OPTIMAL RESOURCE ALLOCATION ALGORITHMS FOR EFFICIENT OPERATION OF WIRELESS NETWORKS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ÖMÜR ÖZEL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

JULY 2009

Approval of the thesis:

OPTIMAL RESOURCE ALLOCATION ALGORITHMS FOR EFFICIENT OPERATION OF WIRELESS NETWORKS

submitted by **ÖMÜR ÖZEL** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen ________________
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. İsmet Erkmen ________________
Head of Department, **Electrical and Electronics Engineering**

Assoc. Prof. Dr. Elif Uysal Bıyıkoğlu ________________
Supervisor, **Electrical and Electronics Engineering, METU**

Examining Committee Members:

Prof. Dr. Yalçın Tanık
Electrical and Electronics Engineering, METU ________________

Assoc. Prof. Dr. Elif Uysal-Bıyıkoğlu
Electrical and Electronics Engineering, METU ________________

Prof. Dr. Mustafa Kuzuoğlu
Electrical and Electronics Engineering, METU ________________

Assist. Prof. Dr. Cüneyt F. Bazlamaçcı
Electrical and Electronics Engineering, METU ________________

Assist. Prof. Dr. Tolga Girici
Electrical and Electronics Engineering, TOBB ETU ________________

Date: ________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: ÖMÜR ÖZEL

Signature :

ABSTRACT

OPTIMAL RESOURCE ALLOCATION ALGORITHMS FOR EFFICIENT OPERATION OF WIRELESS NETWORKS

Özel, Ömür
M.S., Department of Electrical and Electronics Engineering
Supervisor : Assoc. Prof. Dr. Elif Uysal Bıyıkoğlu

July 2009, 105 pages

In this thesis, we analyze allocation of two separate resources in wireless networks: transmit power and buffer space. Controlled allocation of power can provide good performance for both users and the network. Although centralized mechanisms are possible, distributed power control algorithms are preferable for efficient operation of the network. Viewing distributed power allocation as the collection of rational decisions of each user, we make game theoretic problem formulations, devise distributed algorithms and analyze them. First, equilibrium analysis of a vector power control game based on network energy efficiency in a multiple access point wireless network is presented. Then, a distributed mechanism is proposed that can smooth admission control type power control so that every user can stay in the system. Introducing a new externality into utility function, a game theoretic formulation that results in desired distributed actions is made. Next, the proposed externality is investigated in a control theoretic framework. Convergence of gradient based iterative power updates are investigated and stability of corresponding continuous time dynamical system is established. In the final part of the thesis, allocation of buffer space is addressed in a wireless downlink using a queueing theoretic framework. An efficient algorithm that finds optimal buffer partitioning is proposed and applications of the algorithm for different scenarios are illustrated. Implications of the results about cross layer design and multiuser diversity are discussed.

Keywords: Distributed Power Control, Game Theory, Utility Function, Buffer Management, Resource Allocation

ÖZ

KABLOSUZ AĞLARDA VERİMLİ İŞLEYİŞ İÇİN OPTİMAL KAYNAK AYIRMA ALGORİTMALARI

Özel, Ömür
Yüksek Lisans, Elektrik-Elektronik Mühendisliği Bölümü
Tez Yöneticisi : Doç. Dr. Elif Uysal-Bıyıkoğlu

Ţemmuz 2009, 105 sayfa

Bu tez çalışmasında, kablosuz ağlar için iki farklı kaynağın kullanıcılara atanması analiz edilmektedir: veri iletim gücü ve hafıza. Güç kontrolü hem kullanıcılar hem de ağın başarımı için önemlidir. Güç atama merkezi bir şekilde yapılabilse de ağın verimli işleyişi açısından güç atamasının dağıtık yapılması tercih edilmektedir. Dağıtık güç kontrolü kullanıcıların kendi kendilerine aldıkları kararlar bütünü olarak yorumlanmıştır. Oyun teorisinin kavramları ile problemler formule edilmiş, dağıtık algoritmalar geliştirilmiş ve analizler yapılmıştır. İlk olarak, ağın enerji verimliliğine dayanan çok erişim noktalı bir kablosuz ağdaki güç kontrolü oyununun denge analizi sunulmuştur. Daha sonra, kullanıcıların sistemde var olup olmadığına dayanan kontol yöntemlerine alternatif olabilecek yumuşak geçişlere dayanan dağıtık bir mekanizma önerilmiştir. Kazanç fonksiyonu üzerinde yeni bir modifikasyon önerilmiş, istenilen dağıtık kontrolü elde eden oyun formülasyonu yapılmıştır. Bir sonraki bölümde, önerilen yeni modifikasyon kontrol teoretik bir yaklaşımla ele alınmıştır. Gradyan temelli yinelemeli algoritmaların yakınsaması incelenmiş, bu yinelemelere ait dinamik sistemin stabilite analizi yapılmıştır. Son olarak, bir kablosuz merkez uç birim sisteminde kuyruk teorisi kullanılarak veri yastığı ataması problemi çalışılmıştır. Toplam veri yastığını optimal olarak bölmelere ayıran verimli bir algoritma önerilmiş ve algoritmanın uygulaması çeşitli senaryolar üzerinde örneklerle gösterilmiştir. Elde edilen sonuçların katmanlararası tasarım ve çok kullanıcı çeşitliliği konusunda içerdiği anlamlar tartışılmıştır.

Anahtar Kelimeler: Dağıtık Güç Kontrolü, Oyun Teorisi, Kazanç Fonksiyonu, Hafıza Yönetimi, Öz Kaynak Tahsisi

To my family

ACKNOWLEDGMENTS

Coming very close to the end of a six year long period in Middle East Technical University Department of Electrical and Electronics Engineering, I am quite sentimental about the past six years at METU. There is no doubt that the people I have known here are the ones who make these six years valuable and it is my debt and pleasure to thank them via this thesis work.

First of all, I acknowledge my supervisor Prof. Elif Uysal-Biyikoglu. I am profoundly grateful to her as this thesis would not have been completed without her guidance, assistance, patience and tolerance. From the beginning of our joint work, she allowed me to make research freely on topics of my personal interests and this, along with her friendly approach, made the two year work more instructive and enjoyable than it would be otherwise.

In the last few months of my thesis work, Prof. Tolga Girici from TOBB ETU started to collaborate with us. I am sincerely thankful to Prof. Girici for his interest and support on my thesis work. I also thank Dr. H. Emre Kankaya from Bilkent University for his comments and critics about the work in the thesis.

I appreciate faculty members of the department who contributed my education. In particular, I am indebted to Prof. Mustafa Kuzuoğlu, Prof. Ali Özgür Yılmaz, Prof. Yalçın Tanık and Prof. Mübeccel Demirekler for their teaching, advices and sometimes tolerance.

