.9: A convergence behavior of the primal-dual algorithm in the case of noisy estimates in the adjoint network.

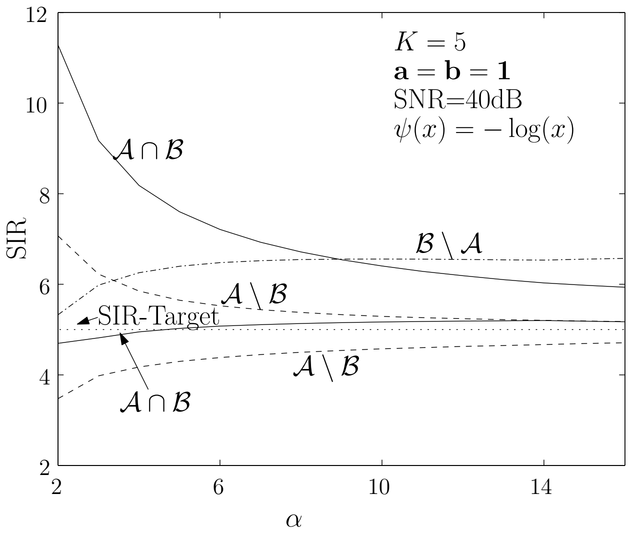

Figure 2.10: The SIR performance over α under a power vector given by (2.35).

improve the convergence speed, α is increased gradually as the algorithm proceeds.

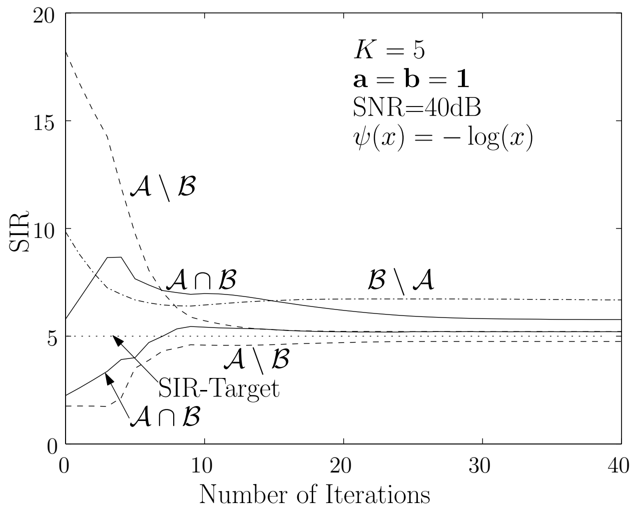

Figure 2.11: Convergence behavior of (2.51) with α being increased gradually as the algorithm proceeds.

Remark: We point out that in practical systems, increasing the SIR above a certain threshold corresponding to the maximum feasible data rate will not improve the rate performance (but will in general reduce the bit error rate). Thus, it may be reasonable to impose additional upper bounds on the SIR values. The problem is not addressed here but the reader will realize that our results can be easily extended to incorporate any upper bounds on the SIR values. Such a problem generalization implies a slight modification of the primal-dual algorithms presented in Section 2.3. In fact, augmenting the new SIR constraints to an associated Lagrangian function would only increase the dimensionality of the dual variable. Alternatively, we could follow similar ideas to those in Section 2.4 to make large SIR values less "attractive" by introducing a suitable penalty function.

2.6 Appendix

A. Distributed Feedback Scheme

In this section we describe a distributed feedback scheme which realizes the power control iteration (2.8). It has been introduced in [SWB07]. The ideas are however applicable to other algorithms and are utilized later in the sections 2.3 (Alg. 1), 2.4.3, 3.1.4, 3.1.6 (Alg. 3) and 3.2.4 (Alg. 4) to develop other distributed gradient-based algorithms as well as distributed primal-dual algorithms. Before we explain the scheme in more detail we first define our understanding of a distributed network organization and optimization.

Definition 2 *We understand a solution of a network optimization problem as a decentralized or distributed one if it consists of decoupled actions executed seperately for each link. We require a coarse synchronization between the links to be able to measure the interference at each link. Further, we assume that the local knowledge at each link is provided by a low rate signaling and feedback channel on each link.*

The major achievement of distributed feedback scheme was to parallelize the computation of $\nabla F_e(\mathbf{s})$ in such a way that each node can calculate $\nabla_k F_e(\mathbf{s}) = \frac{\partial F_e}{\partial s_k}(\mathbf{s})$ for all $k \in \mathcal{K}$, without requiring extensive communication between nodes. More precisely, the algorithm is separated into K local algorithms operating concurrently at different transmitter–receiver pairs.

Consider the nth iteration in (2.8) and assume that $\mathbf{p} = \mathbf{p}(n) = e^{\mathbf{s}(n)}$ is the nth power vector. It is easy to see, that $\nabla_k F_e(\mathbf{s}) = \nabla_k F(\mathbf{p})e^{s_k}$. Thus, in what follows, we focus on $\nabla F(\mathbf{p})$. Considering (2.9) reveals that the gradient vector can be rewritten as follows

$$\nabla F(\mathbf{p}) = \underbrace{(\mathbf{I} + \boldsymbol{\Gamma}(\mathbf{p}))\mathbf{g}(\mathbf{p})}_{\boldsymbol{\eta}(\mathbf{p})<0} - \underbrace{(\mathbf{I} + \mathbf{V}^T)\boldsymbol{\Gamma}(\mathbf{p})\mathbf{g}(\mathbf{p})}_{\boldsymbol{\theta}(\mathbf{p})<0} = \boldsymbol{\eta}(\mathbf{p}) - \boldsymbol{\theta}(\mathbf{p})\,, \tag{2.54}$$

where $\boldsymbol{\eta}(\mathbf{p}) = (\eta_1(\mathbf{p}), \ldots, \eta_K(\mathbf{p}))$, $\boldsymbol{\theta}(\mathbf{p}) = (\theta_1(\mathbf{p}), \ldots, \theta_K(\mathbf{p}))$, $\mathbf{g}(\mathbf{p}) := (\phi_1(\mathbf{p}), \ldots, \phi_K(\mathbf{p}))$ with $\phi_k(\mathbf{p}) = w_k\psi'(\mathrm{SIR}_k(\mathbf{p}))/I_k(\mathbf{p})$ defined by (2.10) and

$$\boldsymbol{\Gamma}(\mathbf{p}) := \mathrm{diag}\big(\mathrm{SIR}_1(\mathbf{p}), \ldots, \mathrm{SIR}_K(\mathbf{p})\big)\,.$$

So the computation of the gradient $\nabla_k F(\mathbf{p})$ at the kth transmitter consists of a local part

$$\eta_k(\mathbf{p}) = \big(1 + \mathrm{SIR}_k(\mathbf{p})\big)\phi_k(\mathbf{p}) = \big(1 + \mathrm{SIR}_k(\mathbf{p})\big)\frac{w_k\psi'(\mathrm{SIR}_k(\mathbf{p}))}{I_k(\mathbf{p})}\,. \tag{2.55}$$

and a global part

$$\theta_k(\mathbf{p}) = \mathrm{SIR}_k(\mathbf{p})\phi_k(\mathbf{p}) + \sum_{l\in\mathcal{K}} v_{l,k}\mathrm{SIR}_l(\mathbf{p})\phi_l(\mathbf{p}). \tag{2.56}$$

Assuming that the transmitters are coarsely synchronized and transmit a pseudorandom training sequence each receiver may estimate the signal-to-interference ratio and feed the estimate back to the corresponding transmitter using a reliable low-rate feedback channel. Based on this information and assuming that the kth transmitter–receiver pair can also extract the interference $I_k(\mathbf{p})$, each transmitter k is able to calculate the estimate of $\phi_k(\mathbf{p}) < 0$ and thus the local part $\eta_k(\mathbf{p})$.

However, the main ingredient to the scheme and the key to its distributed character is the concept of the so-called adjoint network that allows to determine the global part $\theta_k(\mathbf{p})$.

Definition 3 (Adjoint Network [SWB07, SWB09]) *Consider an arbitrary wireless network with K (logical) links and the gain matrix $\mathbf{V}$. Let us call it the primal network. Then, a network with K (logical) links and the gain matrix $\mathbf{U} \in \mathbb{R}_+^{K \times K}$ is said to be adjoint to the primal network if $\mathbf{U} = \mathbf{V}^T$.*

We see that $\boldsymbol{\theta}(\mathbf{p})$ in (2.56) results from the multiplication of the vector $\boldsymbol{\Gamma}(\mathbf{p})\mathbf{g}(\mathbf{p})$ with $(\mathbf{I} + \mathbf{V}^T)$. This suggests that the entries of the vector $\boldsymbol{\theta}(\mathbf{p})$ may be made available to the nodes in the network by transmitting appropriately scaled pilot symbols over an adjoint network. The key feature of this concept is that one can mimic the corresponding adjoint network:

Observation 7 *Given a network with interference matrix $\mathbf{V} = (V_{k,l})$ as is used in (1.1), an adjoint network on the same link set $\mathcal{K}$ is obtained by*

i) applying a reversed network, where the roles of transmitters and receivers on each link in a primal network are reversed. (In a reversed network, link $k \in \mathcal{K}$ is a link between the kth receiver and the kth transmitter in the primal network.);

ii) inverting the channel multiplying the transmit symbols of each transmitter $k \in \mathcal{K}$ in the reversed network by $1/\sqrt{V_{k,k}}$.

Using the definitions and observations above the authors of [SWB07, SWB09] proposed the following distributed handshake protocol. The basic idea is to use the primal network and the adjoint network alternately to obtain an estimate of $\nabla_k F$ at the kth transmitter in the primal network. It is assumed that the kth transmitter in the primal network in advance informs the kth receiver about its weight w_k and ψ'. This must be done only once before starting the iteration process. Now, in Algorithm 2 the protocol is summarized.

B. Some auxiliary results

Lemma 5 *$\mathbf{X} \geq 0$ is irreducible if and only if, for each (i, j) with $1 \leq i, j \leq n$, there is $k \geq 0$ such that $x_{i,j}^{(k)} := (\mathbf{X}^k)_{i,j} > 0$.*

Proof: See [Sen81]. ∎

Theorem 1 *Let $\mathbf{X} \geq 0$ be arbitrary and let $\mu > 0$ be any scalar. A necessary and sufficient condition for a solution $\mathbf{p} \geq 0, \mathbf{p} \neq \mathbf{0}$ to $(\mu\mathbf{I} - \mathbf{X})\mathbf{p} = \mathbf{b}$ to exist for any $\mathbf{b} > 0$ is that $\mu > r = \rho(\mathbf{X})$. In this case, there is only one solution $\mathbf{p}$, which is strictly positive and given by $\mathbf{p} = (\mu\mathbf{I} - \mathbf{X})^{-1}\mathbf{b}$.*

Proof: See [SWB09]. ∎

Algorithm 2 Distributed gradient projection algorithm [SWB07, SWB09]

Require: $\mathbf{w} > 0, n = 0, \mathbf{s}(0) \in \mathrm{S}$, constant or non-increasing step size sequence $\{\delta(n)\}_{n\in\mathbb{N}_0}$.

Ensure: $\mathbf{s} \in \mathrm{S}$

1: **repeat**

2: Concurrent transmission of training sequences at powers $(p_1(n), \ldots, p_K(n))$.

3: Receiver-side estimation of the signal-to-interference ratios and interferences. Based on these estimations, each receiver calculates $\phi_k(\mathbf{p}(n)), k \in \mathcal{K}$.

4: All receivers feed the signal-to-interference ratios back to the corresponding transmitters using a per-link control channel. Transmitter-side computation of $\phi_k(\mathbf{p}(n))$, and then $\eta_k(\mathbf{p}(n))$ for each $k \in \mathcal{K}$.

5: Concurrent transmission of sequences of zero-mean independent symbols X_k with

$$E[|X_k|^2] = |\mathrm{SIR}_k(\mathbf{p}(n)) \cdot \phi_k(\mathbf{p}(n))|,\ k \in \mathcal{K}$$

over the adjoint network. Note that the transmission over the adjoint network involves channel inversion.

6: Transmitter-side estimation of the received power and subtraction of noise variances from the estimates to obtain $\theta_k(\mathbf{p}(n))$. Since $\eta_k(\mathbf{p}(n))$ and $\theta_k(\mathbf{p}(n))$ are known at transmitter k, the transmitter computes

$$\begin{aligned}\nabla_k \hat{F}(\mathbf{p}(n)) &= \eta_k(\mathbf{p}(n)) - \theta_k(\mathbf{p}(n)) \\ &= \phi_k(\mathbf{p}(n)) - (\mathbf{V}^T \mathbf{\Gamma}(\mathbf{p}(n))\mathbf{g}(\mathbf{p}(n)))_k\end{aligned}$$

where we assumed that all the variables have been estimated perfectly.

7: Update of transmit powers according to (2.8) with $\mathbf{s}(n) = \log(\mathbf{p}(n))$

8: $n = n + 1$

9: **until** $|F_e(\mathbf{s}(n)) - F_e(\mathbf{s}(n-1))| < \epsilon$

10: $\mathbf{s} = \mathbf{s}(n)$

Theorem 2 *Let* $\mathbf{X} \geq 0$ *be irreducible and let* $q(\lambda) = \prod_{k=1}^{s}(\lambda - \lambda_k(\mathbf{X}))^{m_k}$ *be its minimal polynomial with the eigenvalues ordered so that* $|\lambda_1(\mathbf{X})| \geq |\lambda_2(\mathbf{X})| \geq \cdots \geq |\lambda_s(\mathbf{X})|$. *Then, there exist matrices* $\mathbf{Z}_{k,j} \in \mathbb{R}^{n\times n}, 1 \leq j \leq m_k, 2 \leq k \leq s$, *such that*

$$(\mu\mathbf{I} - \mathbf{X})^{-1} = \frac{1}{\mu - \rho(\mathbf{X})}\mathbf{p}\mathbf{q}^T + \sum_{k=2}^{s}\sum_{j=1}^{m_k}\frac{(j-1)!}{(\mu - \lambda_k(\mathbf{X}))^j}\mathbf{Z}_{k,j} \tag{2.57}$$

where $\mathbf{q} > 0$ *and* $\mathbf{p} > 0$ *with* $\mathbf{q}^T\mathbf{p} = 1$ *are left and right positive eigenvectors of* $\mathbf{X}$ *associated with* $\rho(\mathbf{X})$.

Proof: See [Lan69]. ■

Theorem 3 *Suppose that* $\bar{\mathrm{U}} = \{\mathbf{X} : (\mu\mathbf{I} - \mathbf{X})^{-1}\mathbf{b} \in \mathrm{P}\}$ *is not an empty set for a given bounded set* $\mathrm{P} \subset \mathbb{R}^n_{++}$, *a positive vector* $\mathbf{b}$ *and some constant* $\mu > 0$. *Then, there exists a constant* $c = c(\mu, \mathbf{b}, \mathrm{P}) > 0$ *such that* $0 \leq \rho(\mathbf{X}) \leq \mu - c$ *for all* $\mathbf{X} \in \bar{\mathrm{U}}$. *So, we have* $\sup_{\mathbf{X}\in\bar{\mathrm{U}}} \rho(\mathbf{X}) < \mu$.

Proof: Let $\mathbf{X} \in \bar{\mathrm{U}}$ be arbitrary, and let $\mathbf{x} \geq 0$ be any right eigenvector of $\mathbf{X}$ associated with $\rho(\mathbf{X})$ and normalized such that $\|\mathbf{x}\|_1 = \mathbf{1}^T\mathbf{x} = 1$. Note that by Theorem 1, we must have $0 \leq \rho(\mathbf{X}) < \mu$. If $\rho(\mathbf{X}) = 0$, then the theorem holds trivially. So, assume that $\rho(\mathbf{X}) \in (0, \mu)$. As $\mathbf{b}$ is positive, it is clear that there is a constant $c_1 > 0$ such that $\mathbf{x} \leq \mathbf{b}/c_1$. Moreover, from the boundedness of the set P, we can conclude that there is a constant $c_2 > 0$ such that $\mathbf{1}^T(\mu\mathbf{I} - \mathbf{X})^{-1}\mathbf{b} \leq c_2$. Considering these two inequalities yields

$$0 \leq \mathbf{1}^T(\mu\mathbf{I} - \mathbf{X})^{-1}\mathbf{p} \leq \frac{1}{c_1}\mathbf{1}^T(\mu\mathbf{I} - \mathbf{X})^{-1}\mathbf{b} \leq \frac{c_2}{c_1}\,.$$

On the other hand, by the Neumann series, one obtains

$$\begin{aligned}\mathbf{1}^T(\mu\mathbf{I} - \mathbf{X})^{-1}\mathbf{x} &= \frac{1}{\mu}\mathbf{1}^T\big(\mathbf{I} - \frac{1}{\mu}\mathbf{X}\big)^{-1}\mathbf{x} = \frac{1}{\mu}\mathbf{1}^T\sum_{l=0}^{\infty}\big(\frac{1}{\mu}\mathbf{X}\big)^l\mathbf{x} \\ &= \frac{1}{\mu}\sum_{l=0}^{\infty}\big(\frac{1}{\mu}\rho(\mathbf{X})\big)^l\mathbf{1}^T\mathbf{x} = \frac{1}{\mu}\frac{1}{1 - \rho(\mathbf{X})/\mu} = \frac{1}{\mu - \rho(\mathbf{X})}\,.\end{aligned}$$

Combining this identity with the previous inequality yields $1/(\mu - \rho(\mathbf{X})) \leq c_2/c_1 > 0$ or, equivalently, $\rho(\mathbf{X}) \leq \mu - c_1/c_2$. Since $\mathbf{X}$ has been chosen to be any matrix in $\bar{\mathrm{U}}$, this proves the theorem with $c = c_1/c_2 \in (0, \mu)$. ■

3 Resource Allocation in Multiantenna Wireless Networks

Beamforming provides in addition to power control another powerful mechanism to manage the interference in wireless networks. In order to ensure a high utilization of wireless resources, transmit powers and beamformers should be optimized jointly to exploit interdependencies between them. As is widely known, the overall network can be optimized with respect to different optimization goals. In general, there exist two main approaches attracting interest during the last years. The classical QoS-based approach aims at satisfying a certain quality-of-service (QoS) requirement with minimum power. To circumvent the feasibility problem a related approach is to solve the so-called max-min SIR-balancing or relative max-min problem. In contrast to this stands the utility-based resource allocation problem that has been considered in more detail in the previous chapter with power as the resource to be optimized. Remember, here the network operator aims at optimizing a weighted aggregate utility so as to maximize the overall network performance. Throughput, fairness, or a tradeoff between throughput and fairness can be achieved by choosing the utility function adequately.

Optimizing over the joint space of powers and receive beamformers is a more intricate task than optimizing over powers only. Even more, in addition to efficiently managing wireless resources, it is a challenging task to distributedly assign these resources, especially in the case of the non-differentiable relative max-min objective function. Further, in practice we need to apply stochastic algorithms that deal with real-world noisy measurements and estimations. Based on the above discussion, in this chapter we discuss the following two topics to better understand the interdependencies between powers and beamformers as well as the optimization approaches:

(i) Optimization of an aggregate utility function jointly over powers and receive beamformers in real-world distributed wireless networks (section 3.1).

(ii) Solve the max-min SIR-balancing problem by optimizing a certain aggregate utility function over the joint space of powers and receivers (section 3.2).

3.1 Joint Utility-Based Power Control and Receiver Control

As mentioned in the introduction, in this section we focus on the following problem: maximize an aggregate utility jointly over powers and receive beamformers in real-world distributed wireless networks.

Existing works on power control, distributed power control and the topic of noisy measurements are summarized at the beginning of chapter 2. For the sake of completeness, we point out that the following section is an extension of the work on utility-based power control in [SWB07], where the authors propose a distributed utility-based power control algorithm for general wireless networks, touch the problem of stochastic approximation and show how to cope with it in practice.

Independently from and simultaneously to our work the authors of [LHC07] proposed a distributed utility-based joint power control and receive beamforming algorithm for cellular uplinks applying the scheme of [HRCW08]. Apart from that, so far most work on joint power control and beamforming focusses on the so-called max-min SIR-balancing and its related problem. For example in [RTL98, SB04, SB06] the duality between uplink and downlink channels is exploited. Another strategy shows that the problem can be embedded in semidefinite and conic optimization programs [BO01, WES06]. The work presented in [3] (see also section 3.2) extends the work of [SB06] by solving the max-min SIR-balancing problem under general power constraints. However, apart from [SWB09] the above set of publications considers only the deterministic case. First works incorporating imprecise knowledge of received waveforms include [HMV95, UY98, Var99]. Recently, stochastic algorithms for joint QoS-based power control and receive beamforming and their convergence analysis have been proposed by [LUE05, DV07].

To be more precise, in this section we consider the problem of joint power control and receive beamforming in order to maximize a certain aggregate utility function that represents the QoS attained and is a function of the SIR. In section 3.1.2 we first reformulate the joint power and receiver control problem as a pure power control problem under implicit optimal linear receivers. However, in contrast to the pure utility-based power control problem [SWB07], it is not known which class of utility functions allows a convex formulation of this joint optimization problem and thus enables an efficient global solution in distributed wireless networks. In particular, in the case of the lo-garithmic function, the aggregate utility function appears to have relatively many local maxima. Though, if we confine our attention to utility functions whose relative conca-vity is larger than that of the logarithmic function, numerical experiments suggest that then the proposed algorithm may converge to a global maximum for a large set of initial SIRs. In section 3.1.3 we start by assuming perfect synchronization. Due to the fact that an optimal linear receiver can be obtained in closed-form solution for any power vector, we analyze the power control problem under implicit optimal linear receivers. However, an efficient implementation of the equivalent gradient projection algorithm is notoriously difficult to achieve in decentralized wireless networks. Thus, in section 3.1.4 we decompose the problem into two coupled subproblems and propose an alternating algorithm that converges to a stationary point. As already mentioned, in real-world wireless networks noisy measurements and estimations occur. We embed the proposed alternating algorithm into the framework of stochastic approximation, thus discussing practical implementation aspects. Then, in section 3.1.5 we investigate the convergence properties of the proposed stochastic algorithm by simulations. Finally and for the sake of completeness, in section 3.1.6 we compare our joint utility-based approach with the recent pure adaptive beamforming scheme via bi-directional training [SBH10].

The results presented in this section appeared in part in [10, 4]. Parts of these works also appeared in [SWB09].

3.1.1 Problem Statement

In the following we consider a multiple-antenna wireless network with K users or active links as introduced in section 1.4. For simplicity we assume that the transmit powers of the users are subject to individual power constraints $P_1, \ldots, P_K > 0$ so that $\mathbf{p} \in \mathrm{P}$ must hold, where $\mathrm{P} = (\mathbf{p} \in \mathbb{R}_+^K : \forall_k p_k \leq P_k) \subset \mathbb{R}_+^K$. As introduced in section 1.4.5 the receivers are collected in a matrix $\mathbf{C} \in \mathrm{C}$ where $\mathrm{C} = \{\mathbf{C} = (\mathbf{c}_1, \ldots, \mathbf{c}_K) \in \mathbb{C}^{M \times K} : \forall_k \|\mathbf{c}_k\|_2 = 1\}$ denotes the set of all receive beamforming matrices. Each column of $\mathbf{C}$ is normalized meaning that each receiver is a vector on a unit sphere. Also note that both P and C are compact sets, so is also their Cartesian product $\mathrm{P} \times \mathrm{C}$. Finally, we define $\mathrm{P}_+ = \mathrm{P} \cap \mathbb{R}_{++}^K$. In words, P_+ is the set of positive power vectors satisfying the power constraints. The main figure of merit is the SIR at the output of each receiver and has been defined in (1.6) assuming perfectly synchronized users. A crucial property of receive beamforming is that the kth receiver only influences the SIR of the k-th user. Thus the SIR of user k depends only on the kth receive beamformer. In contrast, transmit powers and also transmit filters have in general impact on all users. That is why we write $\mathrm{SIR}_k(\mathbf{p}, \mathbf{c}_k)$.

One way to control both transmit powers of the users and their receive beamformers is to apply a utility-based framework. Let $\Psi(\mathrm{SIR}_k(\mathbf{p}, \mathbf{c}_k))$ be the utility of user k that represents the QoS level or the degree of user satisfaction of user k under power vector $\mathbf{p}$ and receive beamformer $\mathbf{c}_k$. Throughout this section we assume that Ψ satisfies (A.1)-(A.3). Now, the joint utility-based power control and receive beamforming problem can be written as follows. Given any weight vector $\mathbf{w} = (w_1, \ldots, w_K) > 0$, we search for a power vector $\mathbf{p}^* \in \mathrm{P}$ and a beamforming matrix $\mathbf{C}^* = (\mathbf{c}_1^*, \ldots, \mathbf{c}_K^*)$ such that

$$(\mathbf{p}^*, \mathbf{C}^*) = \underset{(\mathbf{p}, \mathbf{C}) \in \mathrm{P} \times \mathrm{C}}{\arg\max}\ G(\mathbf{p}, \mathbf{C}) \tag{3.1}$$

where the maximum is assumed to exist[1] and

$$G(\mathbf{p}, \mathbf{C}) = \sum_{k\in\mathcal{K}} w_k \Psi\big(\mathrm{SIR}_k(\mathbf{p}, \mathbf{c}_k)\big)\,. \tag{3.2}$$

We know that [SWB07], for any fixed $\mathbf{C} \in \mathrm{C}$, $G_e(\mathbf{s}, \mathbf{C}) := G(e^{\mathbf{s}}, \mathbf{C})$ is concave in the logarithmic power vector $\mathbf{s} := \log \mathbf{p} \in \mathrm{S}$ with $\mathbf{p} \in \mathrm{P}_+$ and $\mathrm{S} := \{\mathbf{s} = \log \mathbf{p} : \mathbf{p} \in \mathrm{P}_+\} \subset \mathbb{R}^K$. Therefore, for $G_e(\mathbf{s}, \mathbf{C})$ to be concave in $\mathbf{s}$, it is sufficient that $\mathrm{SIR}_k(e^{\mathbf{s}})$ is a log-concave function of $\mathbf{s} \in \mathbb{R}^K$. Now the question is what happens if we use this class of utility functions $\Psi(x)$ defined by (2.3) in the joint power control and receive beamforming problem (3.1).

3.1.2 Problem Properties

First of all we point out that the joint problem (3.1) is equivalent to a power control problem once the optimal receivers are known explicitly for a given power vector. Under optimal receivers we understand the receivers that maximize the SIR. A crucial ingredient to reason this statement comes from the fact, that the kth SIR depends only on the kth receiver. Because each receiver has a unit norm, it is a vector on a unit sphere of $\mathbb{C}^M$ denoted by C^{M-1}, such that $\mathbf{c}_k \in \mathrm{C}^{M-1}, k \in \mathcal{K}$. Further, let us assume that

(A.11) $\mathbb{R}_+^K \times \mathrm{C}^{M-1} \to \mathbb{R}_+^K : (\mathbf{p}, \mathbf{c}_k) \mapsto \mathrm{SIR}_k(\mathbf{p}, \mathbf{c}_k)$ is continuous,

which is reasonable in practice. From this it follows immediately that the SIR has a maximum on the compact set $\mathrm{P} \times \mathrm{C}^{M-1}$. Thus we can observe:

Observation 8 *For any fixed $\mathbf{p} \geq 0$, there exists $\mathbf{c}_k^*(\mathbf{p})$ such that*

$$\mathrm{SIR}_k^*(\mathbf{p}) := \max_{\mathbf{c}_k \in \mathrm{C}^{M-1}} \mathrm{SIR}_k(\mathbf{p}, \mathbf{c}_k). \tag{3.3}$$

Now we show that the joint problem (3.1) can be written as a pure power control problem under optimal adaptive receivers. This is because

$$\begin{aligned}
\max_{(\mathbf{p},\mathbf{C})\in \mathrm{P}\times\mathrm{C}} G\big(\mathbf{p}, \mathbf{C}\big) &\overset{(a)}{=} \max_{\mathbf{p}\in\mathrm{P}} \sum_{k\in\mathcal{K}} \max_{\mathbf{c}_k\in\mathrm{C}^{M-1}} w_k \Psi(\mathrm{SIR}_k(\mathbf{p}, \mathbf{c}_k)) \\
&\overset{(b)}{=} \max_{\mathbf{p}\in\mathrm{P}} \sum_{k\in\mathcal{K}} w_k \Psi\Big(\max_{\mathbf{c}_k\in\mathrm{C}^{M-1}} \mathrm{SIR}_k(\mathbf{p}, \mathbf{c}_k)\Big) \\
&\overset{(c)}{=} \sum_{k\in\mathcal{K}} w_k \Psi\Big(\mathrm{SIR}_k^*(\mathbf{p})\Big) \\
&= \max_{\mathbf{p}\in\mathrm{P}} G\big(\mathbf{p}, \mathbf{C}^*(\mathbf{p})\big),
\end{aligned}$$

[1] Since the noise variance is strict positive, standard arguments can be used to show that with our choice of the utility functions, the maximum exists.

where (a) follows from the fact, that the kth receiver impacts only the kth SIR, (b) is an immediate consequence of the strictly increasing function $\Psi(x), x > 0$, and the positive vector $\mathbf{w}$ and (c) results from (3.3). Remind, that $\mathbf{C}^*(\mathbf{p}) = (\mathbf{c}_1^*(\mathbf{p}), \ldots, \mathbf{c}_K^*(\mathbf{p}))$ denotes an optimal receiver matrix for a given power vector $\mathbf{p} \geq 0$. We summarize the result in the following lemma.

Lemma 6 *For any fixed $\mathbf{p} \geq 0$, the receive beamforming vectors maximizing the individual SIRs (or solves (3.3)) solve at the same time the original utility-based problem*

$$\max_{\mathbf{C}\in\mathrm{C}} \sum_{k\in\mathcal{K}} w_k \Psi(\mathrm{SIR}_k(\mathbf{p}, \mathbf{c}_k)).$$

Now, using the substitution $\mathbf{s} = \log \mathbf{p}, \mathbf{p} \in \mathrm{P}_+$, (in accordance with the power control problem of [SWB07]) and defining

$$F(\mathbf{s}) := \sum_{k\in\mathcal{K}} w_k \Psi\Big(\mathrm{SIR}_k^*(e^{\mathbf{s}})\Big) \tag{3.4}$$

it follows that a solution to (3.1) is any pair $(\mathbf{p}^*, \mathbf{C}^*) \in \mathrm{P}_+ \times \mathrm{C}$ given by $\mathbf{p}^* = e^{\mathbf{s}^*}$ and $\mathbf{C}^* = \mathbf{C}^*(\mathbf{p}^*)$ where

$$\mathbf{s}^* = \arg\max_{\mathbf{s}\in\mathrm{S}} F\big(\mathbf{s}\big)\,. \tag{3.5}$$

In words, the problem reduces to a power control problem except that now each SIR is assumed to attain its maximum over all receive beamformers. Since $\mathbf{c}_k^*(\mathbf{p})$ maximizes the SIR for a given power vector $\mathbf{p}$, it indeed minimizes the interference at the receiver of user k.

Convexity Discussion

Unfortunately, so far it is not known which class of utility functions allows a convex formulation of the problem (3.5), such that a global solution can be found using standard optimization methods. We assume a perfectly synchronized network with two links such that the SIR_k^* in (3.4) can be expressed by (1.12). For both exemplary utility functions $\Psi(x) = \log(x)$ and $\Psi(x) = \log[x/(1+x)]$, $x > 0$ it can be easily shown by a counter example that the problem is not concave in general (see section 3.3 Appendix A). However, by simulations we have observed that the standard gradient projection algorithm deduced in 3.1.3 for this power control problem converges to a local maximum in significantly fewer cases with increasing concavity of the utility function.

In economics, the quantity

$$r(x, \Psi) = -\frac{x\Psi''(x)}{\Psi'(x)} \geq 0 \tag{3.6}$$

is known as the coefficient of relative risk aversion [MCWG95]. Here $\Psi'(x) > 0$ and

$\Psi''(x) < 0, x > 0$, denote the first and second derivatives of Ψ, respectively. $r(x, \Psi)$ is used to measure the relative concavity of $\Psi(x)$. The larger the value of $r(x, \Psi) \geq 0$ is, the larger is the relative concavity of $\Psi(x)$ at $x > 0$, and therefore a better fairness performance (at the cost of the throughput performance) can be expected. For the logarithmic function $\Psi(x) = \log(x), x > 0$ we have $r(x, \Psi) = 1, x > 0$, for $\Psi(x) = \log[x/(1+x)],\ x > 0$ it is $r(x, \Psi) = (1+2x)/(1+x) \in (1, 2)$ while for $\Psi_\alpha(x) := \frac{x^{1-\alpha}}{1-\alpha},\ \alpha \geq 2, \alpha \in \mathbb{N}$ it can be easily verified that $r(x, \Psi_\alpha) = \alpha$.

Further, for any fixed $\mathbf{C} \in \mathrm{C}$ we observe the following:

Observation 9 *$r(x, \Psi) \geq 1$ if and only if $\Psi(e^x)$ is concave on $\mathbb{R}$.*

Proof: Since $\Psi(x), x > 0$, is twice continuously differentiable, $\Psi(e^x), x \in \mathbb{R}$, is concave if and only if $e^x\big(\Psi''(e^x)e^x + \Psi'(e^x)\big) \leq 0, x \in \mathbb{R}$. This in turn holds if and only if $\Psi''(x)x + \Psi'(x) \leq 0$, which is equivalent to $r(x, \Psi) \geq 1$ since $\Psi'(x) > 0$ for all $x > 0$. ■

However, as we have seen by the counter example the problem (3.5) is not concave. Yet, numerical experiments with the utility function $\Psi(x) = \log[x/(1+x)],\ x > 0$ suggest that in this case, the gradient projection algorithm (see Section 3.1.3) converges to a global maximum for a relatively large set of initial SIR values compared to the utility function $\Psi(x) = \log(x)$. Now, an interesting problem is whether a global convergence (if not for all starting points, then at least for most of them) of the gradient projection algorithm can be achieved by requiring that $r(x, \Psi) \geq c, x > 0$, for some sufficiently large constant $c \geq 2$. Remember, increasing the constant c leads to utility functions with larger relative concavities. In particular, as shown below, if there is a utility function for which each addend in (3.4) is concave on $\mathbb{R}^K$, then $F(\mathbf{s})$ is concave for all utility functions with a larger coefficient $r(x, \Psi)$.

Observation 10 *Let $g : \mathbb{R}_{++} \to \mathrm{Q}_1$ be any utility function for which (3.6) holds, and suppose that each addend in $F(\mathbf{s})$ with $\Psi(x) = g(x), x > 0$, is concave on $\mathbb{R}^K$. Then, $F(\mathbf{s})$ with $\Psi(x) = f(x), x > 0$, is concave for any utility function $f : \mathbb{R}_{++} \to \mathrm{Q}_2$ such that $r(x, g) \leq r(x, f)$ for all $x > 0$.*

Proof: Since g and f are bijective utility functions, there is a twice continuously differentiable and strictly increasing function $h : \mathrm{Q}_1 \to \mathrm{Q}_2$ such that $f(x) = h(g(x)), x > 0$. So, the observation follows if $h(x)$ is concave. Considering the fact that $f'(x) = h'(g(x))g'(x)$ with $g'(x) > 0$ for all $x > 0$, the second derivative of $f(x)$ yields $f''(x) = h''(g(x))(g'(x))^2 + f'(x)g''(x)/g'(x), x > 0$. By $r(x, g) \leq r(x, f)$, we have $f''(x)g'(x) \leq f'(x)g''(x), x > 0$, so that $f''(x) \geq h''(g(x))(g'(x))^2 + f''(x), x > 0$. This implies that $h''(g(x)) \leq 0, x > 0$, and hence one obtains $h''(x) \leq 0, x \in \mathrm{Q}_1$ due to the bijectivity of g. ■

Applying this observation to the class in (2.3) reveals that if there was some $\alpha \geq 2$ such that $\Psi_\alpha(\mathrm{SIR}_k^*(e^{\mathbf{s}}))$ is concave on $\mathbb{R}^K$ for each $k \in \mathcal{K}$, then the problem (3.5) would be a convex problem for all $\Psi_{\alpha'}(x)$ with $\alpha \leq \alpha'$. Then we would be able to arbitrarily close approximate the max-min fair power allocation for any power constraints. In detail, from [MW00, Lemma 3]) we may deduce that $\mathbf{p}^* = e^{\mathbf{s}^*}$ with $\Psi(x) = \Psi_\alpha(x)$ converges to the max-min power allocation as α tends to infinity. Moreover, for every α, $\nabla F(\mathbf{s})$ exists and is continuous on $\mathbb{R}^K$ so that efficient gradient projection algorithms could be used to approximate the max-min power allocation for any power constraints if the algorithms were global convergent for some sufficiently large α (as discussed before).

3.1.3 Centralized Algorithm in Case of Perfect Synchronization

In this section, we derive a gradient projection algorithm for the problem (3.5) and prove its convergence. To this end, let us first identify optimal receive beamformers. As has been shown, an optimal receive beamformer of user k is exactly that beamformer for which the kth SIR attains its maximum. Hence, under the assumption of perfect synchronization it follows from (1.11) and (1.12), that the power control problem with implicit optimal linear receivers becomes

$$\begin{aligned}\mathbf{s}^* &= \arg\max_{\mathbf{s}\in\mathrm{S}} F(\mathbf{s}) \\ &= \arg\max_{\mathbf{s}\in\mathrm{S}} \sum_{k\in\mathcal{K}} w_k \Psi\Big(e^{s_k}\mathbf{b}_k^H \mathbf{Z}_k^{-1}\big(e^{\mathbf{s}}\big)\mathbf{b}_k\Big)\end{aligned} \tag{3.7}$$

and ($\mathbf{s}^* = \log \mathbf{p}^*$)

$$\mathbf{C}^* = \Big(a_1 \mathbf{Z}_1^{-1}(\mathbf{p}^*)\mathbf{b}_1, \dots, a_K \mathbf{Z}_K^{-1}(\mathbf{p}^*)\mathbf{b}_K\Big) \in \mathrm{C} \tag{3.8}$$

with appropriately chosen constants $a_1, \dots, a_K > 0$ and $\mathbf{Z}_k$ defined by equation (1.10).

If we assume the utility functions $\Psi(x)$ of (2.3), then $F(\mathbf{s})$ in (3.7) can be written using the inverse of

$$\mathbf{Z}(\mathbf{p}) = \mathbf{Z}_k(\mathbf{p}) + p_k \mathbf{b}_k \mathbf{b}_k^H$$

which is independent of the index k. Indeed, by the Sherman-Morrison formula [Mey00], it follows that

$$\mathbf{Z}_k^{-1}(\mathbf{p}) = \mathbf{Z}^{-1}(\mathbf{p}) + \frac{\mathbf{Z}^{-1}(\mathbf{p})\mathbf{b}_k\mathbf{b}_k^H\mathbf{Z}^{-1}(\mathbf{p})p_k}{1 - p_k\mathbf{b}_k^H\mathbf{Z}^{-1}(\mathbf{p})\mathbf{b}_k}$$

and hence

$$\frac{\mathrm{SIR}_k^*(\mathbf{p})}{1 + \mathrm{SIR}_k^*(\mathbf{p})} = p_k \mathbf{b}_k^H \mathbf{Z}^{-1}(\mathbf{p})\mathbf{b}_k\,.$$

So, if $\Psi(x) = \log(x/(1+x)), x > 0$, and using this equation the aggregate utility function in (3.4) yields

$$F(\mathbf{s}) = \sum_{k\in\mathcal{K}} w_k \log\Big(e^{s_k}\mathbf{b}_k^H\mathbf{Z}^{-1}(e^{\mathbf{s}})\mathbf{b}_k\Big)$$
$$= \sum_{k\in\mathcal{K}} w_k \log\Big(e^{s_k}\mathbf{b}_k^H\mathrm{adj}(\mathbf{Z}(e^{\mathbf{s}}))\mathbf{b}_k\Big) - \|\mathbf{w}\|_1 \log\det\big(\mathbf{Z}(e^{\mathbf{s}})\big)\,.$$

Choosing $\Psi(x) = \Psi_\alpha(x) := \frac{x^{1-\alpha}}{1-\alpha},\ \alpha \geq 2, \alpha \in \mathbb{N}$ gives

$$F(\mathbf{s}) = \frac{1}{1-\alpha}\sum_{k\in\mathcal{K}} w_k\Big(\frac{1-e^{s_k}\mathbf{b}_k^H\mathbf{Z}^{-1}(e^{\mathbf{s}})\mathbf{b}_k}{e^{s_k}\mathbf{b}_k^H\mathbf{Z}^{-1}(e^{\mathbf{s}})\mathbf{b}_k}\Big)^{\alpha-1}$$

where $\alpha \geq 2$ and the constant $1/(1-\alpha)$ can be neglected as it has no impact on the maximizer.

Gradient projection algorithm

All partial derivatives of $\mathrm{SIR}_k^*(e^{\mathbf{s}})$ with $\mathrm{SIR}_k^*(\mathbf{p})$ given by (1.12) exist and are continuous functions on $\mathbb{R}^K$ because the inverse matrix $\mathbf{Z}_k^{-1}(e^{\mathbf{s}})$ exists for all $\mathbf{s} \in \mathbb{R}^K$, regardless of the choice of the transmit vectors, and the entries in $\mathbf{Z}_k^{-1}(e^{\mathbf{s}})$ vary continuously with the entries in $\mathbf{Z}_k(e^{\mathbf{s}})$. Hence, we can consider a gradient projection algorithm with a constant step size $\delta > 0$ (sufficiently small):

$$\mathbf{s}(n+1) = \Pi_s\Big[\mathbf{s}(n) + \delta\nabla F(\mathbf{s}(n))\Big], \mathbf{s}(0) \in \mathbb{R}^K \tag{3.9}$$

where $\Pi_s(\mathbf{x})$ is the projection of $\mathbf{x} \in \mathbb{R}^K$ on the closed convex set S [Ber03, BT89] and the kth partial derivative $\nabla_k F(\mathbf{s}) = \frac{\partial F}{\partial s_k}(\mathbf{s})$ yields

$$\nabla_k F(\mathbf{s}) = w_k\Psi'\Big(\mathrm{SIR}_k^*(e^{\mathbf{s}})\Big)\mathrm{SIR}_k^*(e^{\mathbf{s}}) - e^{s_k}\sum_{l\neq k} w_l e^{s_l}\Psi'\Big(\mathrm{SIR}_l^*(e^{\mathbf{s}})\Big)\big|\mathbf{b}_l^H\mathbf{Z}_l^{-1}(e^{\mathbf{s}})\mathbf{b}_k\big|^2 \tag{3.10}$$

where the following identity was used: For an invertible and differentiable matrix function $\mathbf{Y}(x)$, there holds

$$\frac{d\mathbf{Y}^{-1}(x)}{dx} = -\mathbf{Y}^{-1}(x)\frac{d\mathbf{Y}(x)}{dx}\mathbf{Y}^{-1}(x)\,.$$

Hence, due to the individual power constraints on each user $\forall_k\, s_k \leq \log P_k$, the algorithm (3.9) takes the form

$$\begin{aligned} s_k(n+1) = \min\Big\{&\log P_k, s_k(n) + \delta\Big[w_k\Psi'\Big(\mathrm{SIR}_k^*(e^{\mathbf{s}})\Big)\mathrm{SIR}_k^*(e^{\mathbf{s}}) - \\ &e^{s_k}\sum_{l\neq k} w_l e^{s_l}\Psi'\Big(\mathrm{SIR}_l^*(e^{\mathbf{s}})\Big)|\mathbf{b}_l^H\mathbf{Z}_l^{-1}(e^{\mathbf{s}})\mathbf{b}_k|^2\Big]\Big\} \end{aligned} \tag{3.11}$$

where $\mathrm{SIR}_k^*(e^{\mathbf{s}})$ is defined by (1.12).

Lemma 7 (Convergence) *For a sufficiently small step size $\delta > 0$, the sequence $\{\mathbf{p}(n)\}$ generated by the algorithm (3.11) with $\mathbf{p}(n) = e^{\mathbf{s}(n)}$ converges to a stationary point.*

Proof: By standard results [BT89, Ber03], the gradient projection algorithm converges to a stationary point for sufficiently small values of $\delta > 0$ if $F(\mathbf{s})$ is bounded above, continuously differentiable on S and the gradient $\nabla F(\mathbf{s})$ is Lipschitz continuous on any bounded subset of S. The first condition is clearly satisfied due to the power constraints. The second condition holds as well since, by assumption, the utility function $\Psi(x)$ is twice continuously differentiable. Hence, the Hessian of $F(\mathbf{s})$ is bounded in the matrix 2-norm on any bounded subset of S. This implies that $\nabla F(\mathbf{s})$ is Lipschitz continuous on any bounded subset of S [OR00, page 70]. ■

Note that the maximum feasible step size in the algorithm may depend on the choice of the starting point $\mathbf{s}(0)$.

3.1.4 Distributed Alternating Implementation

The computation of the gradient in (3.11) seems be too expensive to be implemented in a distributed environment. The question to be answered is: How should $|\mathbf{b}_l^H \mathbf{Z}_l^{-1}(e^{\mathbf{s}})\mathbf{b}_k|^2$ be measured at the l-th link without costly data exchange or signaling between the nodes? Due to the definition of $\mathbf{Z}_l^{-1}$ this part of the gradient depends on all effective transmit vectors associated with receiver k and hence seems not to be amenable to distributed implementation. In order to circumvent this problem and the demand on perfect synchronization, in this section, we slightly modify the algorithm so that it can be implemented in a distributed manner. The basic idea is to increase the value of the function $G(\mathbf{p}, \mathbf{C})$ in the following alternating fashion: For some given receive beamforming matrix $\mathbf{C}(n)$ and power vector $\mathbf{p}(n)$, a new power vector $\mathbf{p}(n+1)$ is chosen such that $G(\mathbf{p}(n), \mathbf{C}(n)) \leq G(\mathbf{p}(n+1), \mathbf{C}(n))$. Then, the beamforming matrix is updated by $\mathbf{C}(n+1)$ such that $G(\mathbf{p}(n+1), \mathbf{C}(n)) \leq G(\mathbf{p}(n+1), \mathbf{C}(n+1))$. This alternating process is repeated until convergence.