I thank my friends with whom I made projects, shared knowledge and spent time together at METU. Being an outstanding and a modest person, Tuğcan Aktaş has been a role model for me. I am quite lucky to be a friend of him at the same time. He has been interested and helpful in the success of my work. I owe him a lot. In addition to Tuğcan, my colleagues Gökhan Güvensen and Alper Bereketli were close friends, either in the department or in the dormitory. I appreciate their earnest friendship.

I also thank people from the research group such as Neyre Tekbıyık, Sinan Akyürek and Akif Antepli. It was also very nice to have known Baran Öztan, Özgür Ergül and Cüneyt Baykal.

My good old friends Demet and Çiğdem from hometwon Adana has appeared in the last two months of the thesis work. I appreciate their keen interest on my success and their sincere friendship.

I acknowlegde sponsorship of TÜBİTAK under Kariyer grant 106E119.

Last but not the least, through the completion of this thesis I present my special thanks to my mother, father and brother, Nimet, Yaşar and Özgür for their endless support and trust on me. The thesis is dedicated to them.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

Due to channel variations and interference of nodes on each other, wise allocation and control of resources is important in wireless networks. Moreover, scalability and cost considerations drive the interest in finding efficient algorithms for networks. In this thesis, we predominantly study allocation of transmit power and buffer space.

In wireless networks, interference induced interaction make a user's performance dependent on other users' actions. Hence, coordination and control is necessary in wireless networks for acceptable level of communication with minimal cost for users sharing the network. Transmit power, one of possible actions of users, is an important determinant of user performance in the network. On one side, transmit power is supposed to be high for a good level of communication; on the other side, it should not be too high in order not to waste limited energy resource and create extra load (interference) on other members of the network. In this context, power controlled data transmission can optimize both network and user performance: unnecessary interference is avoided and energy consumption is made more efficient.

Transmit power control can be performed using either distributed or centralized mechanisms. Distributed power control is more preferable than its centralized counterpart as it does not require central processing for the network and communication overhead to disseminate the control data. In a distributed control mechanism, users use their own observations, make their own processing and decide on their own. Hence, it is natural to use a game model for distributed transmit power control: users are players and the transmit powers are strategies.

Game theoretic models for power control has recently attracted considerable interest [2,3,11,12,24,38–40,53,54,60,70]. Analysis using tools of game theory gives insight to understand why and how the coordination is required in the network and provides methods to generate distributed algorithms [6, 15]. Selfish and rational players try to maximize their own utility and this way distributed actions are obtained. Depending on the definition of the utility, this utility based power control may drive the network to an inefficient equilibrium [53]. To avoid inefficiency, utility functions of users can be introduced externalities such as prices, that act as a punishment mechanism [3,53,70]. In power control context, by means of externalities, users are encouraged to transmit with less power than they would like to do for obtaining full satisfaction from the network, which may sometimes lead to harsh decreases and hence to admission control as in [3, 53, 70]. Being a way of cooperation in distributed power control, these externalities introduced to the utility functions may sometimes make users discarded from the system. However, smooth distributed power control mechanisms, unlike admission control, may be more natural and useful [42, 43]. We addressed smoothness of distributed actions in a game theoretic framework.

Another resource that is crucial in the performance of a wireless network is buffer [9, 13, 36, 61]. In a finite buffer system shared by multiple sessions, the question of how much data should be buffered from each session becomes important for throughput performance. Moreover, if the link is wireless, storing certain amount of data affects multiuser diversity gain, which has not yet been addressed in the literature. Novel problem formulations are possible if buffer allocation is considered jointly with channel assignment and/or scheduling. Cross layer approaches gained popularity [17]. In particular, joint link and physical layer operations for user scheduling in a wireless downlink constitute an important portion of recent research trends [30, 59, 73, 74]. However, cross layer approaches come with a high complexity cost. When finite buffer constraint is also taken into formulation, the optimal joint strategy is not known in the literature and it is possible that separation of layers performs better. In this context, benefits of cross layer approaches are questioned and alternative algorithms are developed in this thesis.

1.1 Outline of the Thesis

In Chapter 2, we present a game theoretic study of a wireless network with multiple access points, where users can control their uplink transmit power targeted to any or all of the access points. Distributed mechanisms that achieve a tradeoff between energy efficiency and spectral efficiency in the network are obtained through the use of suitably designed utility functions. A user's utility is a function of throughput and average transmission power. A key assumption in the chapter is that throughput is modeled as a sigmoidal function of SINR (signal to interference-plus-noise ratio). Each user, being selfish and rational, acts to maximize its utility in response to instantaneous SINR by adjusting its power. The resulting mechanism is a distributed power control scheme that can incline towards energy-efficient or spectrally efficient operating points depending on the choice of utility function. Existence and uniqueness of Nash Equilibrium [15] (NE) points in this game are shown via convergence of the distributed power iterations. The best-response strategy is shown to converge to the strategy of each user selecting a single access point, and the corresponding power levels used for various priorities between energy and spectral efficiency are characterized.

In Chapter 3, we study distributed power control in an interference network. We define quality of service (QoS) objectives of users based on fading induced outage probabilities. Starting from the objectives, we construct a utility function. Feasibility and optimality of the power allocation in equilibrium of resulting distributed mechanism are investigated using a game theoretic approach. Then, we modify the utility function such that users smoothly decrease their objectives with increasing interference in the form of non-monotonic reaction curves. This modified game has a unique NE and resultant modified power control algorithm will be called non-monotonic power control (NMPC). As users self-select to reduce power smoothly, NMPC is power-efficient when their own SIR objectives are made unfeasible by the channel gains available. This algorithm leads to a smooth transfer of resources from disadvantaged user to the advantaged in terms of channel state. By means of a numerical example, NMPC is shown to increase the number of users that achieve the objective without removing any user. Besides, it considerably decreases total power consumption.

In Chapter 4, we extend the approach in Chapter 3 to a control theoretic setting.

When rate expectations of users in a wireless network cannot all be continuously satisfied, one choice is to discard some users from the system, in a mechanism usually called admission control. However, in a data network, users have a certain tolerance to occasional rate outages. We argue that reducing users' rate objectives smoothly, by considering an outage probability tolerance, may be more preferable to the users than not being provided any service at all. We propose a distributed utility based algorithm for doing this. The smoothness of reactions is maintained by transmit power reaction curves to be absolutely subhomogeneous as a function of interference. This is done through introducing an "objective reduction factor", in addition to a linear price. We first provide conditions for a unique NE. Assuming that distributed nodes use gradient based optimization, convergence and error sensitivity of gradient based iterative algorithms are analyzed [2]. Lastly, the continuous-time counterpart of the problem is considered and a stability condition is established for the system [28].

Finally, in Chapter 5, we analyze throughput performance of a wireless downlink using a queueing theoretic framework. In particular, we consider a finite buffer shared by multiple packet queues. Using results about buffer sharing under Poisson arrival and service in previous literature [13, 23, 25], we observe that throughput can be improved by partitioning the buffer space among the queues judiciously, especially under unbalanced and high load regime. We formulate optimal buffer partitioning as a resource allocation problem, the solution of which is found through a greedy incremental algorithm in polynomial time. The rest of the work is devoted to applying the optimal buffer allocation strategy in different scenarios modeling a wireless downlink. First, the strategy is applied in a general parallel $M/M/1/m_i$ system and a numerical study verifies that the strategy may boost the throughput considerably. Then, the buffer allocation problem to queues belonging to users with different arrival rates is tackled jointly with assignment of users to the set of available channels(bands). The joint optimization problem is solved by a best channel highest arrival rate policy under certain assumptions along with On/Off channel model. Lastly, buffer allocation is considered for scheduled service. By means of numerical studies, separate buffer allocation and user scheduling is shown to operate close in throughput to joint queue aware policies.

We conclude the thesis in Chapter 6 by emphasizing key points addressed and possible extensions for future work.

CHAPTER 2

DISTRIBUTED CONTROL OF TRANSMIT POWER FOR ENERGY EFFICIENCY IN WIRELESS NETWORKS

2.1 INTRODUCTION

The material in this chapter partially appears in Proceedings of IEEE 17th Signal Processing and Communications Applications Conference [44]. A fuller version of this chapter is under review [45]. The work in this chapter has been performed as a part of the project [66].

Due to the broadcast nature of wireless communication, and the interference this tends to cause, the performance of a user in a wireless network can be highly dependent on other users' actions. One of the possible actions is the choice of transmission power. In order to achieve a certain rate, for example, a user may need to increase its transmit power as interference level increases. This in turn, can increase the interference on others, who respond, and so on. This interaction may culminate at a stable operating point where every user is satisfied with its own level of signal-to-interference ratio. However, this operating point may not always be energy-efficient. There is a well known tradeoff between energy efficiency and transmission rate [65, 67], and in the context of a wireless network, increasing transmission power only makes sense if, considering the network's response, it will ultimately lead to an appreciable gain in rate.

In principle, a network can be engineered to converge to a desired operating point by using a suitable power control algorithm. Of course, for implementability, distributed algorithms are attractive. Distributed power control, where wireless nodes make their own power control decisions (possibly asynchronously) has been the focus of a large body of previous studies [14, 19, 22, 70, 72, 75].

In recent years, game theory [15] has been used to model the interference-induced interaction in wireless communication, and to obtain distributed algorithms [3, 12, 39]. In fact, communication networks form an increasingly popular setting for the application of game theory [43]: For one thing, the terminals (nodes) are quite truly rational, and usually selfish players.

A power control game arises when users are able to adjust their power in response to the interference they cause on each other's receiver, with the goal of maximizing the *utility* of their communication with their intended receiver. Depending on the constraints and the utility function, there may be an equilibrium or several equilibria in this game. The network designer's problem is to design the utility function to give rise to a desirable equilibrium from the perspective of the whole network. One particular goal for the network designer is to maximize network energy efficiency, that is, the total power used per overall throughput in the network.

One approach through which users can be driven to be efficient while also trying to maximize their rate has been *pricing*. In this case, users try to maximize a net utility, which is utility minus a price, where price is a function of, say, power. However, pricing is not always natural in settings where there is no center to collect them or if the center is not also a player in the game. In this chapter, we set up a utility model that does not have an explicit price in it in order to especially address ad-hoc or sensor networks where network-wide energy-efficiency is important, yet prices would be a bit out of place.

The single base-station uplink power control game has been addressed and solved in the literature for various utility functions [53, 57]. The main contribution of this chapter is to solve the multiple base-station (access-point) uplink power allocation problem. The best-response strategies (reaction curves) are found, and the existence and uniqueness of Nash equilibrium is shown. Iterative methods to reach the equilibrium are presented. Secondarily, the behavior of equilibria under utility functions with varying degrees of energy-efficiency or throughput-maximization priorities are investigated.

2.1.1 Related Work

In [53], power control games in a single cell system are considered with a utility function in the form of rate/power. A "socially optimum" operating point is derived and prices are introduced to obtain a point closer to the social optimum. In [3], CDMA power control is modeled as a non-cooperative game with utility proportional to rate. Linear prices may result in admission control, since users may opt out of the network as they try to unilaterally optimize net utilities. In [39, 40], power control games in a CDMA system are established with an energy-efficiency goal. With different types of receivers, adaptive modulation and coding, hybrid games are obtained and equilibrium points are analyzed. Analysis of non-cooperative power control in a single cell multi-carrier CDMA system is presented in [38]. The multiple base-station problem addressed in this chapter, while carrying similarities to the multi-carrier CDMA problem, does not reduce to it, as users that select different base stations still potentially cause interference on each other (whereas users selecting different carriers do not).

The rest of the chapter is organized as follows: the next section describes the system model and some definitions. In §2.3, the expression and properties of the utility function are given. The single access point system is analyzed in §2.4. In §2.5 the multiple base-station vector power control game is analyzed. Considering different priorities for different applications, the trade-off between energy and spectral efficiency is considered in § 2.6. Conclusions are presented in §2.7.

2.2 SYSTEM MODEL

We consider a wireless network of K users and M access points (Fig. 2.1). Users can transmit with a rate up to R bps in a common frequency band B Hz. Let the channel gain between user i and access point b be h_{ib}. Channel gains are assumed constant during operation. More generally, h_{ib} can be considered as average channel gain in a fading environment.