Let us first consider the power vector update. To this end, let $\mathbf{C}$ be fixed and define $G_{\mathbf{C}}(\mathbf{s}) := G(e^{\mathbf{s}}, \mathbf{C})$. Then, the power vector can be updated according to the following algorithm

$$s_k(t+1) = \min\Big[\log P_k, s_k(t) + \delta_p(t)\nabla_k \hat{G}_{\mathbf{C}(t)}(\mathbf{s}(t))\Big],\ t \in \mathbb{N}_0 \tag{3.12}$$

for some $\mathbf{s}(0) \in \mathbb{R}^K$, where, with some abuse of notation, $\nabla \hat{G}_{\mathbf{C}(t)}(\mathbf{s}(t))$ is used to denote a noisy estimation of the gradient vector $\nabla G_{\mathbf{C}(t)}(\mathbf{s}(t))$ and $\{\delta_p(t)\}$ with $\delta_p(t) > 0$ is an appropriately chosen sequence of diminishing step sizes [KY03]. If $\{s_k(t)\}_{t=1}^{L_p}$ is a sequence generated by (3.12) for a certain number of power iterates $L_p \geq 1$, then we put $\mathbf{s}(n+1) = (s_1(L_p), \ldots, s_K(L_p))$. Note, that the estimate $\nabla_k \hat{G}_{\mathbf{C}(t)}(\mathbf{s}(t)), k \in \mathcal{K}$ can be computed in a distributed manner using the adjoint network of [SWB07] shortly summarized in section 2.6 Appendix A. This scheme enables each transmitter to estimate its current update direction from the received signal power. This mitigates the problem of global coordination of the transmitters when carrying out gradient-projection algorithms in distributed wireless networks. More precisely, instead of each node sending its message separately as in case of classical flooding protocols, nodes transmit simultaneously (only coarse synchronization is required) over the adjoint network such that each node can estimate its gradient component from the received power. The price for this are possible estimation errors that usually can be dealt with using a diminishing step size [KY03] as is shortly discussed in the next subsection 'Stochastic Approximation View'.

Now assume that $\mathbf{s} = \log \mathbf{p}$ is fixed. Distributed algorithms for computing optimal receive beamformers defined by (3.3) are widely established. These algorithms are based either on blind or pilot-based estimation methods [Ver98]. In the latter case, if X_k is a pilot symbol[2] of user k with $E[|X_k|^2] = e^{s_k}$ and $\mathbf{r}_k \in \mathbb{C}^M$ represents the observations at receiver k, then $\mathbf{c}_k^*$ given by (1.11) minimizes the mean square error $\theta_k(\mathbf{c}_k) = E[|\kappa X_k - \langle \mathbf{c}_k, \mathbf{r}_k \rangle|^2]$ over $\mathbb{C}^M$, where $\kappa > 0$ is a normalizing constant[3] chosen such that, in the minimum, $\|\mathbf{c}_k\|_2 = 1$. Note that the expectation is taken with respect to $\mathbf{r}^{(k)} := (\mathbf{r}_k, X_k)$, which depends on the logarithmic power vector $\mathbf{s} \in \mathbb{R}^K$. Now if the convex function $\theta_k(\mathbf{c}_k)$ was explicitely known, then the algorithm (with the complex gradient operator ∇ which gives the direction of steepest ascent of $\theta_k : \mathbb{C}^M \to \mathbb{R}_+$)

$$\mathbf{c}_k(t+1) = \mathbf{c}_k(t) - \delta_r \nabla \theta_k(\mathbf{c}_k(t)), k \in \mathcal{K}, t \in \mathbb{N}_0 \tag{3.13}$$

would converge to $\mathbf{c}_k^*$ defined by (1.11) for a sufficiently small step size $\delta_r > 0$. The problem is that the function θ_k is usually not known since the distribution of $\mathbf{r}_k$ is not known [Ver98]. Therefore, $\nabla \theta_k(\mathbf{c}_k(t))$ cannot be computed and the algorithm must be modified using the framework of stochastic approximation [KY03]. The idea is to consider the functions $\theta_{\mathbf{r}^{(k)}}(\mathbf{c}_k) = |\kappa X_k - \langle \mathbf{c}_k, \mathbf{r}_k \rangle|^2$ for all $\mathbf{r}^{(k)}$ as noisy estimations of $\theta_k(\mathbf{c}_k)$. Then, under some conditions on the estimation error and for any $\mathbf{c}_k(0) \in \mathbb{C}^M$, the algorithm

$$\mathbf{c}_k(t+1) = \mathbf{c}_k(t) - \delta_r(t) \nabla \theta_{\mathbf{r}^{(k)}(t)}\big(\mathbf{c}_k(t)\big), t \in \mathbb{N}_0 \tag{3.14}$$

converges to $\mathbf{c}_k^*$ (in some probabilistic sense), provided that the step size $\delta_r(t) > 0$ with $\lim_{t\to\infty} \delta_r(t) = 0$ and $\sum_{t=0}^{\infty} \delta_r(t) = +\infty$ is chosen suitably [HMV95].

Now combining these two ingredients leads to the following joint power control and receive beamforming algorithm: At the beginning of every frame, $\mathbf{s}(0)$ and $\mathbf{C}(0)$ are set to be equal to the current transmit powers and receive beamformers. Then, all users

[2] A zero-mean random variable that is known to the estimator

[3] For practical implementation, we can assume $\kappa = 1$, and then normalize the beamformers so that their l^2-norms are equal to one.

concurrently execute $N \geq 1$ updates of their transmit powers and receive beamformers. The nth update consists of the following intermediate steps:

(i) For fixed $\mathbf{C}(n)$ and a number of power iterates $L_p \geq 1$, each user $k \in \mathcal{K}$ generates a sequence $\{s_k(t)\}_{t=1}^{L_p}$ by carrying out (3.12) and defines $s_k(n+1) = s_k(L_p)$.

(ii) For a number of receiver iterates $L_r \geq 1$ and with $E[|X_k|^2] = e^{s_k(n+1)}$, each user $k \in \mathcal{K}$ executes L_r iterations of the algorithm (3.14) to obtain the sequence $\{\mathbf{c}_k(t)\}_{t=1}^{L_r}$. It defines $\mathbf{c}_k(n+1) = \mathbf{c}_k(L_r)$.

The convergence of the algorithm (in some probabilistic sense) strongly depends on the choice of the step sizes in (3.12) and (3.14) as well as on the properties of the estimation errors in (3.12) and (3.14). However, we point out that the algorithm is motivated by the following observation: If the estimates in (3.12) are known perfectly[4] and (3.13) is used instead of (3.14), then the sequence $\{(\mathbf{s}(n), \mathbf{C}(n))\}$ generated by the resulting algorithm converges to a stationary point. This is because, under this assumption, (3.12) and (3.13) are both monotonic, and hence we have (for all $n \in \mathbb{N}_0$)

$$G(\mathbf{p}(n), \mathbf{C}(n)) \leq G\big(\mathbf{p}(n+1), \mathbf{C}(n)\big) \leq G\big(\mathbf{p}(n+1), \mathbf{C}(n+1)\big).$$

This implies that the sequence $\{G\big(\mathbf{p}(n), \mathbf{C}(n)\big)\}$ is monotonically increasing, provided that the step sizes are sufficiently small. Moreover, it is bounded since $G\big(\mathbf{p}(n), \mathbf{C}(n)\big) \leq G\big(\mathbf{p}^*, \mathbf{C}^*\big)$ for all $n \in \mathbb{N}_0$. Therefore, the algorithm converges to a stationary point. In addition, to show that this stationary point is a local maximizer for the problem (3.1), we need to verify the second order sufficiency condition.

Due to scarce resources in wireless networks, it is reasonable to choose the number of updates $N = 1$ in every frame. In addition, instead of transmitting pilot signals in the intermediate step (ii), the optimal receive beamformers can be estimated during the data transmission using some blind estimation method (see [Ver98] and references therein). So, at the beginning of every frame, the step (i) is executed only once. Then, the resulting transmit powers are used for data transmission. During this time, the receive beamformers are updated online after each transmitted symbol. However, numerical experiments suggest that the scheme should not exclusively rely on blind methods to estimate the optimal receivers with a sufficient accuracy.

Stochastic Approximation View

As already mentioned, in real-world networks estimation errors and other distorting factors as quantization noise occur. Now the interesting question is, what is the impact of these stochastic noisy measurements on the convergence properties. Does the proposed

[4] In this case, $\delta_p(t) = \delta_p$ for sufficiently small $\delta_p > 0$.

algorithm still converge and under what conditions? In the case of such distorting factors, the proposed algorithm has to be analyzed applying stochastic approximation theory. However, this topic is too wide to be considered in more detail. We refer to [KY03] as a comprehensive reference. In the following we give only a few known insights and present some exemplary numerical results (section 3.1.5).

We assume that the estimated gradient component $\nabla\hat{G}_{\mathbf{C}(t)}(\mathbf{s}(t))$ is a random variable of the form

$$\nabla_k\hat{G}_{\mathbf{C}(t)}(\mathbf{s}(t)) = \nabla_k G_{\mathbf{C}(t)}(\mathbf{s}(t)) + M_k(t)$$

where $M_k(t), k \in \mathcal{K}$ is the estimation noise process that fulfills the following conditions [KY03]:

(A.12) The estimation noise process depends on the receiver noise process which is assumed to be a martingale difference that is uncorrelated with transmit symbols and has a finite variance.

(A.13) The estimation noise is zero-mean and exogeneous, meaning that $M_k(t), k \in \mathcal{K}$ is independent of the iterate value.

Assuming these two conditions one can deal with the estimation noise applying a diminishing step size sequence that satisfies $\delta_p(t) > 0$ with $\lim_{t\to\infty}\delta_p(t) = 0$ and $\sum_{t=0}^{\infty}\delta_p(t) = +\infty$. A typical choice for a step size sequence is for instance $\delta_p(t) = ct^y$ for some $y \in (0, 1]$. The choice of the step size is central to the effectiveness of the algorithm as is shown by simulations in the next section 3.1.5.

In the above algorithm the powers and beamformers are updated in parallel, meaning that the power control algorithm does not wait for the convergence of the receive beamformers and vice versa. Thus the convergence of this practical stochastic algorithm is only verified by simulations presented in the following section. In addition, note that condition (A.13) is not necessarily fulfilled by the distributed power control algorithm. Thus the estimates $\nabla_k\hat{G}_{\mathbf{C}(t)}(\mathbf{s}(t))$ may be biased by some $b_k(t)$ meaning that $\nabla_k\hat{G}_{\mathbf{C}(t)}(\mathbf{s}(t)) = \nabla_k G_{\mathbf{C}(t)}(\mathbf{s}(t)) + M_k(t) + b_k(t)$. Simulation results indicate that the algorithm still converges to a contraction region around the optimal point provided that the bias is bounded by a scaled version of the true gradients.

3.1.5 Numerical Results

In the following, we show exemplarily the convergence behavior of the proposed scheme for $\Psi(x) = -1/x, x > 0$ and a random channel realization. We consider a wireless system with $M = 2$ transmit and receive antennas, and $K = 4$ users operating at a SNR level of 30 dB. The weight vector is $\mathbf{w} = \mathbf{1}$. The noisy measurements of the gradient are assumed to be $\nabla_k\hat{G}_{\mathbf{C}(t)}(\mathbf{s}(t)) = \nabla_k G_{\mathbf{C}(t)}(\mathbf{s}(t)) + z, k \in \mathcal{K}$, where z is an independent zero-mean Gaussian random variable (and thus fulfils the conditions of a martingale difference noise) whose

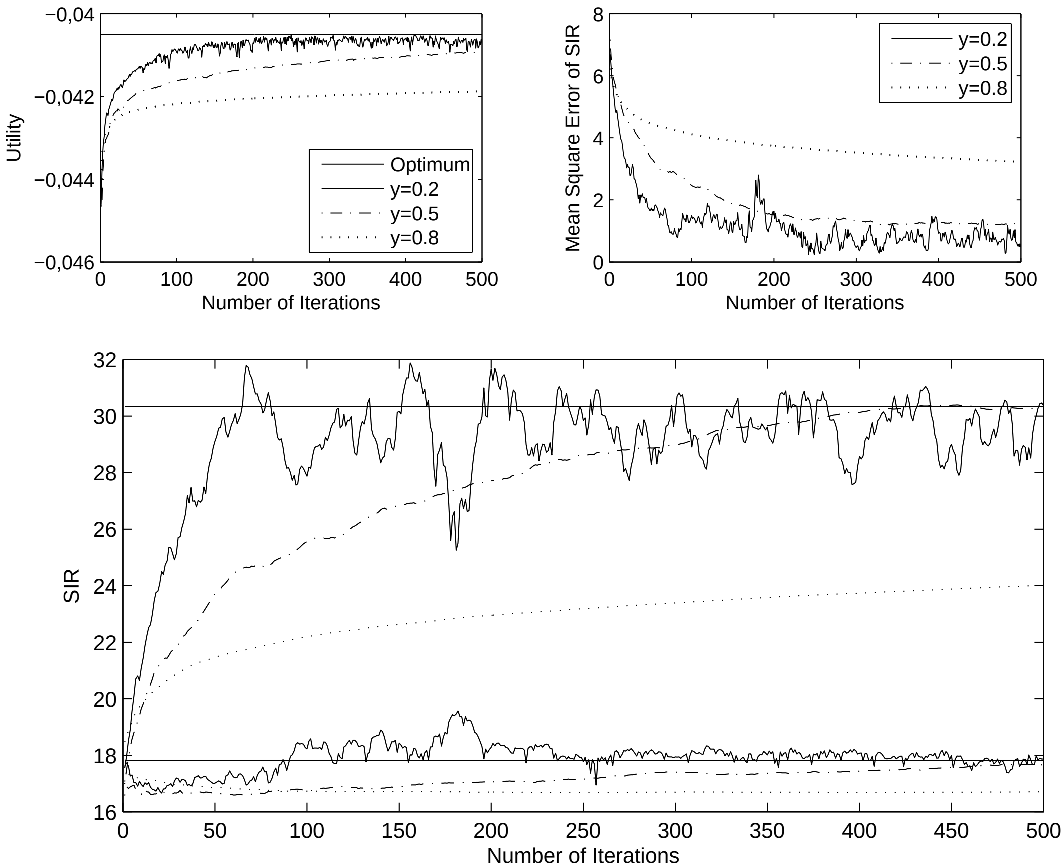

Figure 3.1: Convergence behavior of the distributed algorithm in case of noisy measurements for different values of y, $c_1 = 30$ and $c_2 = 0.5$.

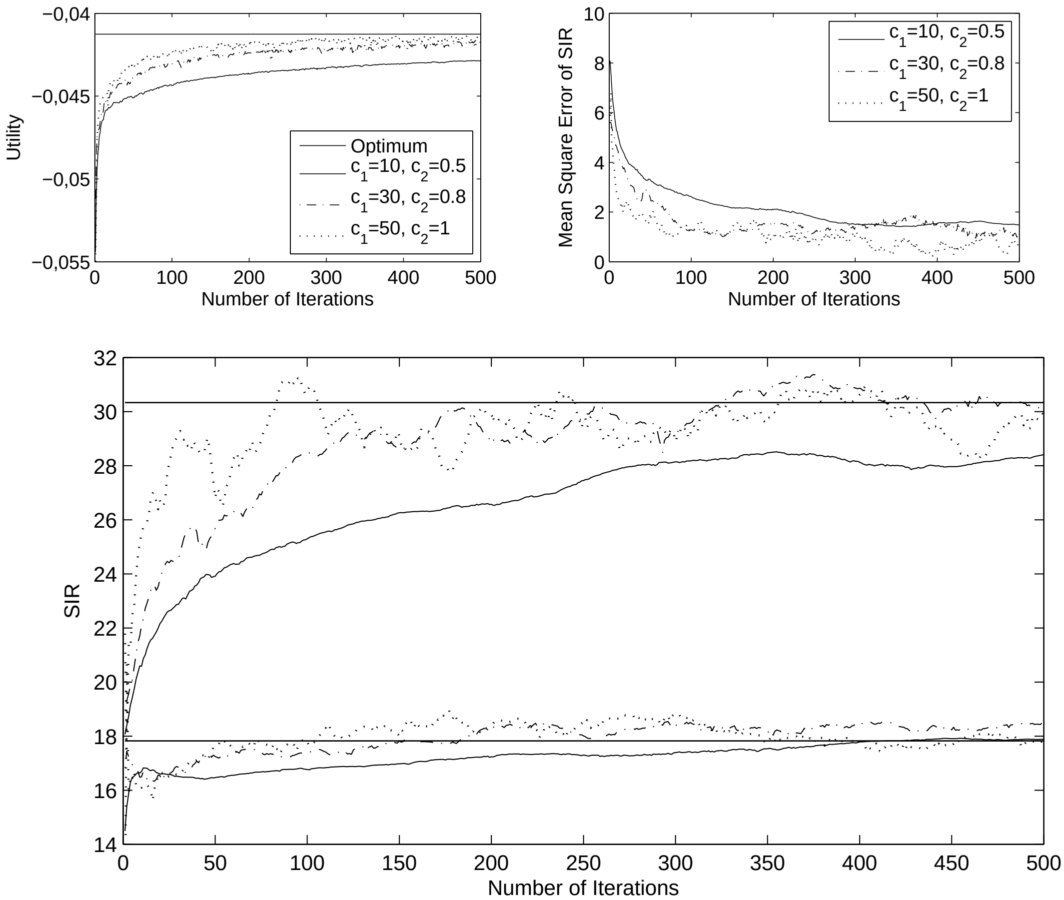

Figure 3.2: Convergence behavior of the distributed algorithm in case of noisy measurements for different values of the step sizes c_1, c_2 and $y = 0.5$.

variance $\sigma_z^2(t)$ depends on t and is 10 percent of the absolute gradient value. We have $L_p = 1$ and $L_r = 16$ steps in (i) and (ii), respectively. Hence, during each iteration step n, the algorithm performs 1 power control step and estimates the beamformers using 16 pilot symbols. The diminishing step sizes for the intermediate steps are $\delta_p(t) = c_1/(t+1)^y$ and $\delta_r(t) = c_2/(t+1)^y$ for some positive constants c_1, c_2 and some exponent $y \in (0,1]$. Figure 3.1 depicts the aggregate utility, the mean square error of the SIR and the SIR for two users over the number of iterations n for different values of y to show the influence of the diminishing step size. As can be easily seen, if the step size vanishes fast the algorithm converges much slower than with a slowly decreasing step size. However, the behavior is very smooth causing nearly no oscillations in contrast to a slowly decreasing step size. Figure 3.2 depicts the aggregate utility, the mean square error of the SIR and the SIR for two users over the number of iterations n for different values of c_1 and c_2 and a fixed $y = 0.5$ to show the influence of the start step size values. Here a higher (but sufficiently small) start step size leads to a faster but oscillating convergence compared to lower start step sizes with a slow but smooth convergence behavior.

Summarizingly, we state the following. It is important that the step sizes c_1 and c_2 are sufficiently small to ensure that the algorithm does not diverge. Besides, the decrease of the step sizes (exponent y) should be not too small to avoid a very slow convergence speed. In case of a dynamic environment where the channel changes over time, y should be chosen to be able to follow the channel changes. This is then obtained at the cost of a more oscillating behavior. Finally note that the length of the pilot sequences also depends on the number of users because the link-specific pilot sequence is typically a pseudo-noise sequence with good autocorrelation poperties.

In figure 3.3 we depict a convergence example for the case that the estimates $\nabla_k \hat{G}_{\mathbf{C}(t)}(\mathbf{s}(t))$ are biased by some $b_k(t)$. Further independent simulations suggest that the proposed algorithm converges to a contraction region around the optimal point if the bias is small enough. Otherwise the algorithm may diverge. However, the conditions on the bias to ensure convergence to a contraction region remain an open question.

3.1.6 Comparison with Pure Beamformer Control

Finally, in this section, we compare our joint utility-based scheme with the scheme of [SBH10] that optimizes transmit and receive beamformers only, with the aim to maximize the sum rate. Therefore, we first derive a heuristic algorithm to additionally update the transmit beamformers or precoders in our utility-based joint power and receive beamforming algorithm. Then we summarize the main ingredients of the scheme in [SBH10]. Finally, we compare both schemes numerically to show the influence of beamforming and power control on the sum rate performance. Note, that due to the fact that the [SBH10] aims at sum rate maximization, we need to choose $\Psi(x) = \log(x)$.

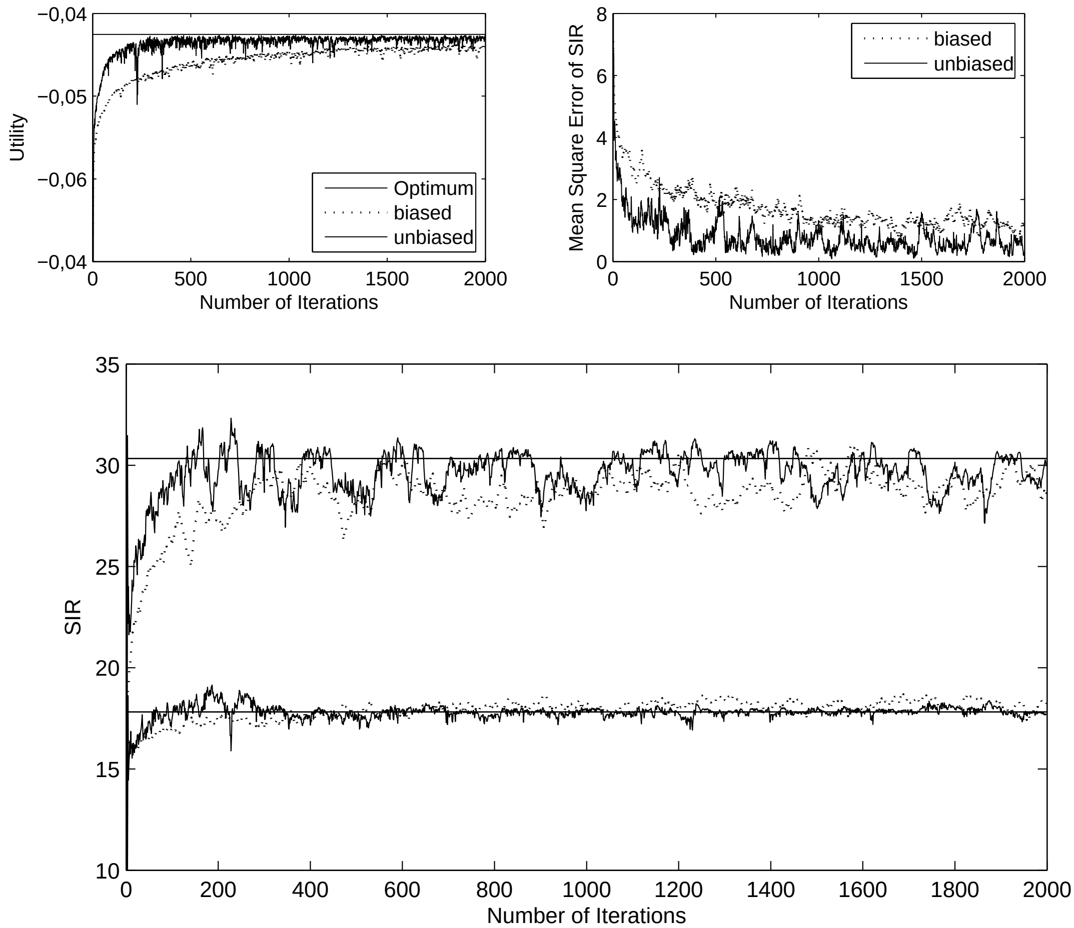

Figure 3.3: Convergence behavior of the distributed algorithm in case of noisy measurements $y = 0.4$, $c_1 = 40$, $c_2 = 1$ for the biased and unbiased case.