Messages are sent from nodes to access points and each access point hears each user's message. Let the message signal of user k to destination access point b be X_{kb}. The signal Y_b, received at base station b is:

$$Y_b = \sum_{l=1}^{K} \sum_{j=1}^{M} \sqrt{h_{lb}} X_{lj} + Z_b \tag{2.1}$$

Z_b is additive noise at access point b. For convenience, we model Z_b as white Gaussian with zero mean and $E|Z_b|^2 = \sigma^2$. P_{ij} is the average power of the message X_{ij}: $E|X_{ij}|^2 = P_{ij}$. Each user is subject to a power constraint:

$$\sum_{j=1}^{M} P_{ij} \leq P_{max} \ \forall i$$

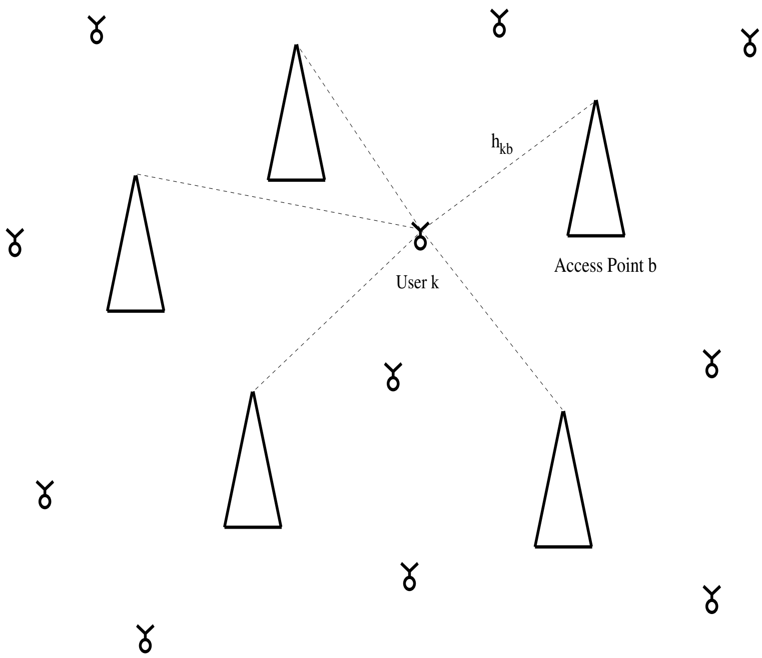

Figure 2.1: Wireless Network with Several Users and Access Points

We consider *single user decoders* in the receivers. Depending on the receiver structure, cross channel gains may be suppressed by additional processing gains. Typical application of this model is DS/CDMA [62] with specific spreading codes for each possible link. As user i sends different data to different users, user i's own message to access points other than b are also treated as interference at access point b. Interference will be treated as noise, as it is usually done in practical receivers, and will be modeled as Gaussian, which can be a good assumption as the number of independent interferers grows. G_p stands for processing gain. Signal to interference plus noise ratio (SINR) of user k in the receiver of access point b is:

$$\gamma_{kb} = G_p \frac{h_{kb} P_{kb}}{\sigma^2 + (\sum_{i=1}^{K} \sum_{j=1}^{M} h_{ib} P_{ij}) - h_{kb} P_{kb}} \tag{2.2}$$

The model with multiple access points has been motivated by a number of communication scenarios: (i) a local area network where wireless nodes may be in the range of multiple access points, (ii) an ad-hoc wireless network with multiple gateway stations that enable connection to a larger, wired network, (iii) the *microdiversity* system in Hanly [21] where multiple access points are considered as a single access point having multiple antennae distributed in space. The understanding in these scenarios is that the connections between access points are wired and the communication among them is straightforward.

In each of these settings, the signaling and coding can take different forms. In addition to the immediate example of a CDMA system given in the previous paragraph, other relevant models include multicarrier signaling and time division: In an OFDM multiplexing strategy, users can allocate different subcarriers to different access points, and divide their total instantaneous power among the subcarriers. In this case, the structure of the problem is somewhat different than the single-carrier version in that users are only subject to interference from users on the same subcarrier. Similarly, users could allocate different time slots to access to different base stations, allocating a long term average power constraint between time slots. Again, interference is between subsets of users using the same time-slot. While characterizing the equilibrium points are more involved in the multicarrier and multislot models, some of the results in this chapter continue to hold, as will be argued later in the chapter.

Note that the effects of strategies of the other users are observed in the denominator of the SINR expression in Eq. (2.2). Hence, users are in such an interaction that performance of one user is degraded when another user attempts to increase its power. This interaction is observed not only in single user decoders but also in multiuser detectors such as MMSE [40] and MMSE SIC [8]. In order to analyze this interaction among users, we will employ static noncooperative game theory.

2.2.1 Static (One-Shot) Games and Equilibrium [15]

A static game is denoted as $\Omega = [U, \{S_i\}, \{u_i\}]$. Three components to define a static game [15] are:

1. User set U
2. Action or strategy set S_i, $\forall i \in U$
3. Utility u_i as a function of elements of S_i, $\forall i \in U$

The user set is the index set of players: $U = \{1, 2, ..., K\}$. Given the other users' actions, users unilaterally maximize their utility in their strategy set. The notion of equilibrium that captures the non-cooperative nature of the problem is called *Nash equilibrium.* An operating point such that no user can achieve higher utility by unilateral changes in action is *Nash equilibirum.*

Definition 2.2.1 *A Nash equilibrium (NE) is the vector of strategies* $\vec{s^*} = [s_1^*, s_2^*, s_3^*, \ldots, s_K^*]$ *such that*

$$u_i(s_i^*, \vec{s^*}_{-i}) \geq u_i(s_i, \vec{s}_{-i}) \ \forall s_i \in S_i$$

is satisfied for all user i *where* $\vec{s}_{-i} = (s_1, s_2, .., s_{i-1}, s_{i+1}, .., s_K)$.

Note that a NE is not necessarily pareto-optimal, i.e. there may be a point with utilities u_i' that is feasible and yet $u_i' > u_i^* \ \forall i$ where u_i^* is the value of user i's utility at NE. Actually it is possible to obtain higher total utility using a cooperative mechanism such as pricing [53, 70]. Since (selfish) users are not interested in overall performance of the network, each user optimizes its *own* utility in its *own* action space.

Given actions of users other than k, $\vec{s}_{-k}$, best response (in other words, the *reaction curve*) of user k, r_k, is:

$$r_k(\vec{s}_{-k}) \triangleq \arg\max_{s_k \in S_k} u_k$$

NE can also be defined in terms of best responses. $\vec{s^*}$ is NE iff $s_k^* = r_k(\vec{s^*}_{-k}) \ \forall k$. In other words, NE is a fixed point of best responses. Consequently, the concept of NE is well-suited to the wireless network power control problem, and we will analyze stable operating points through examining the existence and properties of NE.

2.3 UTILITY FUNCTION

In game theoretic terms, utility function u_i is a mapping from the Cartesian product of action sets of users to real numbers, $u_i : \prod_{j=1}^{K} A_j \rightarrow \mathbf{R}$. The value of the function u_i represents the level of satisfaction of user i with respect to some goal. Usually, in a communication scenario, satisfaction of a node is related to the communication performance such as throughput, outage probability, BER, SINR and power or energy cost. The choice of utility can also depend on external conditions: when spectral resources are scarce, throughput carries high utility, whereas if energy is limited, a utility that decreases with transmit power is appropriate. However, a combination of these parameters must determine the level of satisfaction for mobile data users. *Bits successfully sent per joule of energy spent* has been a well known utility function [40, 53, 54] that appropriately combines throughput and cost terms, encouraging energy-efficient behavior.

The standard definition of throughput is the long term average data rate (bits per transmission) achieved. Taking into account link-layer framing and error control mechanisms whereby a data packet (say, a constant number of bits) is declared unsuccessful if more than a certain number of bit errors occur and resulting packet drops which happen with finite probability, throughput by definition is upper bounded by the long term average coding rate, R. In previous literature, throughput was often modeled as a sigmoidal function of SIR, or SINR (see Fig. 2.2).) [34]. The main reason for this is, as a certain threshold in SNR (SINR) γ is exceeded, packet success probability quickly rises toward 1 with many practical as well as optimal modulation and coding schemes. As a very simple example for the occurrence of the sigmoid, consider the following: packets of length L symbols, are sent using BPSK, and the code rate is R bits per symbol. Each bit is decided erroneously with probability $BER(\gamma)$. Then, the long term average throughput T is:

$$T = R(1 - BER(\gamma))^L \tag{2.3}$$

When $BER(\gamma)$ is decreasing with concave and convex regions, as it is typically the case, T is a sigmoid (e.g., [34]), that is, there is an inflection point $\overline{\gamma}$ such that $T(\gamma)$ is convex in $[0, \overline{\gamma})$ and concave in $(\overline{\gamma}, \infty)$. Note that Eq. (2.3) is in the form of an effective rate, that is, rate multiplied by an *efficiency function* (packet success rate)

$f(.)$.

$$T = Rf(\gamma) \tag{2.4}$$

The sigmoidal assumption for $f(\gamma)$ is valid in many communication scenarios. In a system with fixed coding and modulation, such as an ARQ scheme with CRC check [27], $f(\gamma)$ has sigmoidal shape. Even if messages heard by each access point were decoded in a common center, sigmoidal assumption still holds for a fixed coding scheme. Moreover, in case modulation and rate are adapted to changes in SINR, the sigmoidal shape still applies [39]. Consistent with observations from many theoretical and practical research results [8, 27, 38, 40, 52], $f(\gamma)$ will be assumed to have sigmoidal shape in this chapter.

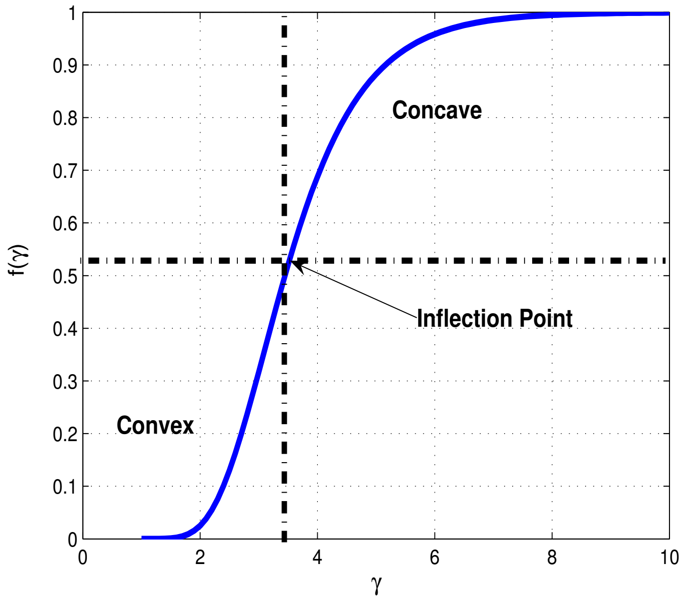

Figure 2.2: Function $f(\gamma)$ v.s. γ in normal scale. It is plotted for BPSK modulation with packet length $L = 400$ bits.

Let T_{kb} and p_{kb} be the throughput and power of user k for communication with access point b respectively. The utility function u_k is defined as the ratio of total goodput to total dissipated power:

$$u_k = \frac{\sum_{b=1}^{M} T_{kb}}{\sum_{b=1}^{M} p_{kb}} \tag{2.5}$$

Note that the motivation for having the power term in the denominator of the utility function is to encourage energy-efficient behavior of users.

To avoid associating a positive utility with no transmission, we must have $u_k \to 0$ when $\sum_{b=1}^{M} p_{kb} = 0$ for all k. Because, the utility in no transmission case must be zero. In order to satisfy this condition, we assume that the following two hold:

$$\lim_{\gamma \to 0} f(\gamma) = 0 \ , \ \lim_{\gamma \to 0} f'(\gamma) = 0$$

Before approaching the general problem, we will first tackle a simpler case. To that end, in the next section, we formulate and analyze the power control game in a single access point system.

2.4 POWER CONTROL GAME IN A SINGLE ACCESS POINT SYSTEM

The notation of this section is obtained by letting $M = 1$ in the general setup. The strategy set of each user i is $S_{1i} = [0, P_{max}]$, where P_{max} is the maximum power level allowed for each user. Utility function of user k is

$$u_{1k} = \frac{T_k}{p_k} \tag{2.6}$$

T_k is the long term average rate as in Eq. (2.3). Let $\Omega_1 = [U, \{S_{1i}\}, \{u_{1i}\}]$ be the one-shot game in which each user unilaterally performs the following optimization

$$\max_{p_k \in S_{1k}} u_{1k}(p_k, P_{-k}) \ \textit{for all}\ k \in \ U$$

An important property possessed by the utility functions u_i that plays a key role in the existence and uniqueness of equilibrium is quasiconcavity.

Definition 2.4.1 *A function $f : \Re \to \Re$ is quasiconcave if $\exists x_0$ such that f is non-decreasing in $x < x_0$ and f is non-increasing in $x > x_0$*

It is observed and can be verified that given P_{-k}, $u_i(p_k, P_{-k})$ $\forall i$ are quasiconcave with respect to p_k (Fig. 2.3).

Theorem 2.4.2 Ω_1 *has a Nash Equilibrium.*

Proof: The result follows from compactness and convexity of strategy set S_i and quasiconcavity of utility functions u_i of users (see Theorem 1 in [53]). ■

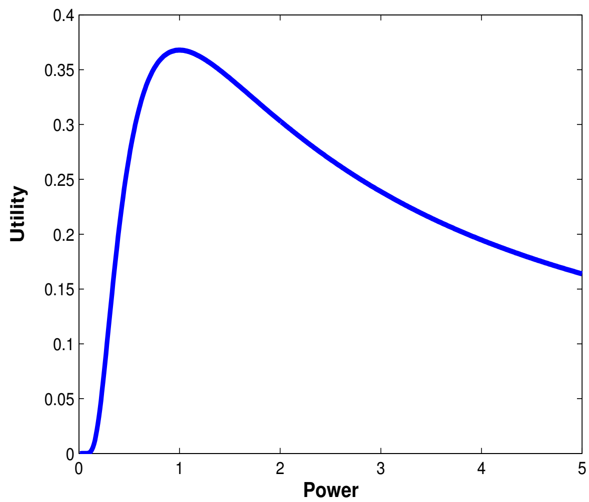

Figure 2.3: Typical variation of utility function u_k with power p_k given other users' powers. It is quasiconcave: monotone increasing up to some value of power, and monotone decreasing afterward. Note that, depending on P_{max}, the decreasing regime may not be observed, but this does not violate quasiconcavity.