Joint Utility-based Power and Transceiver Optimization

For comparison reasons we now extend the joint utility-based power control and receive beamforming algorithm and heuristically incorporate the transmit filters [3]. We point out, that the effective transmit vector is affected by the transmitter structure, the channel state etc. These vectors determine how the transmit signals direct into and through the space. Assuming the transmitter structure to be a transmit beamformer any user can influence its effective transmit vectors $\mathbf{b}_l^{(k)}$ by changing its transmit beamformer $\mathbf{u}_l$ defined in section 1.4.5. The optimization of transmit beamforming is crucial, because in contrast to receive beamforming it impacts the interference power to all other receivers and thus the performance of the other users. However, depending on the interference scenario a transmit beamformer optimization may achieve high performance gains. Related work to both transmitter and joint transmitter and receiver optimization can be found in [RLT02, CTR98, SB06]. However, it has been shown that the virtual uplink concept to calculate transmit beamformers and powers cannot be applied in networks with general power constraints [RLT02]. Next we present a heuristic scheme that recalculates the transmit beamformers as long as the utility improves.

As in [CTR98] the basic idea is to optimize the transmit and receive beamformers in an alternating way. For a better understanding we use the notion of the primal and the reversed network. In contrast to the primal network the reversed network is the network that is obtained by reversing or swapping the roles of transmitters and receivers. More precisely, transmitters and transmit beamformers are assumed to be receivers and receive beamformers, respectively. Vice versa receivers and receive beamformers are assumed to be transmitters and transmit beamformers. Each iteration of the joint optimization of powers, receive and transmit beamformers consists of 4 steps, where 2 steps are executed in the primal and in the reversed network, respectively. These 2 steps include the update of the power vector and the beamformer optimization. The proposed algorithm is summarized in Alg. 3. Next we describe the single steps in detail. Note that the transmit beamforming vectors are collected in the transmit beamforming matrix $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_K) \in \mathbb{C}^{M \times K}$, $\mathbf{U} \in \mathrm{U}$ where $\mathrm{U} = \{(\mathbf{u}_1, \ldots, \mathbf{u}_K) : \|\mathbf{u}_k\|_2 = 1\}$ denotes the set of all transmit beamforming matrices. The algorithm performs the following steps. First each user calculates for fixed transmitters and receivers the power allocation $\mathbf{p}(t)$. The power computation was explained in section 3.1.4. After that each user updates its receive beamformers $\mathbf{c}_k(t)$ in the primal network such that the SIR is maximized. Therefore each user fixes its transmit power $\mathbf{p}(t-1)$ and transmit beamforming vector $\mathbf{u}_k(t-1), k \in \mathcal{K}$. Then we switch to the reversed network. Remember, in the reversed network the transmit beamformers $\mathbf{u}_k, k \in \mathcal{K}$ are assumed to be receive beamformers and the receive beamformers $\mathbf{c}_k, k \in \mathcal{K}$ are assumed to be transmit beamformers. The transmit powers of the reversed network are collected in the power vector $\mathbf{q} = (q_1, \ldots, q_K) \in \mathbb{R}_+^K$ that is subject to some individual power constraints $q_k \leq Q_k, k \in \mathcal{K}$. Now, in the reversed network first we perform the power update for fixed transmit and receive beamformers. The power allocation $\mathbf{q}(t)$ for the reversed network is computed again such that the sum of weighted utilities of SIRs is maximized. Then the receive beamformers of the reversed network $\mathbf{u}_k(t), k \in \mathcal{K}$ are updated such that the SIR of each user is maximized. The algorithm terminates if the sum utility values decrease or a sufficiently small relative improvement is achieved. In addition the algorithm may also be stopped after a predefined number of iterations.

Algorithm 3 Heuristic Joint Power and Transceiver Optimization

Require: $\delta > 0, t = 0, \mathbf{p}(0) \in \mathrm{P}_+, \mathbf{C}(0) \in \mathrm{C}, \mathbf{U}(0) \in \mathrm{U}$

repeat
    $t = t + 1$
    primal network: compute power allocation $\mathbf{p}(t)$
    primal network: given $\mathbf{p}(t)$ and $\mathbf{U}(t-1)$ compute receive beamformers $\mathbf{c}_k(t), k \in \mathcal{K}$
    reversed network: compute power allocation $\mathbf{q}(t)$
    reversed network: given $\mathbf{q}(t)$ and $\mathbf{C}(t)$ compute receive beamformers $\mathbf{U}(t)$
until some termination condition is satisfied

Adaptive Beamforming via Bi-directional Training

Recently the authors of [SBH10] designed an adaptive distributed algorithm for upda-ting the precoders and receive beamformers to maximize the sum rate. It consists of two steps that are repeated as long as the sum rate increases: (i) Fix the precoders and optimize the receivers in the primal network, (ii) reverse the direction of the transmission and optimize the transmit filters (now receivers) in the adjoint network. The optimization criterion in each step is to maximize the SINR of each user. The authors assume that in each step the set of transmitters or the set of receivers transmit coarse synchronously pilot or training sequences in each direction. Consequently this scheme is quite similar to our distributed receive beamforming scheme and to the heuristic transceiver optimization proposed above and in [3]. We emphasize that in addition the authors numerically investigated the influence of the training length and the number of bi-directional iterations.

Numerical Results

Finally we compare both schemes. We consider a network with K users, $M = 4$ transmit and receive antennas and a channel matrix which entries are iid complex Gaussian distributed. Each user is subject to the same individual power constraint and operates at SNR= 30dB.

Figures 3.4 and 3.5 illustrate for an exemplary network the gain that can be achieved by different resource allocation schemes. More precisely, the sum rate averaged over 1000 channel realizations is depicted over the number of users in the network for the following 5 resource allocation strategies. Note, with increasing user number the interference in the

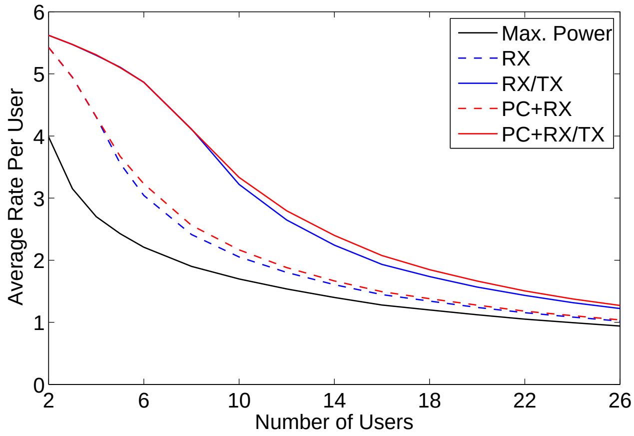

Figure 3.4: Average rate per user over the number of users for different resource allocation schemes

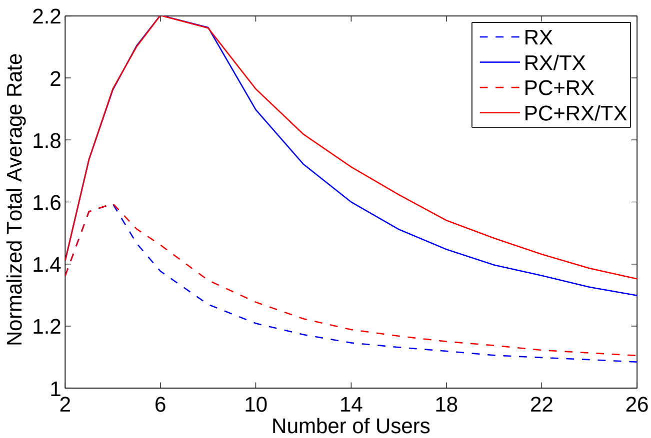

Figure 3.5: Relative improvement of the average rate per user with respect to the maximum power scheme (Max. power) for different resource allocation schemes

network increases. If not otherwise stated the beamformers of the schemes adapt to the channel. The first strategy, that is also used as reference strategy in figure 3.4, allocates the maximum power to each user (max. power). The same is true for the scheme of [SBH10]. However, here the receivers only (denoted by RX) or both the transmitters and receivers (denoted by RX/TX) are optimized. Finally we show for $\Psi(x) = \log(x)$ the sum rate that can be achieved by a joint power and receiver control (denoted by PC+RX) and by a joint power and transceiver control (denoted by PC+RX/TX). As can be easily seen, optimizing the receive and/or transmit filters may significantly improve the network performance. In contrast and as is known, the mechanism power control provides only small gains in case of a sum rate maximization. Consequently, in case of sum rate maximization ($\alpha = 1$) a pure and low complex beamformer control is preferrable in distributed wireless networks. In the following chapter which investigates the max-min approach, numerical results will show that power control plays an important role.

3.2 Max-Min Fair Joint Power and Receiver Control

This section discusses the possibility of solving the max-min SIR-balancing problem by optimizing a certain aggregate utility function over the joint space of powers and receivers. As emphasized before, in contrast to the max-min SIR problem, an advantage of a utility-based approach is certainly the existence of distributed power control algorithms for a given set of weights.

Recent results [SKB10] obtain a simple characterization of the max-min SIR-balanced power vector under general power constraints. The authors formulate an eigenvalue problem of the same dimension as the original problem and give a procedure to compute the max-min SIR power vector. The authors also reveal the connection between the max-min SIR-balancing and the utility-based power control problem. We briefly explain this connection in section 3.2.2, where we review how to choose the weight vector such that the sum utility optimization achieves the max-min SIR-balancing solution. In addition we discuss the properties of this connection in order to derive an algorithmic solution for our joint power control and receive beamforming problem. Based on these insights, in section 3.2.4, we propose an iterative algorithm that solves the joint power and beamformer optimization problem. We prove important properties as monotonicity and convergence to the optimal max-min SIR-balancing solution. Thus, in this section we generalize the results of [BS06b] to noisy channels under general power constraints and to a larger class of utility functions. More precisely, the authors of [BS06b] describe a similar alternating scheme to solve the joint max-min fair power control and preequalizer problem in the noiseless case for a multiuser downlink channel. They achieve the max-min SIR-balancing solution by minimizing the sum of weighted inverse SIRs. Thus, due to their problem structure the single steps of the algorithm are different. On the other hand, we only provide optimal powers and receive beamformers. For transmit beamformer optimization we propose a heuristic algorithm in section 3.2.5 because the theory of uplink-downlink duality is not applicable in case of general power constraints. Numerical results illustrate the gain achievable by a sophisticated resource allocation. Finally, in this section we discuss the amenability of our algorithm to distributed implementation. Unfortunately, so far a fully-distributed algorithm is not available. On the positive side, we refer to recent results that mitigate this problem considerably.

The results presented in this section appeared in [3].

3.2.1 System Model, Assumptions and Definitions

We consider a multiantenna wireless network with general power constraints as defined in (1.4) so that $\mathbf{p} \in \mathrm{P}$ where $\mathrm{P} = \{\mathbf{p} \in \mathbb{R}_+^K : \mathbf{A}\mathbf{p} \le \mathbf{p}', \mathbf{A} \in \{0,1\}^{N \times K}\} \subset \mathbb{R}_+^K$. Remember, as defined in section 1.4.5 the main figure of merit is the SIR given by

$$\mathrm{SIR}_k(\mathbf{p}, \mathbf{c}_k) := \frac{p_k}{I_k(\mathbf{p}, \mathbf{c}_k)} = \frac{p_k}{\sum_{l=1}^K v_{k,l}(\mathbf{c}_k) p_l + z_k(\mathbf{c}_k)}.$$

Throughout this section we also assume that

(A.14) each entry of the gain matrix $\mathbf{V}(\mathbf{C})$ is a continuous function of $\mathbf{C} \in \mathrm{C}$. Moreover, for any fixed $\mathbf{C} \in \mathrm{C}$, $\mathbf{V} = \mathbf{V}(\mathbf{C}) \ge 0$ is an irreducible matrix.

Remark 1 *In practice, the irreducibility property can be guaranteed by approximating the gain matrices* $\mathbf{V}(\mathbf{C}), \mathbf{C} \in \mathrm{C}$, *with matrices of the form* $\mathbf{V}(\mathbf{C}) + \epsilon\mathbf{A}$ *for some sufficiently small* $\epsilon > 0$ *where* $\mathbf{A} = (\mathbf{1}\mathbf{1}^T - \mathbf{I}) \in \{0,1\}^{K \times K}$.

Let $\gamma_1, \dots, \gamma_K > 0$ be the SIR targets of the users, and let $\mathbf{\Gamma} := \mathrm{diag}(\gamma_1, \dots, \gamma_K)$.

Definition 4 *Given any* $\mathbf{\Gamma}$, *we say that* $(\mathbf{p}^*, \mathbf{C}^*)$ *is a max-min SIR-balanced solution if*

$$(\mathbf{p}^*, \mathbf{C}^*) := \arg\max_{(\mathbf{p},\mathbf{C}) \in \mathrm{P} \times \mathrm{C}} \min_{k \in \mathcal{K}} \left(\mathrm{SIR}_k(\mathbf{p}, \mathbf{c}_k)/\gamma_k\right). \tag{3.15}$$

The maximum exists as the SIRs are continuous on the compact set $\mathrm{P} \times \mathrm{C}$. Moreover, it may be easily verified that $\mathbf{p}^* > 0$, allowing us to focus on $\mathrm{P}_+ = \mathrm{P} \cap \mathbb{R}_{++}^K$.

Notice that γ_k is not necessarily met under a max-min SIR-balanced solution (3.15). However, it is well known [SWB09] that $\mathrm{SIR}_k(\mathbf{p}, \mathbf{c}_k) \ge \gamma_k$ for each k and some $(\mathbf{p}, \mathbf{C}) \in \mathrm{P} \times \mathrm{C}$ if and only if $\mathrm{SIR}_k(\mathbf{p}^*, \mathbf{c}_k^*) \ge \gamma_k$ for each k. In other words, $\mathbf{\Gamma}$ is feasible if and only if each γ_k is met under a max-min SIR-balanced solution.

3.2.2 Characterization of Max-Min SIR-Balancing for Fixed Beamformers

Throughout this section we assume that in addition to the transmit filters, the receive beamformers $\mathbf{C} \in \mathrm{C}$ are arbitrary but fixed. Note that by (A.14), $\mathbf{V} := \mathbf{V}(\mathbf{C})$ is a nonnegative irreducible matrix. In what follows, we briefly review the connection between the max-min SIR-balancing power control problem (3.15) and a utility-based power control problem [SKB10]. More precisely, we show how to choose the weight vector in the utility maximization problem so that a solution to this problem is a max-min SIR-balancing solution. For the proofs, the reader is referred to [SKB10, SWB09]. Let us first define the max-min SIR-balancing power control problem. For any given $\gamma_k > 0, k \in \mathcal{K}$, a max-min SIR-balanced power vector is defined to be

$$\begin{aligned} \bar{\mathbf{p}} := \bar{\mathbf{p}}(\mathbf{C}) = \arg\max_{\mathbf{p}\geq 0} \min_{k\in\mathcal{K}} \big(\mathrm{SIR}_k(\mathbf{p},\mathbf{c}_k)/\gamma_k\big) \\ \text{s.t.} \quad \max_{n\in\mathcal{N}} g_n(\mathbf{p}) \leq 1\,, \end{aligned} \tag{3.16}$$

where

$$g_n(\mathbf{p}) := (1/P_n)\mathbf{a}_n^T\mathbf{p}, \quad n \in \mathcal{N} \tag{3.17}$$

and $\mathbf{a}_n \in \{0,1\}^K$ is a (column) vector equal to the nth row of the power constraint matrix $\mathbf{A}$. Note that $\mathbf{p} \geq 0$ and $\max_{n\in\mathcal{N}} g_n(\mathbf{p}) \leq 1$ holds if and only if the power constraints are satisfied: $\mathbf{p} \in \mathrm{P}$. Moreover, for any power vector $\bar{\mathbf{p}}$ that solves (3.16), we have $\max_{n\in\mathcal{N}} g_n(\bar{\mathbf{p}}) = 1$, which means that at least one power constraint is active.

The irreducibility assumption of $\mathbf{V}(\mathbf{C})$ implies that the max-min SIR-balanced power vector $\bar{\mathbf{p}}$ is unique and that

$$\forall_{k\in\mathcal{K}}\ \mathrm{SIR}_k(\bar{\mathbf{p}})/\gamma_k = \beta \tag{3.18}$$

for some $\beta > 0$. From these observations, it follows that [SKB10]

$$\frac{1}{\beta}\bar{\mathbf{p}} = \mathbf{\Gamma}\mathbf{V}\bar{\mathbf{p}} + \mathbf{\Gamma}\mathbf{z} \qquad\qquad \max_{n\in\mathcal{N}} g_n(\bar{\mathbf{p}}) = 1\,. \tag{3.19}$$

Now let

$$\mathcal{N}(\mathbf{p}) := \big\{m \in \mathcal{N} : m = \arg\max_{n\in\mathcal{N}} g_n(\mathbf{p}) = 1\big\}$$

be the set of power constraints that are active under the power vector $\mathbf{p}$. Note that due to the fact that at least one power constraint is active in the optimum, $\mathcal{N}(\bar{\mathbf{p}})$ is not empty. An important observation is that (see [SKB10])

$$\bar{\mathcal{N}} := \mathcal{N}(\bar{\mathbf{p}}) = \big\{m \in \mathcal{N} : m = \arg\max_{n\in\mathcal{N}}\ \rho(\mathbf{B}^{(n)})\big\} \tag{3.20}$$

where $\rho(\mathbf{B}^{(n)})$ is used to denote the spectral radius of $\mathbf{B}^{(n)}$.[5] Using this definition, we can

[5]Since $\mathbf{B}^{(n)}, n \in \mathcal{N}$, is irreducible, it follows from Perron-Frobenius theorem [Sen81] that $\rho(\mathbf{B}^{(n)})$ is a simple eigenvalue of $\mathbf{B}^{(n)}$. Moreover, left and right eigenvectors of $\mathbf{B}^{(n)}$ associated with $\rho(\mathbf{B}^{(n)})$ are positive and unique up to positive multiples. These eigenvectors are called (left and right) principle eigenvectors of $\mathbf{B}^{(n)}$.

rewrite (3.19) as

$$\frac{1}{\beta}\bar{\mathbf{p}} = \mathbf{\Gamma}\mathbf{V}\bar{\mathbf{p}} + \mathbf{\Gamma}\mathbf{z} \cdot g_n(\bar{\mathbf{p}}),\ n \in \bar{\mathcal{N}} \tag{3.21}$$

which is simply because $g_n(\bar{\mathbf{p}}) = 1$ for each $n \in \bar{\mathcal{N}}$.

Thus, with (3.17), we have (for each $n \in \bar{\mathcal{N}}$)

$$\frac{1}{\beta}\bar{\mathbf{p}} = \mathbf{B}^{(n)}\bar{\mathbf{p}}, \quad \beta > 0, \bar{\mathbf{p}} \in \mathbb{R}^K_{++}, g_n(\bar{\mathbf{p}}) = 1 \tag{3.22}$$

where $\mathbf{B}^{(n)} \in \mathbb{R}^{K\times K}_{+}$ is defined to be (for each $n \in \mathcal{N}$)

$$\mathbf{B}^{(n)} := \mathbf{\Gamma}\mathbf{V} + \frac{1}{P_n}\mathbf{\Gamma}\mathbf{z}\mathbf{a}_n^T = \mathbf{\Gamma}\big(\mathbf{V} + 1/P_n\mathbf{z}\mathbf{a}_n^T\big)\,. \tag{3.23}$$

So, given $\bar{\mathbf{p}}$ defined by (3.16), the condition (3.22) must hold for each $n \in \bar{\mathcal{N}}$. In other words, if $\mathbf{V}$ is irreducible, a max-min SIR-balanced power vector (3.16) must satisfy (3.22) for each node $n \in \mathcal{N}$ whose power constraint is active at the maximum. Since $\mathbf{B}^{(n)}, n \in \mathcal{N}$, is irreducible, the converse can be concluded from the Perron-Frobenius theorem [Sen81, Mey00].

Proposition 4 ([SKB10, SWB09]) *Let (A.14) be satisfied, and let $\beta > 0$ be given by (3.18). Then, $\mathbf{p} = \bar{\mathbf{p}} > 0$ if and only if $\mathbf{p}$ is the right principle eigenvector of $\mathbf{B}^{(n)}, n \in \bar{\mathcal{N}}$, normalized such that $g_n(\mathbf{p}) = 1$. Moreover, $1/\beta = \rho(\mathbf{B}^{(n)}), n \in \bar{\mathcal{N}}$.*

We point out, that these results lead to a simple procedure for computing the max-min SIR-balanced power vector. Besides notice, in matrix $\mathbf{B}$ the power constraints provide an additional interference term to $\mathbf{V}$.

3.2.3 Connection to Utility-based Power Control

By the above results and definitions we are in the position to explain the aforementioned connection between the max-min SIR-balancing problem and a utility maximization problem for any fixed beamformers. In words, we show how to compute the max-min SIR-balanced power vector by solving the corresponding utility-based power control problem. To this end, given any $\mathbf{C} \in \mathrm{C}$, consider the following aggregate utility function

$$U(\mathbf{p}, \mathbf{w}) := \sum_{k\in\mathcal{K}} w_k \Psi\Big(\frac{\mathrm{SIR}_k(\mathbf{p}, \mathbf{c}_k)}{\gamma_k}\Big) \tag{3.24}$$

where Ψ fulfills (A.1)–(A.3). Moreover, without loss of generality, it is assumed that

(A.15) $\mathbf{w} \in \Pi^+_K := \{\mathbf{w} \in \mathbb{R}^K_{++} : \|\mathbf{w}\|_1 = 1\}$.

The utility maximization problem of interest is to maximize the aggregate utility function (3.24) over P_+:

$$\mathbf{p}(\mathbf{w}) := \arg\max_{\mathbf{p}\in\mathrm{P}_+} U(\mathbf{p},\mathbf{w}) \tag{3.25}$$

where the maximum can be shown to exist. Furthermore, it is obvious that in the maximum at least one power constraint is active:

$$\forall_{\mathbf{w}\in\Pi_K^+} \exists_{n\in\mathcal{N}}\, g_n(\mathbf{p}(\mathbf{w})) = 1\,. \tag{3.26}$$

Now, the following connection between (3.16) and (3.25) can be observed.

Proposition 5 ([SKB10, SWB09]) *Suppose that* $\mathbf{y}^{(n)}$ *and* $\mathbf{x}^{(n)}$ *are left and right principle eigenvectors of* $\mathbf{B}^{(n)}, n \in \mathcal{N}$. *Define*

$$\mathbf{w}^{(n)} := \mathbf{y}^{(n)} \circ \mathbf{x}^{(n)},\ (\mathbf{y}^{(n)})^T\mathbf{x}^{(n)} = 1, n \in \mathcal{N}\,. \tag{3.27}$$

Then, $\bar{\mathbf{p}} = \mathbf{p}(\mathbf{w})$ *whenever* $\mathbf{w} = \mathbf{w}^{(m)}$ *for some* $m \in \bar{\mathcal{N}}$.