Given p_j $j \neq k$, γ_k changes linearly with p_k. Letting $\tilde{h}_k$ be *effective channel gain* of user k, SINR expression in Eq. (2.2) is:

$$\gamma_k = \tilde{h}_k p_k \tag{2.7}$$

$$\tilde{h}_k = \frac{G_p \; h_k}{\sum_{i \neq j} h_i p_i + \sigma^2} \tag{2.8}$$

p_k^* that optimizes u_k over the compact set $S_k = [0, P_{max}]$ is such that either it is on the boundary or it satisfies

$$\frac{\partial u_k}{\partial p_k} = 0 \ , \ \ p_k \in S_k$$

Proceeding by taking the derivative and using the linearity of SINR with transmit power, the best response $r_k(P_{-k})$ of user k is found as:

$$r_k(P_{-k}) = min \left\{ \frac{\gamma^*}{\tilde{h}_k}, P_{max} \right\} \tag{2.9}$$

γ^* is unique positive solution of the following equation [51]:

$$f(\gamma) = f'(\gamma)\gamma \tag{2.10}$$

The value of γ^* depends on the sigmoidal function $f(\gamma)$ such that the horizontal component of the intersection point in Fig. 2.4 is strictly greater than the inflection point of the sigmoid [51]. Note that the shape of the sigmoidal function is determined by modulation and coding scheme.

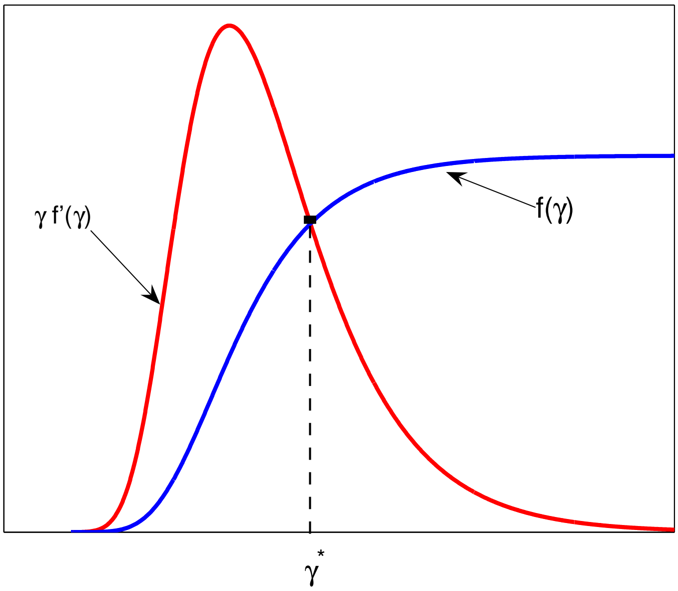

Figure 2.4: γ^* is unique positive valued solution of Eqn. (2.10)

By definition, solution(s) of the fixed point equations $p_k = r_k(P_{-k})$ is (are) NE(s). Consider the corresponding fixed point iteration $p_k(t+1) = r_k(P_{-k}(t))$ where the power of user i at iteration t is $p_i(t)$. The iterations converge to the unique fixed point iff NE is unique. To investigate the convergence of the fixed point iterations, it is useful to view the iterations as a power update algorithm $I(.)$ such that $\mathbf{p}(t+1) = I(\mathbf{p}(t))$ where $\mathbf{p}(t) = [p_1(t), p_2(t), ..., p_K(t)]$. In our problem, the explicit form of $I(.)$ is such that $I_i(\mathbf{p}(t)) = \min\{\hat{p}_i(t), P_{max}\}$ where

$$\hat{p}_i(t) = \frac{\gamma^* \sum_{j \neq i} h_j p_j(t) + \gamma^* \sigma^2}{G_p\ h_i}, \ \ i = 1, 2, ..., K$$

It is evident from the above expression that our $I(.)$ satisfies the *standard* power update algorithm definition of Yates [71]:

Definition 2.4.3 *$I(.)$ is a standard update algorithm if it satisfies*

1. *Positivity, $I(\boldsymbol{0}) > 0$*

2. Monotonicity, $I(\boldsymbol{p}_1) > I(\boldsymbol{p}_2)$*, whenever* $\boldsymbol{p}_1 > \boldsymbol{p}_2$

3. Scalability, $I(\alpha\boldsymbol{p}) < \alpha I(\boldsymbol{p})$, $\forall\alpha > 1$

From [71], if algorithm $\hat{I}(.)$ with $\hat{I}_i(\mathbf{p}(t)) = \hat{p}_i(t)$ is a standard algorithm, then $I_i(\mathbf{p}(t)) = \min\{\hat{p}_i(t), P_{max}\}$ has a unique fixed point. Hence, we conclude that our power update algorithm $I(.)$ has a unique fixed point, and consequently Ω_1 has a unique Nash equilibrium.

Theorem 2.4.4 Ω_1 *has a unique Nash equilibrium.*

In general, $p_k = r_k(P_{-k})\ \forall k$ form a system of K *non-linear* equations. In our particular problem in Eq. (2.9), non-linearity of r_k is due to clipping with P_{max}. If P_{max} is assumed sufficiently large, NE is a solution to the following system of K *linear* equations:

$$\frac{h_k p_k}{\sum_{i \neq k} h_i p_i + \sigma^2} = \gamma^*\ \forall k = 1, 2, .., K$$

The linear system above may have no solution, a unique solution or infinitely many solutions. If γ^* is feasible, then the system has a unique solution. The feasibility of γ^* can be determined using Perron-Frobenius Theorem[1] [56]. If the problem is analyzed in terms of received powers, one can show that the feasibility condition is $\gamma^* < \frac{G_p}{K-1}$. If this condition is not satisfied, power update algorithm $\hat{I}(.)$ diverges and for some of the users $p_i^* = P_{max}$ and $\gamma_i < \gamma^*$ while other users achieve γ^* at NE. On the other hand, the feasibility condition is necessary (but not sufficient) for all users to achieve γ^*.

Now that the power control game in a single access point system has been analyzed and properties of NE derived, we turn our attention to the general multiple access point system in the next section.