Interestingly, this connection is independent of the choice of the utility function Ψ, provided that Ψ fulfills (A.1)–(A.3). The basic idea behind Proposition 5 is illustrated in Figure 3.6. For some Ψ satisfying (A.1)–(A.3), the figure depicts the set $\mathrm{F} \subset \mathrm{Q}^K$ (for the two-user case) defined to be

$$\mathrm{F} = \big\{\boldsymbol{\omega} \in \mathrm{Q}^K : \omega_k = \Psi\big(\mathrm{SIR}_k(\mathbf{p})/\gamma_k\big), k \in \mathcal{K}, \mathbf{p} \in \mathrm{P}_+\big\}\,. \tag{3.28}$$

Note that F can be interpreted as a feasible QoS region where the QoS value for link k is $\Psi(\mathrm{SIR}_k(\mathbf{p})/\gamma_k)$. We define the boundary of F (denoted by $\partial\mathrm{F}$) to be the set of all points of F such that, if $\mathbf{p}$ is the corresponding power vector in (3.28), then $\mathbf{A}\mathbf{p} \le \mathbf{p}'$ holds with at least one equality. Let

$$\bar{\omega}_k = \Psi(\mathrm{SIR}_k(\bar{\mathbf{p}})/\gamma_k), k \in \mathcal{K}$$

and

$$\omega_k(\mathbf{w}) = \Psi\big(\mathrm{SIR}_k(\mathbf{p}(\mathbf{w}))/\gamma_k\big), k \in \mathcal{K}$$

where $\bar{\mathbf{p}}$ and $\mathbf{p}(\mathbf{w})$ are defined by (3.16) and (3.25), respectively. Both

$$\bar{\boldsymbol{\omega}} = (\bar{\omega}_1, \ldots, \bar{\omega}_K) \in \partial\mathrm{F}$$

and

$$\boldsymbol{\omega}(\mathbf{w}) = (\omega_1(\mathbf{w}), \ldots, \omega_K(\mathbf{w})) \in \partial\mathrm{F}_\gamma$$

are boundary points as at least one power constraint is active in the optimum. The figure shows an example of the feasible QoS region under individual power constraints ($\mathbf{A} = \mathbf{I}$). The boundary is given by the following condition: $\boldsymbol{\omega} \in \partial\mathrm{F}_\gamma$ if and only if

$\max_{n\in\mathcal{N}}\rho(\mathbf{Q}(\boldsymbol{\omega})\mathbf{B}^{(n)}) = 1$, where $\mathbf{Q}(\boldsymbol{\omega}) := diag(g(\omega_1), \ldots, g(\omega_K))$ with $g(x) = \Psi^{-1}(x)$. The point $\bar{\boldsymbol{\omega}} = \boldsymbol{\omega}(\mathbf{w}^{(m)})$ for some $m \in \bar{\mathcal{N}}$, corresponds to the unique max-min SIR-balanced power allocation. The weight vector $\mathbf{w}^{(m)}, m \in \bar{\mathcal{N}}$, is normal to a hyperplane which supports the feasible QoS region at $\bar{\boldsymbol{\omega}} \in \partial\mathrm{F}_\gamma$. Note that in Figure 3.6, $\bar{\mathcal{N}} = \{1\}$.

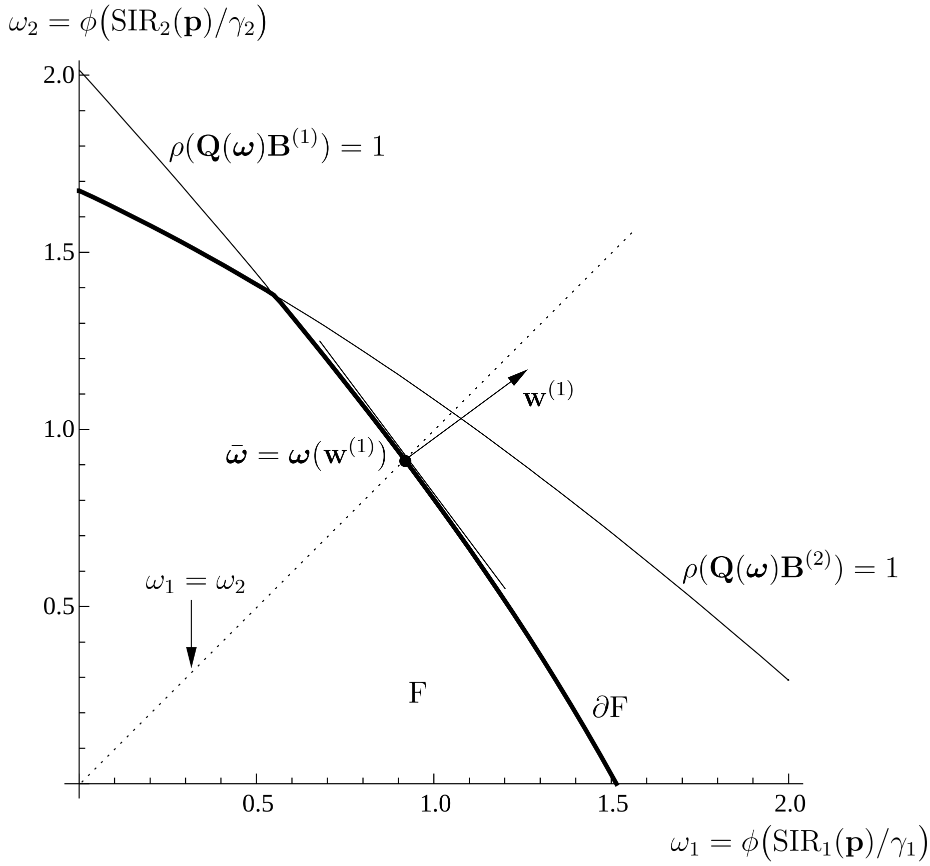

Figure 3.6: An illustration of Proposition 5 [SKB10]. We have $\mathbf{w} = \mathbf{w}^{(1)}$ and $\bar{\mathcal{N}} = \{1\}$: The power constraint of the first link is active under the max-min SIR-balanced power vector $\bar{\mathbf{p}} = \mathbf{p}(\mathbf{w}^{(1)})$.

Next we discuss the properties of this connection in order to derive an algorithmic solution to the joint power control and receive beamforming problem. More precisely, proposition 5 shows how to choose the weight vector $\mathbf{w} \in \Pi_K^+$ in (3.24) so that the max-min SIR-balanced power vector $\bar{\mathbf{p}}$ given by (3.16) is equal to $\mathbf{p}(\mathbf{w})$ given by (3.25). We have the following interesting result that leads to Proposition 5 and is also a key ingredient in the proof of a saddle point of the Perron roots of $\mathbf{B}^{(n)}, n \in \mathcal{N}$ (which are irreducible matrices by assumption): If $\psi(x) = -\Psi(1/x), x > 0$, then [SWB09] (with $n \in \mathcal{N}$)

$$\psi(\rho(\mathbf{B}^{(n)})) \leq G_n(\mathbf{p}, \mathbf{w}) := \sum_{k\in\mathcal{K}} w_k \psi\Big(\frac{(\mathbf{B}^{(n)}\mathbf{p})_k}{p_k}\Big) \tag{3.29}$$

holds for all $\mathbf{p} > 0$ if and only if $\mathbf{w} = \mathbf{w}^{(n)} \in \Pi_K^+$ given by (3.27). Moreover, if $\mathbf{w} = \mathbf{w}^{(n)} \in \Pi_K^+$, then (3.29) holds with equality if and only if $\mathbf{p} = \mathbf{x}^{(n)} > 0$, that is, if and only if $\mathbf{p}$ is the right principal eigenvector of $\mathbf{B}^{(n)}$.

On the other hand, since $\max_{n\in\mathcal{N}} g_n(\mathbf{p}) \leq 1$ for all $\mathbf{p} \in \mathrm{P}_+$, $\psi(x) = -\Psi(1/x), x > 0$, is strictly increasing if (A.1)–(A.3) are satisfied, the diagonal matrix $\mathbf{\Gamma}$ is positive definite and $\mathbf{z} > 0$, we have

$$\begin{aligned} G_n(\mathbf{p}, \mathbf{w}) &= \sum_{k\in\mathcal{K}} w_k \psi\Big(\frac{[\mathbf{\Gamma}(\mathbf{V} + 1/P_n \mathbf{z}\mathbf{a}_n^T)\mathbf{p}]_k}{p_k}\Big) \\ &= \sum_{k\in\mathcal{K}} w_k \psi\Big(\frac{(\mathbf{\Gamma}\mathbf{V}\mathbf{p} + \mathbf{\Gamma}\mathbf{z} g_n(\mathbf{p}))_k}{p_k}\Big) \\ &\leq \sum_{k\in\mathcal{K}} w_k \psi\Big(\frac{(\mathbf{\Gamma}\mathbf{V}\mathbf{p} + \mathbf{\Gamma}\mathbf{z})_k}{p_k}\Big) \\ &= F(\mathbf{p}, \mathbf{w}),\, n \in \mathcal{N}, \mathbf{p} \in \mathrm{P}_+, \mathbf{w} \in \Pi_K^+ \end{aligned} \tag{3.30}$$

with equality if and only if $g_n(\mathbf{p}) = 1$ where

$$F(\mathbf{p}, \mathbf{w}) := \sum_{k\in\mathcal{K}} w_k \psi\Big(\frac{\gamma_k}{\mathrm{SIR}_k(\mathbf{p})}\Big)\,. \tag{3.31}$$

So, considering (3.29), (3.30) and (3.27) shows that (note that $G_n(\mathbf{p}, \mathbf{w})$ has the ray property with respect to $\mathbf{p}$)

$$\begin{aligned} &\psi(\rho(\mathbf{B}^{(n)})) = \min_{\mathbf{p}>0} G_n(\mathbf{p}, \mathbf{w}^{(n)}) \\ &= \max_{\mathbf{w}\in\Pi_K^+} \min_{\mathbf{p}>0} G_n(\mathbf{p}, \mathbf{w}) \\ &= \max_{\mathbf{w}\in\Pi_K^+} \min_{\mathbf{p}\in\mathrm{P}_+} G_n(\mathbf{p}, \mathbf{w}) \\ &\leq \max_{\mathbf{w}\in\Pi_K^+} \min_{\mathbf{p}\in\mathrm{P}_+} F(\mathbf{p}, \mathbf{w}) = \min_{\mathbf{p}\in\mathrm{P}_+} F(\mathbf{p}, \mathbf{w}^*), n \in \mathcal{N} \end{aligned} \tag{3.32}$$

with equality if and only if $g_n(\mathbf{p}(\mathbf{w}^*)) = 1$, where $\mathbf{w}^* \in \Pi_K^+$ is a maximizer of $\min_{\mathbf{p}\in\mathrm{P}_+} F(\mathbf{p}, \mathbf{w})$ and $\mathbf{p}(\mathbf{w}^*)$ is a minimizer of $\mathbf{p} \mapsto F(\mathbf{p}, \mathbf{w}^*)$ over P_+ (see also (3.25)[6]). By (3.26), we know that $\mathcal{N}' = \{m \in \mathcal{N} : g_m(\mathbf{p}(\mathbf{w}^*)) = 1\}$ is a nonempty set. Thus, by (3.30) and (3.32),

$$\psi(\rho(\mathbf{B}^{(n)})) < \psi(\rho(\mathbf{B}^{(m)})) = \min_{\mathbf{p}\in\mathrm{P}_+} F(\mathbf{p}, \mathbf{w}^*) \tag{3.33}$$

[6]Note that any power vector minimizing $\mathbf{p} \mapsto F(\mathbf{p}, \mathbf{w})$ over P_+ for some given weight vector $\mathbf{w} > 0$ is optimal in the sense of the utility maximization problem (3.25) where the utility is defined by (3.24).

for each $m \in \mathcal{N}'$ and $n \notin \mathcal{N}'$. So, by (3.20), we obtain $\mathcal{N}' = \bar{\mathcal{N}}$, which together with (3.29) and (3.33) implies that

$$\mathbf{w}^{(m)} \in \mathrm{W} := \big\{\mathbf{w}^* = \arg\max_{\mathbf{w}\in\Pi_K^+}\min_{\mathbf{p}\in\mathrm{P}_+} F(\mathbf{p},\mathbf{w})\big\}.$$

for each $m \in \bar{\mathcal{N}}$.

Consequently, the max-min SIR-balancing power vector can also be obtained by minimizing $F(\mathbf{p},\mathbf{w}^{(m)})$ for any $m \in \bar{\mathcal{N}}$ over P_+. From (3.33), we see that $\rho(\mathbf{B}^{(m)}) > \rho(\mathbf{B}^{(n)})$ with $m \in \bar{\mathcal{N}}$, $n \notin \bar{\mathcal{N}}$ and $\beta^{(m)} < \beta^{(n)}$. Figure 3.6 illustrates these facts. The line $q_1 = q_2$ intersects both individual boundaries but only the intersection point on $\rho(\mathbf{B}^{(1)}) = 1$ satisfies both power constraints so that we have $\bar{\mathcal{N}} = \{1\}$.

Finally, note that by the above discussion, we can relax the power constraint $\mathbf{p} \in \mathrm{P}_+$ to $g_m(\mathbf{p}) \le 1, \mathbf{p} > 0$, for any $m \in \bar{\mathcal{N}}$. Thus, instead of minimizing $\mathbf{p} \mapsto F(\mathbf{p},\mathbf{w}^{(m)}), m \in \bar{\mathcal{N}}$ over P_+, we can consider

$$\mathbf{p}^{(m)} = \arg\min_{\mathbf{p}>0, g_m(\mathbf{p})\le 1} F(\mathbf{p},\mathbf{w}^{(m)}),\ m \in \bar{\mathcal{N}}\,. \tag{3.34}$$

Alternatively, one can minimize $\mathbf{p} \mapsto G_m(\mathbf{p},\mathbf{w}^{(m)})$ over $\mathbb{R}^K_{++}$, which can be also deduced from (3.32).

3.2.4 Joint Power and Receiver Control

Based on the insights of the previous section we propose an algorithmic solution to find the max-min SIR-balanced resource allocation defined in (3.15) and prove its convergence to the optimum. Instead of directly solving problem (3.15) we find the jointly optimal transmit powers and receive beamformers by incorporating the beamformers into problem (3.34). Hence we need to solve the following utility-based optimization problem

$$\min_{(\mathbf{p},\mathbf{C})\in\mathrm{P}_+\times\mathrm{C}} F(\mathbf{p},\mathbf{C},\mathbf{w})\,, \tag{3.35}$$

where $F(\mathbf{p},\mathbf{C},\mathbf{w})$ is defined in (3.31) with $\mathrm{SIR}_k(\mathbf{p},\mathbf{c}_k)$. The connection between the max-min SIR-balancing power control problem (3.16) and the utility-based power control problem (3.25) assuming a fixed gain matrix $\mathbf{V}(\mathbf{C})$ (i.e. fixed receive beamforming matrix $\mathbf{C}$) was already established in [SKB10] and has been summarized in the previous section. Now, in order to solve (3.35), the basic idea is to perform an iterative algorithm that updates the power vector and the receive beamforming matrix in an alternating way. More precisely, the algorithm keeps one of the variables $\mathbf{C}$ and $\mathbf{p}$ fixed while optimizing with respect to the other. The alternating process is repeated until the termination condition is satisfied. The termination condition may be for instance a predefined number of iterations or a sufficiently small relative improvement $\delta > 0$. The proposed iterative algorithm is summarized in Alg. (4). In the following we describe its single steps in detail.

Algorithm 4 Joint power control and receive beamforming achieving the max-min SIR-balancing solution

Require: $\delta > 0, t = 0, \mathbf{C}(0) \in \mathrm{C}$
repeat
  $t = t + 1$
  determine some $m \in \bar{\mathcal{N}}$
  $\mathbf{w}^{(m)}(t) = c \cdot \mathbf{y}(\mathbf{B}^{(m)}(t-1)) \circ \mathbf{x}(\mathbf{B}^{(m)}(t-1))$
  $\mathbf{p}(t) = \arg\min_{\mathbf{p}>0, g_m(\mathbf{p})\leq 1} F(\mathbf{p}, \mathbf{w}^{(m)}(t))$
  $\mathbf{c}_k(t) = \arg\max_{||\mathbf{c}_k||_2=1} \mathrm{SIR}_k(\mathbf{p}(t), \mathbf{c}_k) \quad \forall_k$
until termination condition is satisfied

Consider iteration t. For some given beamforming matrix $\mathbf{C}(t-1)$ we first update the power vector $\mathbf{p}(t)$. Remember, if we fix the receive beamformers problem (3.35) reduces to a pure power control problem. In order to obtain the max-min SIR-balancing power vector solving the pure power control problem is equivalent to solving problem (3.34) as shown in the previous section. Here, the active power constraint, say m, is determined by (3.20). The weight vector $\mathbf{w}^{(m)}(t)$ is chosen applying proposition 5. The power control problem can be solved by means of a gradient projection algorithm [SWB07].

Then we fix the power vector $\mathbf{p}(t)$ and recalculate the receive beamforming matrix $\mathbf{C}(t)$. Since the k-th receiver influences only the k-th SIR and has no impact on other SIRs we can focus on the single link k. An optimal linear receive beamformer of user k is exactly that beamformer for which the k-th SIR attains its maximum (for a fixed power vector)

$$\mathbf{c}_k(t) = \arg \max_{||\mathbf{c}_k||_2=1} \mathrm{SIR}_k(\mathbf{p}(t), \mathbf{c}_k) \quad \forall_k. \tag{3.36}$$

Then we obtain

$$\min_k \frac{\mathrm{SIR}_k(\mathbf{p}(t), \mathbf{c}_k(t-1))}{\gamma_k} \leq \min_k \frac{\mathrm{SIR}_k(\mathbf{p}(t), \mathbf{c}_k(t))}{\gamma_k}.$$

The proposed algorithm converges to the optimum max-min SIR-balanced resource allocation $(\mathbf{p}^*, \mathbf{C}^*)$ defined in (4). This will be proven in the remainder of this subsection. We start by showing monotonicity.

Lemma 8 *Assume that* $\mathbf{V}(\mathbf{C}(t))$ *is irreducible. Then, the sequence*

$$\beta(t) := 1/\rho(\mathbf{B}^{(m)}(t)) \quad m \in \bar{\mathcal{N}} \tag{3.37}$$

is monotonically increasing in t.

Proof: We have the following

$$
\begin{aligned}
\psi(1/\beta(t)) :&= \psi(\rho(\mathbf{B}^{(m)}(t))) \qquad m \in \bar{\mathcal{N}} \\
&= F(\mathbf{p}(t), \mathbf{w}^{(m)}(t), \mathbf{C}(t)) \\
&\geq F(\mathbf{p}(t), \mathbf{w}^{(m)}(t), \mathbf{C}(t+1))) \\
&\geq \min_{\mathbf{p}>0, g_m(\mathbf{p})\leq 1} F(\mathbf{p}, \mathbf{w}^{(m)}(t+1), \mathbf{C}(t+1)) \\
&= F(\mathbf{p}(t+1), \mathbf{w}^{(m)}(t+1), \mathbf{C}(t+1)) \\
&= \psi(\rho(\mathbf{B}^{(m)}(t+1))) = \psi(1/\beta(t+1))
\end{aligned}
$$

The first inequality follows from the fact that the SIR of user k depends only on the k-th receive beamformer. Thus maximizing the SIR (see (3.36)) decreases the utility measure for a fixed power and weight vector. The second inequality follows from the previous section where it was shown that the max-min SIR-balancing power vector can be obtained by solving (3.34). Now, due to the fact that $\psi(x) = -\Psi(1/x)$ is strictly increasing the sequence $\beta(t)$ is monotonically increasing. ■

Next we conjecture that the monotone sequence $\beta(t)$ converges to the optimum.

Conjecture 1 *The proposed algorithm converges to the optimum max-min SIR-balanced value (3.39), i.e.*

$$\lim_{t\to\infty} \beta(t) = \beta^* \tag{3.38}$$

where

$$\beta^* = \min_{k\in\mathcal{K}} (\mathrm{SIR}_k(\mathbf{p}^*, \mathbf{c}_k^*)/\gamma_k) \tag{3.39}$$

is the minimum SIR value under a max-min SIR-balanced solution (3.15).

It seems that the conjecture can be proven by proceeding essentially as in [BS06b], with the matrix $\mathbf{\Gamma V}$ substituted by $\mathbf{B}^{(m)}$ for some $m \in \bar{\mathcal{N}}$ and the power vectors $\mathbf{p}(t), t \in \mathbb{N}$, normalized such that $g_m(\mathbf{p}(t)) = 1, m \in \bar{\mathcal{N}}$.

3.2.5 Implementation Issues

This section is concerned with practical implementation issues. First of all we review how the receive beamformers and powers can be updated in practice and discuss whether and to what extent a distributed implementation is possible. Then we propose a heuristic algorithm to additionally improve the choice of the transmit beamformers and finally we present a numerical example to show which gains can be achieved by solving the relative max-min optimization problem over powers only and over the joint space of powers and beamformers.

Power and Receiver Updates

We start by reviewing the receive beamformer optimization in practice. It is widely known that an explicit computation of the receive beamformers is intricate due to the lack of perfect synchronization and multipath propagation in many wireless networks. As explained before in section 3.1, adaptive algorithms that converge to the optimal receiver seem to be the only suitable option, especially in distributed wireless networks. Distributed algorithms for computing optimal linear receive beamformers defined by (3.36) are widely established. These algorithms are based either on blind or pilot-based estimation methods as is explained in more detail in subsection 3.1.4.

Next we consider the power updates. As already known, the power control problem can be solved using a gradient projection algorithm that can be implemented in a distributed manner applying the concept of the adjoint network [SWB07]. Unfortunately the calculation of the weight vector is not amenable to distributed implementation as the weights are global variables. An interesting idea to mitigate this problem provides the saddle point characterization of the Perron root of $\mathbf{B}^{(n)}, n \in \bar{\mathbb{N}}$ [SKB10]. This characteri-zation can be used to design a power update where each link updates its weight vector in parallel to the power control recursion. Here the main problem with a distributed implementation comes from the fact that the weight vector needs to be normalized. A partial solution to this problem simplifying the requirements on information exchange provides the recent work of [SKG11]. Both works constitute a basis for developing fully-distributed algorithms in the future.

Finally we emphasize, that in practice it is sufficient to perform some $L_p \geq 1$ steps of the gradient projection algorithm and some $L_r \geq 1$ beamformer updates during the t-th iteration of Alg. 4 to limit the algorithm complexity.

Transceiver Optimization

Similarly to the Alg. 3 presented in section 3.1.6 the algorithm to solve the max-min SIR-balancing problem can be extended, too, with the difference that to compute the powers we use the ingredients summarized in 3.2.4 and in Alg. 4. We point out, that in a network with general power constraints a transceiver optimization has some inherent problems. As mentioned before the uplink-downlink duality does not hold. More considerable is the fact that the power constraints in a general network may not be equivalent in the reversed network. For example, if we consider a downlink that has a sum power constraint, the corresponding reversed network is usually represented by an uplink with individual power constraints for each user. The power constraints for the adjoint network are defined by $\mathrm{Q} = \{\mathbf{q} \in \mathbb{R}_+^K : \mathbf{A}'\mathbf{q} \leq \mathbf{q}', \mathbf{A}' \in \{0,1\}^{N' \times K}\} \subset \mathbb{R}^K$ for some given $\mathbf{q}' = (Q_1, \dots, Q_{N'}) > 0$ and $\mathbf{A}'$ with at least one 1 in each column so that Q is a compact set. N' is the number of power constraints. To be more precise, compared to the primal network, different subsets of links may have (different) sum power constraints. So we cannot ensure an improvement with each iteration step. This should be incorporated in the termination condition.

Numerical Results

We consider a network with K users, $M = 4$ transmit and receive antennas and a channel matrix which entries are iid complex Gaussian distributed. Each user is subject to the same individual power constraint and operates at SNR= 30dB. The SIR targets are chosen to be $\mathbf{\Gamma} = \mathbf{I}$.