[1] Perron-Frobenius Theorem states that the maximum normed eigenvalue of a matrix with non-negative entries is positive valued and corresponding eigenvector has entries with same sign. It has implications about inequalities in the metric space of positive numbers. Th. 3.5.1 in the next chapter briefly explains the theorem. See [56] for a thorough statement of the theorem.

2.5 POWER CONTROL GAME IN MULTIPLE ACCESS POINT SYSTEM

Consider the general model with M access points (Fig. 2.1.) As before, users are subject to power constraint P_{max}. However, now they are allowed to transmit to more than one access point at a time. In other words, users can divide their power budget and transmit (different) data to different access points in order to (possibly) obtain a multiplexing gain.

In this case, the strategy set of a user is

$$S_{2k} = \left\{ [p_{k1} \; p_{k2} \; \ldots \; p_{kM}] \in \mathbf{R}_+^M : \sum_{j=1}^{M} p_{kj} \leq P_{max} \right\} \; \forall k \tag{2.11}$$

The utility function is as in Eq. (2.5):

$$u_{2k} = \frac{\sum_{b=1}^{M} T_{kb}}{\sum_{b=1}^{M} p_{kb}}$$

T_{kb} is the long term average rate of user k in access point b. We will analyze $\Omega_2 = [U, \{S_{2k}\}, \{u_{2k}\}]$ and corresponding user optimization is as follows:

$$\max_{\mathbf{p}_k \in S_{2k}} u_{2k}(\mathbf{p}_k, \mathbf{P}_{-k})$$

where $\mathbf{p}_k = [p_{k1}p_{k2} \ldots p_{kM}]$ and $\mathbf{P}_{-k} = [\mathbf{p}_1\mathbf{p}_2 \ldots \mathbf{p}_{k-1}\mathbf{p}_{k+1} \ldots \mathbf{p}_K]$.

The following theorem will reveal the special structure of the best response strategy, namely, each user transmits to a single access point:

Theorem 2.5.1 *The utility maximizing strategy of user k, $\boldsymbol{p}_k^*$, given $\boldsymbol{P}_{-k}$ in game Ω_2 is such that:*

$$p_{kb}^* = \begin{cases} p_k^*, & \textit{if } b = b_k^* \\ 0, & \textit{otherwise.} \end{cases} \tag{2.12}$$

$$b_k^* = \arg\max_b \left\{ \widehat{h}_{kb} \right\} \tag{2.13}$$

$$p_k^* = min(P_{max}, \frac{\gamma^*}{\widehat{h}_{kb_k^*}}) \tag{2.14}$$

$$\widehat{h}_{kb} = \frac{G_p \; h_{kb}}{\sigma_b^2 + \sum_{i=1 \;\; i \neq k}^{K} h_{ib} \sum_{j=1}^{M} p_{ij}} \tag{2.15}$$

Proof: The proof relies on the results obtained for single access point system. First, the set over which the optimization is performed is extended to $[0, P_{max}]^M$. The result for the optimum in single access point system is used, and a componentwise summation yields the desired conclusion. The details are given in Appendix 6. ■

Theorem 2.5.1 suggests that each user should just transmit to the access point that requires minimum power. The similarity of this strategy to the sum-rate optimum strategy of transmitting to a single user in each channel state in the fading broadcast channel model [63] is notable. The aim of the resource allocation formulation in [63] is maximizing long term average rate, given an average power budget. In contrast, in our formulation, the power budget is optimally divided among base stations to maximize the utility. Put in a different way, optimizing *energy efficiency* requires achieving a target SINR γ^* by choosing the best access point, while rate maximization allocates all power resource to the best channel.

In conclusion of theorem 2.5.1, the problem reduces to a joint access point assignment and power control problem. Consider a new game $\Omega_3 = [U, \{S_{3k}\}, \{u_{3k}\}]$. The strategy set S_{3k} for each user k is:

$$S_{3k} = B \times P \; \forall k = 1, 2, ..., N \tag{2.16}$$

where $B = \{b_1, b_2, \ldots, b_k\}$ and $P = [0, P_{max}]$, b_i being the i^{th} access point. Let b_k and p_k be the access point assignment and transmit power of user k, respectively. $\gamma_{kb_k} = \widehat{h}_{kb_k} p_k$ is the SINR of user k. Each user has the following utility function:

$$u_{3k} = R \frac{f(\gamma_{kb_k})}{p_k} \tag{2.17}$$

Access point assignment and power control game Ω_3 is originally proposed in [54]. In order to find the best-response strategy, optimization is performed in two stages [54]:

1. Given all users' powers are fixed, the selection of user k is $b_k^* = \arg\max_{b \in B} \gamma_{bk}$
2. Given access point b_k^* is assigned, the utility maximizing power level is $p_k^* = min \left\{ \frac{\gamma^*}{\bar{h}_{kb_k^*}}, P_{max} \right\}$

Note that the best response strategies of Ω_2 and Ω_3 coincide. In [54], Ω_3 is proved to have a unique Nash equilibrium. Similar to the single access point problem, the existence proof is based on compactness, convexity of $[0, P_{max}]$ and quasiconcavity of $f(\gamma)$; the uniqueness proof is by direct verification that the best response strategy is such that a standard update algorithm is clipped with power limit P_{max} (this is made precise in Appendix 6).

Theorem 2.5.2 *Ω_2 and Ω_3 have unique Nash equilibria. In NE of Ω_2, user k only transmits to the access point that is assigned in NE of Ω_3 with nonzero power and transmit power p_k in NE of Ω_2 and Ω_3 are equal for all k.*

Earlier work of Yates [72] posed a non-game theoretic integrated power control and access point assignment problem. While the formulation was not that of a game, it applies to the problem at hand. In a K user and M access point system, minimum total transmit power vector (MTP) is found under SINR constraints γ_i' where each user is assigned to only one access point.

$$\gamma_i \geq \gamma_i^{'} \; \forall i = 1, 2, ..., K \tag{2.18}$$

If $\gamma_i^{'}$ are feasible then there exists a unique solution for MTP problem. Perron-Frobenius Theory [56] is again deployed for analyzing feasibility. Assuming that each user is assigned to a fixed access point, SINR constraints and channel gains are combined in one matrix. The feasibility condition is that resultant Perron-Frobenius eigenvalue $\lambda_{PF} < 1$ for some assignment among K^M possible assignments.

Set $\gamma_k' = \gamma^*$. Provided that P_{max} is sufficiently large and feasibility is satisfied, the Nash equilibrium point of Ω_2 and Ω_3 are equivalent to the unique solution of minimum total transmit power problem with $\gamma_k' = \gamma^*$.