Figures 3.7 and 3.8 illustrate for an exemplary network the gain that can be achieved by different resource allocation schemes. More precisely, the max-min SIR-balanced value averaged over 1000 channel realizations is depicted over the number of users in the network for the following 4 resource allocation strategies. If not otherwise stated the beamformers of the schemes adapt to the channel. The first strategy, that is also used as reference strategy in figure 3.8, allocates the maximum power to each user (max. power) whereas the second scheme allocates the max-min SIR power vector to the users (max-min power). The third strategy solves the joint receive beamforming and power allocation problem applying Alg. 4 and the fourth scheme additionally optimizes the transceivers applying Alg. 3. As can be seen a sophisticated resource allocation scheme may significantly improve the network performance. Further we see that in case of a relative max-min objective function, power control seems to be an important mechanism.

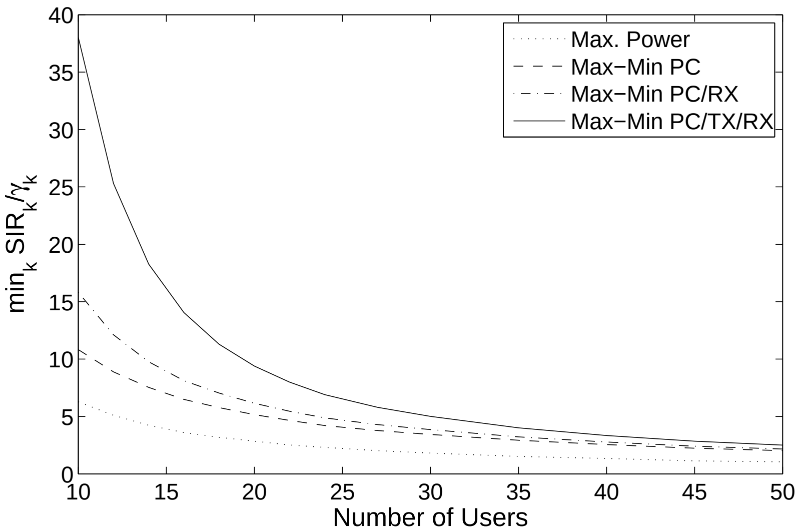

Figure 3.7: Balanced SIR over the number or users for different resource allocation schemes

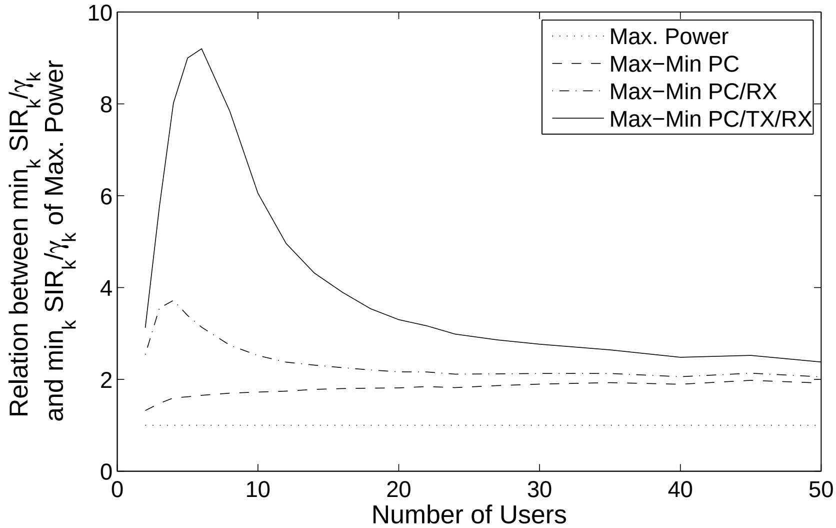

Figure 3.8: Relative improvement with respect to the maximum power scheme (Max. power) for different resource allocation schemes

3.3 Appendix

A. Concavity of (3.4)

We show that $F(\mathbf{s})$ given by (3.4) is not concave in general when $\Psi(x) = \log(x), x > 0$. To this end, consider the 2-user case with $\mathbf{w} = \mathbf{1}$ so that $F(\mathbf{s}) = \sum_{k=1}^{2} \log(e^{s_k} \mathbf{b}_k^H \mathbf{Z}_k^{-1} \mathbf{b}_k)$ where $\mathbf{Z}_1 = e^{s_2}\mathbf{b}_2\mathbf{b}_2^H + \sigma^2\mathbf{I}$ and $\mathbf{Z}_2 = e^{s_1}\mathbf{b}_1\mathbf{b}_1^H + \sigma^2\mathbf{I}$. Define $\|\mathbf{b}_1\|_2^2 = c_{11} > 0, \|\mathbf{b}_2\|_2^2 = c_{22} > 0$ and $|\langle \mathbf{b}_1, \mathbf{b}_2\rangle|^2 = c_{12} \geq 0$. Without loss of generality, assume $\sigma^2 = 1$. Using the Sherman-Morrison formula [Mey00] yields $\mathbf{Z}_1^{-1} = (\mathbf{I} + e^{s_2}\mathbf{b}_2\mathbf{b}_2^H)^{-1} = \mathbf{I} - \frac{e^{s_2}}{1+e^{s_2}c_{22}}\mathbf{b}_2\mathbf{b}_2^H$ and $\mathbf{Z}_2^{-1}$ is obtained in an equivalent way. Therefore,

$$F(\mathbf{s}) = s_1 + s_2 + \log\frac{c_{11} + e^{s_2}\rho}{1 + e^{s_2}c_{22}} + \log\frac{c_{22} + e^{s_1}\rho}{1 + e^{s_1}c_{11}},$$

where $\rho = c_{11}c_{22} - c_{12}$. Assume that $\mathbf{b}_1 \neq \mathbf{b}_2$ and $c_{12} > 0$, in which case $\rho > 0$. Now taking the second derivative either with respect to s_1 or s_2 shows that F is not concave in general.

4 Resource Allocation in Networks with Orthogonal Channels

The medium access control (MAC) dictates how different wireless links share the available resources as power and bandwidth. The wireless spectrum can be divided into different channels that can be assigned to different links by using TDMA, FDMA, OFDMA, CDMA, space division multiple access (SDMA) or hybrid combinations of these methods. The higher the number of simultaneously active links, the higher the level of interference occurring. On the other hand, the performance can be significantly improved by combining power control with link scheduling. These are two central mechanisms for resource allocation and interference management. In this chapter we assume that the links are scheduled such that no interference occurs meaning that all simultaneously active links are orthogonal to each other. More precisely, we consider the problem of joint power control and link scheduling in general wireless networks with orthogonal channels such that the sum of weighted link rates is maximized. The presented results have been partly published in [7, 1] in 2004 and 2005.

In the last years, similar works have emerged further extending and generalizing our results. To the most important ones belong [ASB10, HSAB09b, HSAB09a, HSBA10] that maximize the weighted throughput for CDMA and OFDMA networks, respectively. In contrast to our work, these works use a dual formulation to exploit several structural properties. Applying these properties the authors propose optimal and heuristic algorithms with low complexity. For further related work on the resource allocation topic in orthogonal networks the reader is referred to the references in [ASB10, HSBA10].

As mentioned above, in this chapter the objective is to maximize a weighted sum of link rates. Because of power constraints and a limited bandwidth, we devise strategies for allocating powers and signature sequences / OFDMA-subchannels to nodes depending on their weight and channel states. The chapter is organized as follows: In sections 4.1 and 4.2 we introduce the system model and the problem, respectively. Due to fact that we face a discrete or integer problem, in section 4.3 we first relax the problem and give an iterative solution providing an upper bound to the original discrete problem. Then in section 4.4 we investigate the structure of the discrete problem and derive a simple heuristic algorithm. The weights may depend on some network variables to ensure network stability. In a special case, when the weights are proportional to the queue lengths, such strategies are called throughput optimal [YC03]. In section 4.5 we shortly discuss the choice of the weights of such throughput-optimal strategies in order to also improve the delay performance. Finally in section 4.6 we give an example to decentrally implement our heuristic algorithm and show some numerical results.

4.1 System Model, Assumptions and Definitions

Wireless scheduling approaches can be categorized in TDMA systems and systems that combine TDMA with another multiplexing technique such as CDMA or OFDMA where the transmitter can simultaneously transmit to multiple users in each time slot. Ensuring that links are mutually orthogonal is one way to simplify resource allocation problems. However, establishing and maintaining the orthogonality in wireless networks is a tricky task, especially in the cases of uplink or a general network topology, due to in particular asynchronous multi-path channels of the different links. To be more precise, in OFDMA a proper frequency and timing synchronization is necessary to maintain orthogonality among the active users. On the positive side, in CDMA systems the problem of synchronization can be circumvented by using quasi-synchronous CDMA. Here it is sufficient to allocate signature sequences (or spreading sequences) with zero aperiodic correlation sidelobes within a small window around the zero shift. The window size is determined by the delay of the multiple paths. The drawback is that the number of such aperiodic zero correlation window (AZCW) sequences is strictly limited and inversely proportional to the window size. The interested reader is referred to [FH00, SBH01, SB02], where the problem of designing signature sequences for QS-CDMA systems has been addressed in more detail.

In this chapter we assume that the users are orthogonal. We consider a general wireless network with N nodes and K links that are restricted to the corresponding general power constraints defined in equation (1.4). For each node $n \in \mathcal{N}$ we also remind the definition $\mathcal{K}_n := \{m \in \mathcal{K} : a_{n,m} = 1\} \subseteq \mathcal{K}$ as a set of all active logical links originating in node n. We point out, that the SIR model simplifies due to the orthogonality. That is why we introduce a slightly different notation for this chapter. The channel gain state on link k normalized by its noise power is denoted by $h_k, \forall_{k\in\mathcal{K}}$ and depends on the path attenuation and the multi-path channel coefficients. In case of OFDMA we assume randomly grouped subchannels or a frequency hopping scheme where the subchannels have approximately the same channel quality. Hence, in the OFDMA case the channel vector reduces to a single value that depends only on the link k which is naturally the case using CDMA. Now, assigning one subchannel or one signature sequence to a user it is easy to see that the SIR of link k is

$$\mathrm{SIR}_k = p_k h_k,\ 1 \leq k \leq K. \tag{4.1}$$

Throughout this chapter we will typically talk of signature sequences instead of subchannels. However, due to the system model, we may use the terms subchannel or signature sequence interchangeably.

Queuing Model

Packets are assumed to arrive at the network according to an ergodic stationary process with rates $\lambda_r, 1 \leq r \leq R$ bits per second. The packet lengths are i.i.d. with finite mean and variance. Due to a limited number of signature sequences and power constraints, data packets are not transmitted immediately after their arrivals. The matrix of queue backlogs (expressed in bits) is denoted by $\mathbf{Q} = (q_{ni})_{1 \leq n,i \leq N} \geq 0$, where q_{ni} is the backlog of bits at node n that need to be sent to node i.

Data Rate Model

We consider a channel in which each link can be assigned no, one or several signature sequences/OFDMA-subchannels. Given an arbitrary block, let s_k be the number of signature sequences/subchannels assigned to link k. Clearly, because there are $M > 0$ available signature sequences/subchannels, we need to have $\mathbf{s} = (s_1, \ldots, s_K) \in \mathrm{S}$, where

$$\mathrm{S} := \{\mathbf{s} \in \mathbb{N}_0^K : \|\mathbf{s}\|_1 \leq M\}\,.$$

In addition to p_k and h_k, the data rate on link k depends on s_k. If $s_k = 1$, then the data rate on link k can be often modeled as $a\log(1 + b\mathrm{SIR}_k)$ for some positive constants a, b [NMR03]. For brevity and as before we assume $a = b = 1$. Thus, if there are $s_k \in \mathbb{N}_0$ signature sequences on link k, it follows from (4.1) that

$$r_k = s_k \log\big(1 + \frac{p_k h_k}{s_k}\big),\ 1 \leq k \leq K\,, \tag{4.2}$$

where $0\log(1 + \frac{p_k h_k}{0}) = 0$. Note that r_k is strictly increasing with respect to both $p_k \geq 0$ and $s_k \geq 0$. Furthermore we observe the following:

Observation 11 *Assuming* $\mathbf{s} \in \mathbb{R}^K$. *Then* $r_k(p_k, s_k)$ *defined in (4.2) is a concave function of* (p_k, s_k).

Proof: To see this define

$$f(x, y) = y\log\Big(1 + \frac{c\,x}{y}\Big),\ x, y \geq 0$$

and let $z = (x, y)$ and $z(\mu) = (1-\mu)\hat{z} + \mu\check{z}, 0 \leq \mu \leq 1$ for some arbitrary $\hat{z}, \check{z} \in D = \mathrm{P} \times \mathrm{S}$. We calculate the second derivative of $\tilde{f}(z(\mu)) = f(x, y)$ with $y(\mu) > 0$ to obtain

$$\frac{d^2\tilde{f}(z(\mu))}{d\mu^2} = -\frac{c^2(\check{x}\hat{y} - \hat{x}\check{y})^2}{y(\mu)[y(\mu) + cx(\mu)]^2} \leq 0,\ y(\mu) > 0,$$

for every $0 \leq \mu \leq 1$ and $\hat{z} = (\hat{x}, \hat{y}), \check{z} = (\check{x}, \check{y}) \in D$. Thus, since $\tilde{f}(z)$ is continuous on D, $f(x, 0) = 0$ for any $x \geq 0$, and the Cartesian product of convex sets P and S is convex, we can conclude that $f(x, y)$ is jointly concave. ■

4.2 Problem Statement

In this section, we assume that there is a central network controller having full channel and weight knowledge of each node. More precisely, the controller knows the channel and weight state of each logical link or user. The problem is to maximize a weighted sum of link rates.

At the beginning of every time block, the network controller chooses optimal link transmission rates by allocating transmit powers and signature sequences to the links. To be precise, consider an arbitrary time block and let $w_1, \ldots, w_K$ be given positive weights. Then, the network controller determines $\mathbf{p}^{\text{opt}} \in \text{P}$ and $\mathbf{s}^{\text{opt}} \in \text{S}$ such that

$$(\mathbf{p}^{\text{opt}}, \mathbf{s}^{\text{opt}}) = \arg\max_{\mathbf{p} \in \text{P}, \mathbf{s} \in \text{S}} \sum_{n=1}^{N} F_n(\mathbf{p}_n, \mathbf{s}_n) \tag{4.3}$$

where $\mathbf{s}_n = (s_k)_{k \in \mathcal{K}_n}$ represents the vector of the number of signature sequences allocated to links originating at node n and

$$F_n(\mathbf{p}_n, \mathbf{s}_n) = \sum_{k \in \mathcal{K}_n} w_k s_k \log\big(1 + \frac{p_k h_k}{s_k}\big). \tag{4.4}$$

In a special case, when the weights are equal to the differential queue backlogs (see section 4.5), this approach can be shown to stabilize a wireless network if the link arrival rates are strictly interior to the set of all feasible link rates [NMR03, YC03]. Recall that a wireless network is said to be stable if the probability for any queue length to be infinity is equal to zero.

The problem described in (4.3) is an integer programming problem. Consequently, to find the optimal solution may be a time consuming task. Therefore, in the following we first derive an upper bound and then propose a simple heuristic solution resulting from the problem properties.

4.3 Upper Bound - Problem Relaxation

To derive an upper bound, let $\bar{\text{S}} := \{\mathbf{s} \in \mathbb{R}^K : \|\mathbf{s}\|_1 \leq M\}$ and note that $\text{S} \subset \bar{\text{S}}$. As an immediate consequence it follows that

$$\max_{\mathbf{p} \in \text{P}, \mathbf{s} \in \text{S}} \sum_{n=1}^{N} F_n(\mathbf{p}_n, \mathbf{s}_n) \leq \max_{\mathbf{p} \in \text{P}, \mathbf{s} \in \bar{\text{S}}} \sum_{n=1}^{N} F_n(\mathbf{p}_n, \mathbf{s}_n).$$

In words, an upper bound is obtained assuming that $\mathbf{s} \in \bar{\mathrm{S}}$ is a real-valued sequence allocation vector. Of course, the relaxation implies that the throughput performance can be further improved by employing a suitable time-sharing protocol that allows nodes to share signature sequences / subchannels within one block.

To the best of our knowledge there is no closed-form solution for the relaxed version of Problem (4.3). However, since $F_n(\mathbf{p}, \mathbf{s})$ defined by (4.4) is jointly concave (note that a weighted sum of concave functions is concave [BV04]) and bounded on the convex set $\mathrm{P} \times \mathrm{S}$, the relaxed problem results in a tractable convex optimization problem. Hence, we could use some of the standard optimization methods such as interior point methods to solve this problem numerically. Independently and simultaneously to our work the authors of [ASB10] used a dual approach to characterize the solution, though at this time for a downlink CDMA network only. In the following section, we describe an alternative iterative algorithm that alternately maximizes the rate function with respect to one of the vectors while keeping constant the other one until the algorithm converges.

4.3.1 Iterative Algorithm

In detail, holding the sequence and power allocation alternatingly constant new power and sequence allocation vectors are calculated (see Alg. 5). The algorithm generates a strictly increasing sequence $F(\mathbf{p}^{(t)}, \mathbf{s}^{(t)})$, and converges to the global maximum $\mathbf{p}^{\mathrm{opt}}$ and $\mathbf{s}^{\mathrm{opt}}$ of the relaxed problem (4.3).

Algorithm 5 Dynamic resource allocation

Require: $t := 0, \mathbf{p}^{(0)} \in \mathrm{P}, \mathbf{p} \neq 0$

repeat
  $t := t + 1$
  $\mathbf{s}^{(t)} := \arg\max_{\mathbf{s} \in \mathrm{S}} F(\mathbf{p}^{(t-1)}, \mathbf{s})$
  $\mathbf{p}^{(t)} := \arg\max_{\mathbf{p} \in \mathrm{P}} F(\mathbf{p}, \mathbf{s}^{(t)})$
until $|F(\mathbf{p}^{(t)}, \mathbf{s}^{(t)}) - F(\mathbf{p}^{(t-1)}, \mathbf{s}^{(t-1)})| < \epsilon$ for $\epsilon > 0$

Remark: We start with any point in the interior of $\mathrm{P} \times \mathrm{S}$. If $p_k^{(t)} = 0$ for some k and t, then we set $p_k = s_k = 0$, and remove this user from the optimization (such users are allocated no resources).

Optimal power allocation for a given sequence allocation

Suppose that $\mathbf{s} \in \mathrm{S}$ is given. By the remark above, we can assume that $\mathbf{s} > 0$. Let

$$\mathbf{p}^* = \arg\max_{\mathbf{p} \in \mathrm{P}} F(\mathbf{p}, \mathbf{s}). \tag{4.5}$$

Obviously, the problem is convex and Slater's conditions are satisfied implying that the Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient for the global optimum. Moreover, it is easy to see that in the case of sum power constraints $\|\mathbf{p}^*\|_1 = P_t$ must hold. Writing $F(\mathbf{p}, \mathbf{s})$ as

$$F(\mathbf{p}, \mathbf{s}) = \sum_k w_k s_k \log\big(p_k + \frac{s_k}{h_k}\big) + \sum_k w_k s_k \log \tfrac{h_k}{s_k}$$

reveals that (4.5) is a weighted version of the standard water-filling problem [BV04, Page 245]. Hence,

$$p_k^* = \max\Big\{0, s_k\big(w_k/v^* - 1/h_k\big)\Big\}, h_k > 0, \tag{4.6}$$

where the optimal dual variable v^* can be easily obtained from

$$\sum_k \max\big\{0, s_k\big(w_k/v^* - 1/h_k\big)\big\} = P_t.$$

We emphasize that since $s_k > 0$, $p_k^* = 0$ if $v^* > h_k \cdot w_k$ meaning that in this case user k are assigned no resources. Otherwise, we have $p_k > 0$. Now, in a special case when a certain amount of resources is assigned to each user meaning $p_k^* > 0$ for each $1 \leq k \leq K$, we have

$$p_k^* = w_k s_k \frac{P_t + \sum_{l=1}^K \frac{s_l}{h_l}}{\sum_{l=1}^K s_l w_l} - \frac{s_k}{h_k}.$$

The solution of the power control problem can be easily generalized to the case of general power constraints where we have $1 < N \leq K$ power constraints: Here, instead of one dual variable we have a vector of dual variables $\mathbf{v} = (v_1, \ldots, v_N)$. The optimal power is obtained by

$$p_k^* = \max\Big\{0, s_k\big(w_k/v_n^* - 1/h_k\big)\Big\}, h_k > 0, k \in \mathcal{K}_n, \forall_n \in \mathcal{N} \tag{4.7}$$

where the optimal dual variable v_n^* can be easily obtained from

$$\sum_{k\in\mathcal{K}_n} \max\big\{0, s_k\big(w_k/v_n^* - 1/h_k\big)\big\} = P_n, \forall_n \in \mathcal{N}.$$

Optimal sequence allocation for a power allocation

Given $\mathbf{p} \in \mathrm{P}$ with $\mathbf{p} > 0$, let

$$\mathbf{s}^* = \arg\max_{\mathbf{s}\in\mathrm{S}} F(\mathbf{p}, \mathbf{s}).$$

As before, the KKT conditions are necessary and sufficient for the global optimum (note that if $p_k > 0$, then $s_k^* > 0$). Furthermore, $\|\mathbf{s}^*\|_1 = M$. Eliminating the slack variables from the KKT conditions, we are left with $s^* > 0, \|\mathbf{s}^*\|_1 = M$,

$$v^* \geq w_k \log\Big(1 + \frac{p_k h_k}{s_k^*}\Big) - \frac{w_k h_k p_k}{s_k^* + p_k h_k}, \tag{4.8}$$

and

$$s_k^*\Big[v^* + \frac{w_k h_k p_k}{s_k^* + p_k h_k} - w_k \log\Big(1 + \frac{p_k h_k}{s_k^*}\Big)\Big] = 0. \tag{4.9}$$

The right-hand side of (4.8) tends to $+\infty$ as $s_k^* \to 0$. Thus with v^* chosen so that $\|\mathbf{s}\|_1 = M$, $\mathbf{s}^* > 0$ given by

$$v^* + w_k \frac{p_k h_k}{s_k^* + p_k h_k} - w_k \log\Big(1 + \frac{p_k h_k}{s_k^*}\Big) = 0 \tag{4.10}$$

satisfies the KKT conditions. Using $u_k = (\frac{-s_k^*}{s_k^* + p_k h_k})$ and $z_k = (\frac{v^*}{w_k} + 1)$, we can write (4.10) in form of the inverse Lambert W-function

$$u_k \exp^{u_k} = -\exp^{-z_k},$$

from which the optimal solution $\mathbf{s}^*$ can be obtained by determining first u_k, $\forall_k$. Lambert's W-function is defined to be the inverse of the function $f : \mathbb{C} \to \mathbb{C}$ given by $f(x) := x \exp^x$. That is, W(x) is the complex valued function that satisfies $W(x) \exp^{W(x)} = x$, $\forall x \in \mathbb{C}$. In our case $x = -\exp^{-z_k}$.

4.4 Approximations of the Integer-Programming Problem

One way to find an integer solution is to use the optimal real solution and round the sequence allocation vector in an appropriate way such that the constraint M is met. Next we will propose another heuristic solution by analyzing the problem structure of (4.3). This heuristic is not based on the optimal solution of the relaxed problem. To solve the problem in (4.3), note that if the maximum in (4.3) is attained, we have both $\|\mathbf{s}\|_1 = \sum_k s_k = M$ and $\|\mathbf{p}_n\|_1 = \sum_{k \in \mathcal{K}_n} p_k = P_n$ for each $1 \le n \le N$. This is simply because data rate r_k is continuous and strictly increasing in both $s_k \ge 0$ and $p_k \ge 0$. Thus, we can rewrite (4.3) to obtain

$$\max_{\mathbf{p} \in \mathrm{P}, \mathbf{s} \in \mathrm{S}} \sum_{n=1}^{N} F_n(\mathbf{p}_n, \mathbf{s}_n) \qquad \text{s.t.} \qquad \begin{cases} ||\mathbf{s}_n||_1 = S_n \quad \forall_k \\ \sum_{n=1}^{N} S_n = M \end{cases}. \tag{4.11}$$

Clearly, for any fixed $S_1, \dots, S_N$, Problem (4.11) decomposes into N independent constrained optimization problems, one for each node:

$$(\mathbf{p}_n^*, \mathbf{s}_n^*) = \operatorname*{arg\,max}_{\mathbf{p}_n \in \mathrm{P}_n, \mathbf{s}_n \in \mathrm{S}_n} F_n(\mathbf{p}_n, \mathbf{s}_n), \tag{4.12}$$

where $\mathbf{p}_n^* := \mathbf{p}_n^*(S_n)$, $\mathbf{s}_n^* := \mathbf{s}_n^*(S_n)$, $\mathrm{S}_n := \{\mathbf{s}_n \in \mathbb{N}_0^{|\mathcal{K}_n|} : \|\mathbf{s}_n\|_1 = S_n\}$ and finally $\sum_n S_n = M$ must hold. Thus, an optimal power and sequence allocation $(\mathbf{p}^{\text{opt}}, \mathbf{s}^{\text{opt}})$ follows from

$$(\mathbf{p}_n^{\text{opt}}, \mathbf{s}_n^{\text{opt}}) = \big(\mathbf{p}_n^*(S_n^{\text{opt}}), \mathbf{s}_n^*(S_n^{\text{opt}})\big), 1 \leq n \leq N$$

where

$$(S_1^{opt}, \ldots, S_N^{opt}) = \underset{\substack{S_1,\ldots,S_N \geq 0 \\ \sum_n S_n = M}}{\arg\max} \sum_{n=1}^{N} F_n\big(\mathbf{p}_n^*(S_n), \mathbf{s}_n^*(S_n)\big). \tag{4.13}$$

Now suppose that $\mathbf{p}_n^*(S_n)$ and $\mathbf{s}_n^*(S_n)$ from (4.12) are known for each S_n and $1 \leq n \leq N$. Problem (4.13) can be written as a Multiple-Choice Knapsack Problem (MCKP)

$$\begin{aligned} \mathbf{X}^* = & \underset{\mathbf{X} \in \{0,1\}^{N \times (M+1)}}{\arg\max} \sum_{n=1}^{N} \sum_{i=0}^{M} F_n\big(\mathbf{p}_n^*(i), \mathbf{s}_n^*(i)\big) x_{ni} \\ & s.t. \sum_{n=1}^{N} \sum_{i=0}^{M} i x_{ni} \leq M \text{ and } \forall_n \sum_{i=0}^{M} x_{ni} = 1. \end{aligned} \tag{4.14}$$

Then $S_n^{opt} = i \iff x_{ni} = 1,\ 1 \leq n \leq N, 0 \leq i \leq M$. MCKP problems are weakly NP-hard and can be solved in pseudo-polynomial time through dynamic programming [Pis95]. However, the algorithm for the MCKP problem greatly simplifies if $\forall_n\ F_n\big(\mathbf{p}_n^*(i), \mathbf{s}_n^*(i)\big) - F_n\big(\mathbf{p}_n^*(i-1), \mathbf{s}_n^*(i-1)\big)$ is decreasing with increasing $i, 1 \leq i \leq M$. Unfortunately this property is not always satisfied in general due to the limitation of $\mathbf{s}_n$ on integer values, but we observed that the cases in which the property is not satisfied are quite rare. Therefore, we propose the simplified algorithm for the heuristic (see Alg. 6) and denote its solution as $\hat{S}_n, 1 \leq n \leq N$.

Now return to Problem (4.12). Obviously we can solve (4.12) by checking all the positive integer valued vectors $\mathbf{s}_n \in \mathrm{S}_n$ which may be quite complex. Another way is to solve (4.12) under the assumption that $\mathbf{s}_n \in \bar{\mathrm{S}}_n$, $\bar{\mathrm{S}}_n := \{\mathbf{s}_n \in \mathbb{R}^{|\mathcal{K}_n|} : \|\mathbf{s}_n\|_1 = S_n\}$ is relaxed and can be found using the algorithm presented in the previous section. Searching in the neighborhood $(\mathbf{s}_n^* - 1, \mathbf{s}_n^* + 1)$ of $\mathbf{s}_n^* \in \bar{\mathrm{S}}_n$ seems to yield good results by simulations, but is not optimal in general. Therefore we propose to assign the resources S_n und P_n to the link that maximizes

$$\hat{k} = \underset{k \in \mathcal{K}_n}{\arg\max} \quad w_k S_n \log\big(1 + \frac{P_n h_k}{S_n}\big).$$

and denote the heuristic solution as $(\hat{\mathbf{p}}_n, \hat{\mathbf{s}}_n)$. Note, that in a network with a sum power constraint which maximizes the total throughput ($\forall_l w_l = 1$) it is optimal to assign the resources completely to the link with the best channel. We also emphasize, that the authors of [HSBA10] proposed the same heuristic to decide which user is assigned to a certain bandwidth resource.

The analysis of the problem structure and the assumptions we have made in order to find a simple heuristic algorithm to determine $\big(\hat{\mathbf{p}}_n(\hat{S}_n), \hat{\mathbf{s}}_n(\hat{S}_n)\big)$ result in the algorithm outlined finally (see Alg. 6).

Algorithm 6 Heuristic centralized algorithm

Require: $\forall_n \hat{S}_n = 0, \hat{\mathbf{p}}_n(0) = (0), \hat{\mathbf{s}}_n(0) = (0)$
for $n = 1 : N$ **do**
  calculate $\big(\hat{\mathbf{p}}_n(1), \hat{\mathbf{s}}_n(1)\big)$
  $d_n = F_n\big(\hat{\mathbf{p}}_n(1), \hat{\mathbf{s}}_n(1)\big) - F_n\big(\hat{\mathbf{p}}_n(0), \hat{\mathbf{s}}_n(0)\big)$
end for
for $i = 1 : M$ **do**
  $\hat{n} = \arg\max_n d_n$
  $\hat{S}_{\hat{n}} = \hat{S}_{\hat{n}} + 1$
  calculate $\big(\hat{\mathbf{p}}_{\hat{n}}(\hat{S}_{\hat{n}} + 1), \hat{\mathbf{s}}_{\hat{n}}(\hat{S}_{\hat{n}} + 1)\big)$
  $d_{\hat{n}} = F_{\hat{n}}\big(\hat{\mathbf{p}}_{\hat{n}}(\hat{S}_{\hat{n}} + 1), \hat{\mathbf{s}}_{\hat{n}}(\hat{S}_{\hat{n}} + 1)\big) - F_{\hat{n}}\big(\hat{\mathbf{p}}_{\hat{n}}(\hat{S}_{\hat{n}}), \hat{\mathbf{s}}_{\hat{n}}(\hat{S}_{\hat{n}})\big)$
end for

4.5 Choice of the Weights

In this section we shortly discuss the influence of the weights. Interestingly, by appropriately choosing the weights, one can stabilize the network in case of elastic traffic. On the other hand, in multi-hop wireless networks the weights can be chosen to reduce the delay.

4.5.1 Network Stability

Now we are able to briefly introduce the concept of stability and review a necessary condition for network stability. As mentioned in Section 4.1, packets are assumed to arrive according to an ergodic stationary process. Let $Q_k(t)$ denote the number of unprocessed bits in queue k at time t. Define the overflow function as

$$g_k(V) = \limsup_{t\to\infty} E\Big\{\frac{1}{t}\int_0^t 1_{Q_k(t)>V} dt\Big\}, 1 \le k \le K,$$

for some $V \ge 0$. In other words, $g_k(V)$ is the probability that the queue length is larger than V. Now a network is said to be stable if

$$\lim_{V\to+\infty} g_k(V) = 0,\ 1 \le k \le K.$$

Roughly speaking, in a stable network, the probability that a queue blows up to infinity is zero.

Now, maximizing the weighted sum of link rates as defined by policy (4.3) we can state the following. In a special case, when the weights are equal to the differential queue backlogs (see section 4.5.2), this approach (4.3) can be shown to stabilize a wireless network if the link arrival rates are strictly interior to the set of all feasible link rates [NMR03, YC03].

Numerical Example

In this section we give a numerical example for the downlink of a single cell to compare the throughput optimal policy (4.3) with three other heuristic approaches. As a performance metric, we consider the normalized total average queue size over an arrival rate parameter λ (see Figure 4.1). We assume an asymmetric traffic scenario (4 user groups with $\lambda_1 = 0.4\lambda, \lambda_2 = 0.8\lambda, \lambda_3 = 1.2\lambda, \lambda_4 = 1.6\lambda$) with $K = 16$ users, $M = 8$ signature sequences. The arrival processes are independent Poisson distributed. The LQ (longest queue) policy assigns M signature sequences to M users with the longest queues. The power allocation is equal to (4.6) with $h_1 = \ldots = h_K = 1$. The BC (best channel) policy assigns M signature sequences to M users with the best channels. The power allocation is equal to (4.6). In case of CPT, the resources are evenly distributed among the users.

In comparison with the heuristic approaches, the throughput optimal policy provides significantly larger stability region. To reduce the average queue size under light traffic (small values of λ), the optimal policy determines a new resource allocation whenever one of the queues becomes empty within a time-slot. The performance loss of BC and LQ policies is strongly influenced by the channel statistics and data traffic scenario (asymmetric or symmetric). It turns out that under more realistic traffic and channel scenarios, the LQ policy guarantees lower delays and better stability conditions than the BC strategy. Note that in case of symmetric traffic, the performance of the BC policy (under the additional condition $s_k \leq M, s_k \in \mathbb{N}_0$) approaches that of the throughput optimal policy provided that all users have the same average channel conditions.

4.5.2 Delay Reduction

As mentioned, in multi-hop networks, it is reasonable to choose the weights depen-ding on the differential backlogs, which ensures that packets are moved to destination in directions of decreasing backlogs [TE92a, NMR03]. In order to reduce the delays we additionally may take into account the number of hops from source to destination [NMR03]. In particular, the weights for each link $k := (n, j)$ between node n and node j may be chosen according to

$$w_k = \max[\max_{i\in\{1,\ldots,N\}} \big((q_{ni} - q_{ji}) + \alpha(z_{ni} - z_{ji})\big), 0], k \in \mathcal{K}_n. \tag{4.15}$$

Here and hereafter z_{ni} is equal to the number of hops from source n to destination node i. Thus, data packets are urged to move into the directions of the maximum differential backlogs while being inclined to choose the shortest path route. The factor α prioritizes the route information and may be appropriately chosen in the range of the maximum possible

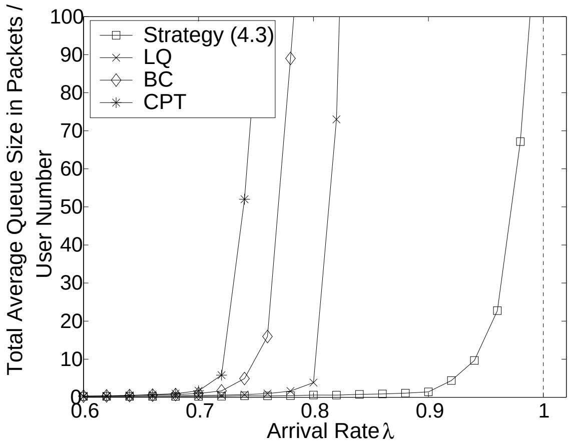

Figure 4.1: Total average queue size vs. arrival rate

transmission rate of a node [NMR03]. We emphasize that (4.15) decides inherently on the route taken by some data packet belonging to a certain source-destination pair.

However, this is only a heuristic algorithm decreasing the delay and avoiding routing loops or unnecessary long paths. The authors of [YSR09] proposed an algorithm (joint traffic-splitting and shortest-path-aided back-pressure algorithm) that in contrast to [NMR03] provably minimizes the average path-lengths. More precisely their algorithm uses long paths only when necessary and thus results in small end-to-end delays. The reader will however realize that each node needs to maintain at most $N(N-1)$ queues. As a remedy, the authors suggest to introduce the cluster-technique proposed in [YST08] to reduce the queue complexity. In [BSS09] the authors deal with both problems. On the one hand they introduce the concept of shadow queues to reduce the number of real queues to be maintained at the nodes. On the other hand they modify the back-pressure algorithm - meaning the choice of the weights - in order to improve the delay performance.

4.6 Resource Allocation: Decentralized Approximation

In general the solution of Problem (4.11) requires a central network controller having the global knowledge of all queue backlogs and channel states. This is also true for the heuristic approach presented in section 4.4, which approximates the optimal solution given by (4.12) and (4.13). In this section, we shortly summarize a possible decentralized implementation to approximate the centralized strategy in a wireless network with a general network topology. Further details can be found in [1].

Summarizingly we propose the following decentralized scheme outlined in Alg. 7.

In detail, we first suppose that the network is based on a clustered organization (step (i)). Due to the properties of AZCW sequences a coarse synchronization between the clusters is

Algorithm 7 Decentralized resource allocation

(i) Introduce a clustered network organization.

(ii) Assignment of a number of signature sequences to each cluster.

(iii) Contention phase for node activation.

(iv) Cluster heads allocate powers and signature sequences determined by Alg. 6 to the cluster nodes activated.

required. We suggest to use the partitioning of the network to assign the available signature sequences among the clusters (step (ii)). In the following we focus on small networks [LG97] and refer to [L.H93] concerning the code assignment with spatial reuse. Further, so far the centralized strategies for resource allocation have been obtained under the assumption that nodes can transmit and receive simultaneously. Unfortunately, in practice this assumption is rarely satisfied. Thus we need a simple decentralized mechanism to choose a set of links that can be activated concurrently (step (iii)). In the decentralized network we suggest to activate nodes with all its outgoing links by a contention phase. Now, provided that a node activation set is chosen, the resources have to be finally assigned to the corresponding links (step (iv)). Under the assumption that the active nodes inform their cluster heads about their link states and weights calculated from (4.15) the cluster head can determine the resource allocation using Alg. 6. We assume that each node n knows the channel states on its outgoing links $k \in \mathcal{K}_n$ and the neighboring queue states. To exchange this control data all nodes are provided with a low rate control channel. Additionally we point out, that for a centralized network the link activation to support the half-duplex communication can be embedded in Algorithm 6 due to the fact that the signature sequences are allocated one by one. Then, the destination node of the link activated and all its outgoing links can be simply removed in each step.

Numerical Example

Finally we present some simulation results to compare the centralized heuristic po-licy with the distributed approximations and with the relaxed optimal solution (upper bound). As a performance measure we consider the total average queue size over an arrival rate parameter $\boldsymbol{\lambda}$ (see Figure 4.2). We restrict the simulations to a small network without spatial code reuse. We assume a network of 16 nodes distributed over a discretized grid with 16 source-destination pairs whereas each node can serve both as source and as destination. During the simulation the nodes can move within a predetermined area. We assume an asymmetric traffic scenario (14 sources with λ_1 and 2 sources with $\lambda_2 = 5\lambda_1$), a channel with 3 paths, a relative time offset parameter $d = 4$ and a spreading gain of 32. Consequently there are $M = 8$ signature sequences (AZCW) of length 40 chips. The required uncoded bit error rate is $BER = 10^{-3}$. In comparison with the optimal solution based

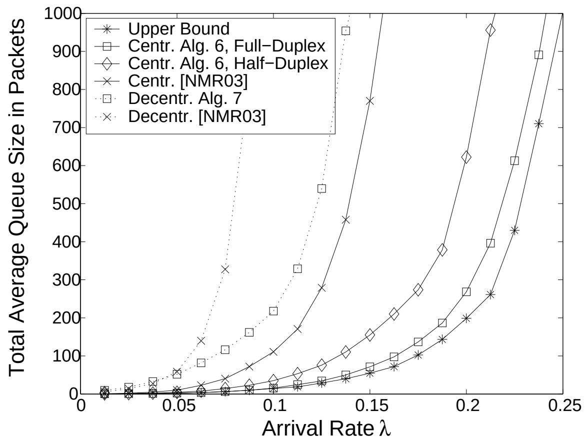

Figure 4.2: Total average queue size in packets vs. arrival rate

on time-sharing the proposed heuristic algorithm (Alg. 6) provides a considerable performance. This is due to the fact that the simple algorithm (Alg. 6) exploits the problem properties of (4.3). In contrast, compared with the centralized half-duplex strategy the decentralized strategy (of Alg. 7) suffers as expected from a performance loss. This is an immediate consequence of the node activation strategy that blocks nodes and therefore links that may be used otherwise.

Furthermore we compare the performance offered by orthogonal QS-CDMA channels with the centralized and distributed implementation of DRPC (Distributed Routing and Power Control) from [NMR03] for a system with random sequences of length 32 chips. Here, the near-far effect is mitigated because the nodes transmit data to the receiving neighbor node which maximizes $w_n \log(1 + SIR_n)$. To determine this neighbor in the distributed case, the receiving nodes which are randomly chosen measure the interference of the transmitting nodes and send this information back to all neighbors. A comparison of networks using either AZCW or random signature sequences depends strongly on the required bit error rate, the spreading gain and the parameter d which includes the coarseness of the synchronization and the multi-path propagation. For the defined parameters the simulations show that the scheme based on AZCW sequences provides larger stability region than the scheme from [NMR03] with random sequences. If $d = 8$, resulting in $M = 4$ signature sequences, both schemes provide a similar performance. Generalizing, if a low bit error rate for data applications is required and it can be ensured that d is much smaller than the spreading gain N, $d \ll N$, then orthogonal QS-CDMA channels outperform a system with random sequences. Vice versa, the same is true if the significant channel impulse response is long and only voice applications have to be supported.

5 Conclusions and Future Work

5.1 Conclusions

The primary objective of this thesis was to develop distributed strategies for resource allocation and interference management in decentralized wireless networks. We considered three main approaches: the utility-based, the utility-based with QoS support and the relative max-min approach. The investigated mechanisms for resource allocation and interference management have been power control, beamforming and link scheduling in time or frequency.

More precisely, in the main part of this thesis, in chapters 2 and 3, we considered a wireless network where interference was treated as noise. In chapter 2 we investigated the so far open problem of how QoS requirements in terms of minimum SIR targets can be incorporated into the traditional utility-based power control problem. A straightforward approach is to maximize the aggregate utility function over the set of feasible power vectors (hard QoS support). Here, an additional challenge in comparison with the traditional approach is to perform the projection operation, which may require a lot of coordination in a distributed environment. Therefore, interesting alternatives are primal-dual algorithms to find stationary points of associated Lagrangian functions. We have shown that a primal-dual algorithm based on the standard Lagrangian function can be efficiently implemented in a distributed manner.

Soft QoS support is of interest since the QoS requirements might be infeasible for some channel states, in which case the above problem with hard QoS support has no solution. To circumvent the feasibility problem and to incorporate possible best-effort users we combined a barrier function with the traditional utility function. In case that the barrier function is chosen suitably and the channel gain matrix is irreducible one can arbitrarily closely approximate a solution to the utility-based power control problem with QoS support. Further, remaining resources are allocated to interested users so as to optimize a given aggregate utility function. In contrast, if the SIR targets are infeasible, they are projected on the feasible SIR region with respect to a weighted maximum norm. Again we have shown that the proposed algorithm is amenable to distributed implementation.

In chapter 3 we proposed a framework for joint power control and receive beamforming in decentralized wireless networks for both typical optimization objectives: maximizing some aggregate utility function of the SIRs and finding the max-min SIR-balancing solution.

Compared to pure power control problems a joint optimization is much more challenging. This chapter is one step forward to a better understanding of the interdependencies. First of all, it is not known for which class of utility functions the problem can be converted into a convex one so that a global solution can be efficiently found in distributed wireless networks. Assuming perfect synchronization we formulated the joint problem as a pure power control problem with implicit optimal linear receivers. However, these implicit optimal receivers result in an algorithm which is not amenable to distributed implementation. A solution was revealed by decoupling the problem in a utility-based power control problem with fixed receivers and a receiver control problem with fixed powers. This algorithm was shown to converge to a stationary point. Simulations suggested, that for a larger concavity the algorithm even converges to a global optimum for a large set of initial points. In addition and as is widely known, in real-world wireless networks noisy measurements and estimations occur. We gave insights especially into practical implementation issues and exemplarily showed the effects of noisy estimations (unbiased and biased) as well as the influence of step size control on the convergence properties.

The second part of chapter 3 considered the possibility of solving the max-min SIR-balancing problem distributedly by optimizing a weighted sum of utilities over the joint space of powers and receivers. The obvious advantage of the utility-based power control approach is that distributed power control schemes are available. Now, in addition to approximating the max-min SIR-balancing solution by appropriately choosing the utility functions, there also exists a connection between both problems. More precisely, the utility-based power vector solves the relative max-min SIR power control problem provided that the weight vector is chosen suitably. We analyzed the connection and proposed a joint power and receiver control algorithm that is conjectured to converge to the optimal relative max-min solution. We emphasized that we would be able to distributedly solve the joint max-min SIR-balancing problem provided that the corresponding weights could be calculated in a distributed fashion. So far calculating the weights requires some control data exchange. On the positive side, recent results mitigated this problem considerably [SKB10, SKG11]. Finally, regarding the second part of chapter 3, we emphasized that the authors of [BS06b] described a similar alternating scheme to solve the joint max-min fair power control and preequalizer problem in the noiseless case for a multiuser downlink channel. They achieved the max-min SIR-balancing solution by minimizing the sum of weighted inverse SIRs. Thus, due to their problem structure the individual steps of the algorithm are different. In fact, the generalization provided by our results has been in two directions. We extended the results of [BS06b] to noisy channels under general power constraints and considered a larger class of utility functions. However, we only provided optimal powers and receive beamformers. For transmit beamformer optimization we described a heuristic approach. This is among other things due to the fact that the theory of uplink-downlink duality is not applicable in case of general power constraints.

While chapters 2 and 3 were concerned with networks with interference, chapter 4 considered a general wireless network with orthogonal channels to separate the users. Orthogonal channels are achieved by applying CDMA or OFDMA (with flat fading channels). As a consequence the problem of joint power control and link scheduling, such that a weighted sum of link rates is maximized, greatly simplifies. On the other hand, the scheduling variables are usually discrete. We established an upper bound of the problem by relaxing these discrete variables. First we showed that the relaxed problem is jointly convex and proposed an iterative algorithm, that maximizes the weighted sum rate by keeping the power and the scheduling vectors alternately constant. The optimal real solution can be simply rounded to result heuristically in an integer solution. However, a more promising way was to analyze the problem structure and then to derive a convenient heuristic based on the problem properties. Interestingly, the discrete problem was shown to be a multiple-choice knapsack problem. We emphasized that the approach of weighted throughput maximization offers throughput-optimality if the weights are chosen accordingly. Moreover, the choice of the weights highly influences the delay performance as well.

Finally, we refer to publications that have not been incorporated in this thesis [5, 12, 6]. All of these works discussed topics in the area of energy-efficient wireless sensor networks.

5.2 Outlook and Future Work

Some of the key success factors of future wireless networks will be the ability to provide cost-effective and affordable wireless bandwidth, better quality of service for innovative services and the ability to integrate sensors. In order to meet these factors the following topics are of interest:

- In order to save infrastructure costs, wireless ad-hoc networks, which are usually decentralized, may be used to flexibly extend wired and cellular wireless networks.
- In many future wireless networks energy-efficient wireless communication is crucial to enhance the lifetime. Furthermore, saving energy fits into the aim of green radio where the carbon emission is intended to be reduced.
- Sophisticated transmission techniques are required to increase the capacity and robustness of wireless links. Here multiple-antenna techniques will play an important role.
- Further research in the area of resource allocation and interference management is necessary to improve the network performance. These developments should incorporate, at most jointly, mechanisms as medium access control in terms of power control and scheduling, as routing and as congestion control.

To stay consistent with the main goal of this thesis, to develop distributed strategies for resource allocation and interference management, one interesting extension might be a utility-per-cost maximization: maximize a weighted sum of utilities of link rates balanced against a cost function that mirrors strict deadlines or limited energy resources. Deterministic MAC protocols should be developed and analyzed, the problem of noisy measurements should be taken into consideration to devise practically applicable algorithms. In addition to deterministic solutions, stochastic protocols are of interest because these cope with fast intractable channels and network dynamics by using the statistics of the network variables. Finally, cross-layer variables from the routing or congestion control layer should be appropriately combined with the MAC scheme to further improve the network performance.

Further open issues that result directly from this thesis include among other things:

- Identify the class of utility functions for which the joint power and receiver control problem is globally solvable.
- Is it possible to develop a fully-distributed power control and receive beamforming algorithm which finds a global solution to the max-min SIR balancing problem?
- Provide a general interference function model and investigate the resource allocation problems assuming a general interference function.
- An interesting theoretical question is: When is it optimal to treat interference as noise?

Last but not least, implementing and verifying these resource allocation algorithms in a simple but real wireless network seems to be crucial for the future of such algorithms and thus provides a significant challenge in future.

Publication list

[1] A. Feistel and S. Stanczak. Dynamic resource allocation in wireless ad-hoc networks based on QS-CDMA. In *Proc. 16th IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Berlin, Germany, Sept. 11-14 2005.

[2] A. Feistel and S. Stanczak. Hop-by-hop congestion control with power control for wireless mesh networks. In *Proc. 65th IEEE Vehicular Technology Conference (VTC-Spring)*, Dublin, Ireland, April 22-25 2007.

[3] A. Feistel, S. Stanczak, and M. Kaliszan. Resource allocation in multiantenna systems under general power constraints. In *Proc. International ITG Workshop on Smart Antennas (WSA)*, Berlin, Germany, Feb. 16–19 2009.

[4] A. Feistel, S. Stanczak, and D. Tomecki. Joint utility-based power control and receive beamforming in decentralized wireless networks. *Special Issue on Wireless Communications and Networking, EURASIP*, June 2010.

[5] A. Feistel, M. Wiczanowski, and S. Stanczak. Optimization of energy consumption in wireless sensor networks. In *Proc. ITG/IEEE International Workshop on Smart Antennas (WSA)*, Vienna, Austria, Feb. 26-27 2007.

[6] I. Koutsopoulos, S. Stanczak, and A. Feistel. Transmit rate control for energy-efficient estimation in wireless sensor networks. In *IEEE Global Communications Conference (GLOBECOM)*, Miami, Florida, USA, Dec. 6-10 2010.

[7] S. Stanczak and A. Feistel. Dynamic resource allocation for frequency-selective orthogonal QS-CDMA systems. In *Proc. 38th Asilomar Conference on Signals, Systems, and Computers*, Monterey, CA, USA, Nov. 7-10 2004.

[8] S. Stanczak, A. Feistel, and H. Boche. QoS support with utility-based power control. In *Proc. 2008 IEEE International Symposium on Information Theory (ISIT)*, Toronto, Canada, July 6–11 2008.

[9] S. Stanczak, A. Feistel, H. Boche, and M. Wiczanowski. Towards efficient and fair resource allocation in wireless networks. In *Proc. 6th Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, Berlin, Germany, March 31–April 4 2008.

[10] S. Stanczak, A. Feistel, and D. Tomecki. On utility-based power control and receive beamforming. In *Proc. 41st Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, USA, March 14-16 2007.

[11] S. Stanczak, A. Feistel, M. Wiczanowski, and H. Boche. Utility-based power control with QoS support. *Wireless Networks, Springer*, 16(6):1691–1705, Aug. 2010.

[12] M. Wiczanowski, A. Feistel, S. Stanczak, and H. Boche. On energy cost of bit and bit/s in multiantenna wireless networks under hardware constraints. In *Proc. ITG/IEEE International Workshop on Smart Antennas (WSA)*, Vienna, Austria, Feb. 26-27 2007.

Bibliography

[AKR$^+$01] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, *Providing quality of service over a shared wireless link*, IEEE Commun. Mag. **39** (2001), no. 2, 150–154.

[ASB10] R. Agrawal, V. G. Subramanian, and R. Berry, *Joint scheduling and resource allocation CDMA systems*, IEEE Trans. Inform. Theory **56** (2010), no. 5, 2416–2432.

[Bam98] N. Bambos, *Toward power-sensitive network architectures in wireless communications: Concepts, issues, and design aspects*, IEEE Person. Commun. Mag. **5** (1998), 50–59.

[BCP00] N. Bambos, S.C. Chen, and G.J. Pottie, *Channel access algorithms with active link protection for wireless communication networks with power control*, IEEE/ACM Trans. Networking **8** (2000), no. 5, 583–597.

[Ber03] D.P. Bertsekas, *Nonlinear programming*, Athena Scientific, Belmont, Massachusetts, 2003.

[BG92] D.P. Bertsekas and R.G. Gallager, *Data networks*, Prentice-Hall, Englewood Cliffs, 1992.

[BO01] M. Bengtsson and B. Ottersten, *Optimal and suboptimal transmit beamforming*, Handbook of Antennas in Wireless Communications (L.C. Godara, ed.), CRC Press, Aug. 2001.

[BS04] H. Boche and S. Stanczak, *Convexity of some feasible QoS regions and asymptotic behavior of the minimum total power in CDMA systems*, IEEE Trans. Commun. **52** (2004), no. 12, 2190–2197.

[BS05] ———, *Log-convexity of the minimum total power in CDMA systems with certain quality-of-service guaranteed*, IEEE Trans. Inform. Theory **51** (2005), no. 1, 374–381.

[BS06a] H. Boche and M. Schubert, *A general theory for SIR-balancing*, EURASIP Journal on Wireless Communications and Networking **2006** (2006), no. 2.

[BS06b] ———, *Resource allocation in multiantenna systems - Achieving max-min fairness by optimizing a sum of inverse SIRs*, IEEE Trans. Signal Processing **54** (2006), no. 6, 1990–1997.

[BSS09] L. Bui, R. Srikant, and A. L. Stolyar, *Novel architectures and algorithms for delay reduction in back-pressure scheduling and routing*, Proc. IEEE INFOCOM Mini-Conference, Rio de Janeiro, Brazil, 2009.

[BT89] D.P. Bertsekas and J.N. Tsitsiklis, *Parallel and distributed computation*, Prentice-Hall, Englewood Cliffs, 1989.

[BV04] S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge University Press, 2004.

[BWS04] H. Boche, M. Wiczanowski, and S. Stanczak, *Characterization of optimal resource allocation in cellular networks*, Proc. 5th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Lisboa, Portugal, July 11-14 2004.

[Chi04] M. Chiang, *To layer or not to layer: Balancing transport and physical layers in wireless multihop networks*, Proc. 23rd IEEE Conference on Computer Communications (INFOCOM), Hong Kong, March 7-11 2004.

[Chi05] ———, *Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control*, IEEE J. Select. Areas Commun. **23** (2005), no. 1, 104–116.

[Chi06] ———, *Geometric programming for communication systems*, Foundation and Trends in Communications and Information Theory **2** (2006), no. 1/2.

[CIM04] D. Catrein, L. Imhof, and R. Mathar, *Power control, capacity, and duality of up- and downlink in cellular CDMA systems*, IEEE Trans. Commun. **52** (2004), no. 10, 1777–1785.

[CTR98] J. Chang, L. Tassiulas, and F. Rashid-Farrokhi, *Joint transmitter receiver diversity for efficient space division multiaccess*, IEEE J. Select. Areas Commun. **16** (1998), no. 8, 1437–1449.

[DV07] D. Das and M.K. Varanasi, *Distributed algorithms for joint optimization of multiuser receivers and power control*, IEEE Trans. Wireless Commun. **6** (2007), no. 9, 3374–3383.

[FH00] P. Fan and L. Hao, *Generalized orthogonal sequences and their applications in synchronous CDMA systems*, IEICE Trans. Fundamentals **E-83** (2000), no. 11, 2054–2069.

[FM93] G.J. Foschini and Z. Miljanic, *A simple distributed autonomous power control algorithm and its convergence*, IEEE Trans. Veh. Technol. **42** (1993), no. 4, 641–646.

[FM05] A. Feiten and R. Mathar, *Optimal power control for multiuser CDMA channels*, Proc. 2005 IEEE International Symposium on Information Theory (ISIT), Adelaide, Australia, Sept. 4-9 2005.

[Fre02] R. Frenkiel, *A brief history of mobile communications*, IEEE Veh. Technol. Society News **49** (2002), no. 2, 4–7.

[GM00] D. Goodman and N. Mandayam, *Power control for wireless data*, IEEE Person. Commun. Mag. **7** (2000), 48–54.

[GMG08] N. Gatsis, A.G. Marques, and G.B. Giannakis, *Utility-based power control for peer-to-peer cognitive radio networks with heterogeneous qos constraints*, Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Las Vegas, Nevada, USA, March 31-April 4 2008.

[HBH05] J. Huang, R. Berry, and M. L. Honig, *A game theoretic analysis of distributed power control for spread spectrum ad hoc networks*, Proc. IEEE International Symposium on Information Theory (ISIT), Adelaide, Australia, Sept. 4–9 2005.

[HBH06] ______, *Distributed interference compensation for wireless networks*, IEEE J. Select. Areas Commun. **24** (2006), no. 5, 1074–1084.

[HMV95] M. Honig, U. Madhow, and S. Verdu, *Blind adaptive multiuser detection*, IEEE Trans. Inform. Theory **41** (1995), no. 4, 944–960.

[HRCW08] P. Hande, S. Rangan, M. Chiang, and X. Wu, *Distributed uplink power control for optimal SIR assignment in cellular data networks*, IEEE/ACM Trans. Networking **16** (2008), no. 6, 1420–1433.

[HSAB09a] J. Huang, V. G. Subramanian, R. Agrawal, and R. Berry, *Downlink scheduling and resource allocation for OFDM systems*, IEEE Trans. Wireless Commun. **8** (2009), no. 1, 288–296.

[HSAB09b] ______, *Joint scheduling and resource allocation in uplink OFDM systems for broadband wireless access networks*, IEEE J. Select. Areas Commun. **27** (2009), no. 2, 226–234.

[HSBA10] J. Huang, V. Subramanian, R. Berry, and R. Agrawal, *Scheduling and resource allocation in ofdma wireless communication systems*, Orthogonal Frequency

Division Multiple Access Fundamentals and Application (T. Jiang, L.Song, and Y. Zhang, eds.), Auerbach Publications, CRC Press, April 2010.

[JXB03] M. Johansson, L. Xiao, and S. Boyd, *Simultaneous routing and resource allocation in cdma wireless data networks*, Proc. IEEE International Conference on Communications, Anchorage, Alaska, May 2003.

[Kel97] F.P. Kelly, *Charging and rate control for elastic traffic*, European Trans. on Telecomm. **8** (1997), 33–37.

[KG05] A. Koskie and Z. Gajic, *A Nash game algorithm for SIR-based power control for 3G wireless CDMA networks*, IEEE/ACM Trans. Networking **13** (2005), no. 5, 1017–1026.

[KL00] D. Kivanc and H. Liu, *Subcarrier allocation and power control for OFDMA, monterey, ca, usa*, Proc. 34th Asilomar Conf. Signals, Systems and Computers, Oct. 29- Nov. 1 2000.

[KLKL01] I. Kim, H.L. Lee, B. Kim, and Y.H. Lee, *On the use of linear programming for dynamic subchannel and bit allocation in multiuser OFDM*, Proc. IEEE Global Telecommunications Conference (GLOBECOM), San Antonio, Texas, USA, Nov. 25–29 2001.

[KMT98] F.P. Kelly, A.K. Maulloo, and D.K.H. Tan, *Rate control for communication networks: Shadow prices, proportional fairness and stability*, J. Oper. Res. Soc. **49** (1998), no. 3, 237–252.

[KV05] K. Kumaran and H. Viswanathan, *Joint power and bandwidth allocation in downlink transmission*, IEEE Trans. Wireless Commun. **4** (2005), no. 3, 1008 – 1016.

[KY03] H.J. Kushner and G.G. Yin, *Stochastic approximation and recursive algorithms and applications*, Springer, Berlin, 2003.

[Lan69] P. Lancaster, *Theory of matrices*, Academic Press, Inc., 1969.

[LG97] C. R. Lin and M. Gerla, *Adaptive clustering for mobile wireless networks*, IEEE J. Select. Areas Commun. **15** (1997), no. 7, 1265–1275.

[LG01] L. Li and A. Goldsmith, *Capacity and optimal resource allocation for fading broadcast channels I. ergodic capacity*, IEEE Trans. Inform. Theory **47** (2001), no. 3, 1083–1102.

[L.H93] L.Hu, *Distributed code assignments for CDMA packet radio networks*, IEEE/ACM Trans. Networking **1** (1993), no. 6, 668–677.

[LHC07] T. Lan, P. Hande, and M. Chiang, *Joint beamforming and power control for optimal sir assignment in cellular uplinks*, Proc. IEEE International Symposium on Information Theory (ISIT), Nice, France, June 2007.

[LPD02] S.H. Low, F. Paganini, and J.C. Doyle, *Internet congestion control*, IEEE Control Systems Mag. (2002).

[LPW02] S.H. Low, L. Peterson, and L. Wang, *Understanding TCP vegas: A duality model*, Journal of the ACM (JACM) **49** (2002), no. 2, 207–235.

[LS06] X. Lin and N.B. Shroff, *The impact of imperfect scheduling on cross-layer congestion control in wireless networks*, IEEE/ACM Trans. Networking **14** (2006), no. 2, 302–315.

[LUE05] J. Luo, S. Ulukus, and A. Ephremides, *Standard and quasi-standard stochastic power control algorithms*, IEEE Trans. Inform. Theory **51** (2005), no. 7, 2612–2624.

[MCWG95] A. Mas-Colell, M. D. Whinston, and J.R. Green, *Microeconomic theory*, Oxford University Press, 1995.

[Mey00] C. D. Meyer, *Matrix analysis and applied linear algebra*, SIAM, Philadelphia, 2000.

[MW00] J. Mo and J. Walrand, *Fair end-to-end window-based congestion control*, IEEE/ACM Trans. Networking **8** (2000), no. 5, 556–567.

[NMR03] M.J. Neely, E. Modiano, and C.E. Rohrs, *Dynamic power allocation and routing for time varying wireless networks*, Proc. 22nd IEEE Conference on Computer Communications (INFOCOM), San Francisco, CA, USA, March 30-April 3 2003.

[OR00] J.M. Ortega and W.C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*, Classics in Applied Mathematics 30, SIAM, Philadelphia, 2000.

[Pis95] D. Pisinger, *A minimal algorithm for the multiple-choice knapsack problem*, European Journal of Operational Research **114** (1995), 528–541.

[RLT02] F. Rashid-Farrokhi, K. J. Liu, and L. Tassiulas, *Transmit beamforming and power control for cellular wireless systems*, IEEE Trans. Wireless Commun. **1** (2002), no. 1, 16–27.

[Roc71] R.T. Rockafellar, *Saddle points and convex analysis*, Differential Games and Related Topics (H.W. Kuhn and G.P. Szego, eds.), Amsterdam, The Netherlands: North-Holland, 1971, pp. 109–127.

[RTL98] F. Rashid-Farrokhi, L. Tassiulas, and K. J. Liu, *Joint optimal power control and beamforming in wireless networks using antenna arrays*, IEEE Trans. Wireless Commun. **46** (1998), no. 11, 1313–1323.

[SB02] S. Stanczak and H. Boche, *On sequence sets for CDMA channels with a small delay spread*, Proc. 36th Asilomar Conference on Signals, Systems, and Computers, Monterey, CA, USA, Nov. 3-6, 2002.

[SB04] M. Schubert and H. Boche, *Solution of the multi-user downlink beamforming problem with individual SINR constraints*, IEEE Trans. Veh. Technol. **53** (2004), no. 1, 18–28.

[SB06] ———, *Qos-based resource allocation and transceiver optimization*, Foundation and Trends in Communications and Information Theory **2** (2006), no. 6.

[SBH01] S. Stanczak, H. Boche, and M. Haardt, *Are LAS codes a miracle?*, Proc. IEEE Global Communications Conference (GLOBECOM), San Antonio, Texas, USA, Nov. 25-29 2001.

[SBH10] C. Shi, R.A. Berry, and M.L. Honig, *Adaptive beamforming in interference networks via bi-directional training*, 44th Annual Conference on Information Sciences and Systems (CISS), Princeton, USA, Mar. 17-19 2010.

[Sen81] E. Seneta, *Non-negative matrices and Markov chains*, Springer, Berlin, 1981.

[SKB10] S. Stanczak, M. Kaliszan, and N. Bambos, *A characterization of max-min SIR-balanced power allocation with applications*, Wireless Networks, Springer **16** (2010), no. 8, 2335–2347.

[SKG11] S. Stanczak, M. Kaliszan, and M. Goldenbaum, *Max-min fair rate control based on a saddle-point characterization of some perron roots*, submitted to the 36th International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 2011.

[SMG02] C.U. Saraydar, N.B. Mandayam, and D.J. Goodman, *Efficient power control via pricing in wireless data networks*, IEEE Trans. Commun. **50** (2002), no. 2, 291–303.

[SS01] S. Shakkottai and A.L. Stolyar, *Scheduling algorithms for a mixture of real-time and non-real-time data in HDR*, Proc. of the 17th International Teletraffic Congress (ITC-17), Salvador da Bahia, Brazil, Sept. 2001.

[ST05] S. Sarkar and L. Tassiulas, *End-to-end bandwidth guarantees through fair local spectrum share in wireless ad hoc networks*, IEEE Trans. Automat. Contr. **50** (2005), no. 9, 1246–1259.

[SW05] S. Stanczak and M. Wiczanowski, *Distributed fair power control for wireless networks: Objectives and algorithms*, Proc. 43rd Annual Allerton Conference on Communications, Control, and Computing, Sept. 28-30 2005, *invited paper.*

[SWB07] S. Stanczak, M. Wiczanowski, and H. Boche, *Distributed utility-based power control: Objectives and algorithms*, IEEE Trans. Signal Processing **55** (2007), no. 10, 5058–5068.

[SWB09] ______, *Fundamentals of resource allocation in wireless networks*, vol. 3, Springer, Berlin, 2009.

[TE92a] L. Tassiulas and A. Ephremides, *Jointly optimal routing and scheduling in packet radio networks*, IEEE Trans. Inform. Theory **38** (1992), no. 1, 165–168.

[TE92b] ______, *Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks*, IEEE Trans. Automat. Contr. **37** (1992), no. 12, 1936–1948.

[TH99] D. Tse and S. Hanly, *Linear multiuser receivers: Effective interference, effective bandwidth and user capacity*, IEEE Trans. Inform. Theory **45** (1999), no. 2, 641–657.

[Tse97] D.N. Tse, *Optimal power allocation over parallel gaussian broadcast channels*, Proc. International Symposium on Information Theory (ISIT), Ulm, Germany, Jun. 29 - Jul. 4 1997.

[UY98] S. Ulukus and R.D. Yates, *Stochastic power control for cellular radio systems*, IEEE Trans. Commun. **46** (1998), no. 6, 784–798.

[Var99] M.K. Varanasi, *Nonlinear multiuser receivers with distributed power control in cellular radio networks*, Proc. 37th Annual Allerton Conference on Communication, Control and Computing, IL, USA, Sept. 22-24 1999.

[VAT99] P. Viswanath, V. Anantharam, and D.N.C. Tse, *Optimal sequences, power control, and user capacity of synchronous CDMA systems with linear MMSE multiuser receivers*, IEEE Trans. Inform. Theory **45** (1999), no. 6, 1968–1983.

[Ver98] S. Verdu, *Multiuser detection*, Cambridge University Press, 1998.

[WCLM99] C.Y. Wong, R.S. Cheng, K.B. Lataief, and R.D. Murch, *Multiuser OFDM with adaptive subcarrier, bit and power allocation*, IEEE J. Select. Areas Commun. **17** (1999), no. 10, 1747 –1758.

[WES06] A. Wiesel, Y.C. Eldar, and S. Shamai, *Linear precoding via conic optimization for fixed MIMO receivers*, IEEE Trans. Signal Processing **54** (2006), no. 1, 161–176.

[WSB08] M. Wiczanowski, S. Stanczak, and H. Boche, *Providving quadratic convergence of decentralized power control in wireless networks – The method of min-max functions*, IEEE Trans. Signal Processing **56** (2008), no. 8, 4053–4068.

[Wu00] Qiang Wu, *Optimum transmitter power control in cellular systems with heterogeneous SIR thresholds*, IEEE Trans. Veh. Technol. **49** (2000), no. 4, 1424–1429.

[XSC03] M. Xiao, N.B. Schroff, and E.K.P. Chong, *A utility-based power control scheme in wireless cellular systems*, IEEE/ACM Trans. Networking **11** (2003), no. 2, 210–221.

[Yat95] R.D. Yates, *A framework for uplink power control in cellular radio systems*, IEEE J. Select. Areas Commun. **13** (1995), no. 7, 1341–1347.

[YC03] E. Yeh and A. Cohen, *Maximum throughput and minimum delay in fading multiaccess communications*, Proc. IEEE International Symposium on Information Theory (ISIT), Yokohama, Japan, 2003.

[YS04] Y. Yi and S. Shakkottai, *Hop-by-hop congestion control over a wireless multihop network*, Proc. 23rd IEEE Conference on Computer Communications (INFOCOM), Hong Kong, March 7-11 2004.

[YSR09] L. Ying, S. Shakkottai, and A. Reddy, *On combining shortest-path and backpressure routing over multihop wireless networks*, Proc. IEEE Conference on Computer Communications (INFOCOM), Rio de Janeiro, Brazil, April 20-25 2009.

[YST08] L. Ying, R. Srikant, and D. Towsley, *Cluster-based back-pressure routing algorithm*, Proc. IEEE Conference on Computer Communications (INFOCOM), Phoenix, USA, April 14-17 2008.

[Zan92a] J. Zander, *Distributed cochannel interference control in cellular radio systems*, IEEE Trans. Veh. Technol. **41** (1992), 305–311.

[Zan92b] ———, *Performance of optimum transmitter power control in cellular radio systems*, IEEE Trans. Veh. Technol. **41** (1992), no. 1, 57–62.

[Zan97] ———, *Radio resource management in future wireless networks: requirements and limitations*, IEEE Commun. Mag. **35** (1997), no. 8, 30–36.

[ZK01] J. Zander and S.-L. Kim, *Radio resource management for wireless networks*, Artech House, Boston, London, 2001.