It is interesting to extend the system model to a multicarrier multiple-access point system. The existence of multiple carriers introduces an extra dimension to the strategy sets: now users, by picking which subcarriers to use to access which base station (and how much of their total power to allocate to it), are picking a *subset* of interferers. The results in [38] and Th. 2.5.1 straightforwardly combine to conclude that the best response strategy would be to transmit to only one access point by putting the total power on a single carrier. However, the game may not have *unique* NE in this case: Due to the orthogonality among the carriers, monotonicity and thus the standardness property cease to hold: as different users can transmit in different carriers, one user may not respond to an increase in another user's power. Therefore, the uniqueness of NE is not guaranteed, and in fact, as observed in [38], for some values of channel gains, multiple Nash Equilibria may exist.

2.6 ENERGY AND SPECTRAL EFFICIENCY

The analyses of games for single access point system in § 2.4 and for multiple access point system in § 2.5 were based on energy efficient utility with unit *bits per joule*. Unilateral optimization of utilities led users to reach a target SINR γ^*, which is the unique solution of the equation $f(\gamma) = f'(\gamma)\gamma$. However, the spectrum resource is inefficiently used in case γ^* has a low value. This observation points to the trade-off between energy and spectral efficiency.

In order to make the trade-off more clear in the game settings, we introduce a priority exponent $\alpha > 0$. Consider the single access point system. We propose to change the utility function as follows:

$$u_k = \frac{T_k}{p_k^\alpha} \tag{2.19}$$

The priority exponent α brings a variable degree of energy efficiency to the utility function. For $\alpha = 1$, the utility function in Eq. 2.6 is obtained. $\alpha < 1$ means that users value spectral efficiency more while $\alpha > 1$ drives users to be more energy-efficient.

The equilibrium SINR γ^* is the *unique* solution of

$$f(\gamma) = \frac{1}{\alpha} f'(\gamma)\gamma$$

The variation of the γ^* for different values of α is shown in Figs. 2.5 and 2.6.

A similar modification can be made to the utility function for multiple access point system:

$$u_k = \frac{\sum_{b=1}^{M} T_{kb}}{\sum_{b=1}^{M} p_{kb}^\alpha} \tag{2.20}$$

In this case, the best response is again transmitting to a single access point that requires lowest power to reach γ^*. Variation of γ^* with respect to α is same as that depicted in Figs. 2.5 and 2.6.

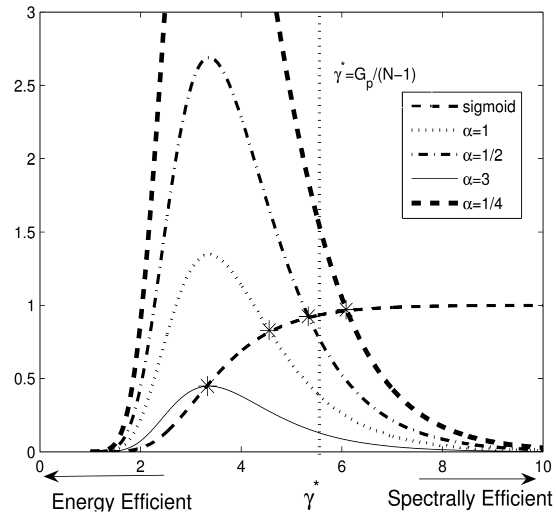

Figure 2.5: γ^* is the intersection of two functions $f(\gamma)$ and $\frac{1}{\alpha}f'(\gamma)\gamma$. The horizontal axis component of intersection points is the equilibrium SINR γ^*. For $\alpha < 1$, as the priority of spectral efficiency is higher, γ^* takes higher values. It is not possible for users to mutually reach γ^* in case $\gamma^* > \frac{G_p}{K-1}$

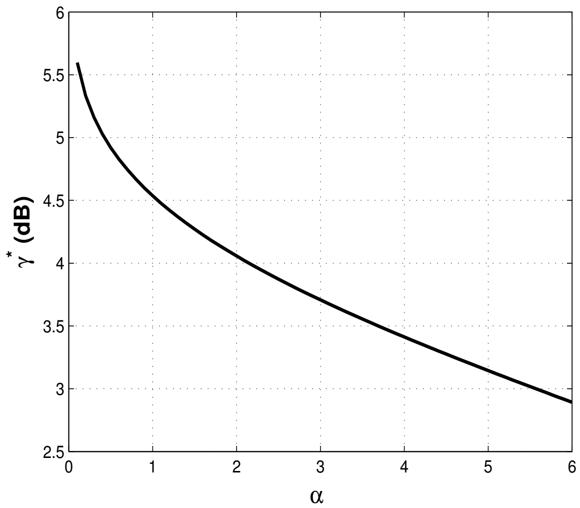

Figure 2.6: Variation of γ^* with respect to the priority exponent α. BPSK modulation with 1000 bit length packets is assumed.

2.7 CONCLUSION

In this chapter, we analyzed vector power control in the uplink of a general multiple access point system using a game theoretic framework. Given the other users' strategies, each user optimizes its utility. Utility function is chosen based on users' priorities. With an energy efficiency motivation, utility was first chosen in the form of rate over power. A vector power control game was proposed using this utility function. Best response strategy of the game was shown to have a special structure: transmitting to a single access point. Hence, the game reduced to access point selection and power control. Existence and uniqueness of the game was established using this special structure.

The best-response strategy basically leads to a target SINR based power control algorithm. Target SINR γ^* is determined by the coding and modulation type. The Nash Equilibrium (NE) of the game, and existence conditions for it were shown. When γ^* is feasible, NE corresponds to the minimum total transmit power vector under quality of service constraint γ^*.

Finally, the utility function was modified to have a variable degree of energy efficiency. The variation in the target SINR γ^* with respect to priority exponent was analyzed and the trade-off between energy and spectral efficiency was verified.

CHAPTER 3

DISTRIBUTED POWER CONTROL USING NON-MONOTONIC REACTION CURVES

3.1 INTRODUCTION

The material in this chapter partially appears in Proceedings of 2009 International Conference on Game Theory for Networks in [43]. The work in this chapter has been performed as a part of the project [66].

In general, a wireless network could be viewed as an interference network. In such a network, power control could be employed to provide an acceptable level of service to as many users as possible with the minimum amount of energy consumption. Distributed power control mechanisms are attractive in networks where central control is not practical.

Among earlier distributed power control algorithms, the proposals of [14,75] should be noted. In [14], the goal is to guarantee all users a certain target signal to interference ratio (SIR^{tar}). The following power iteration (to be performed at each terminal) is at the core of the algorithm:

$$P(t+1) = SIR^{tar}\frac{I(t)}{G} \tag{3.1}$$

where $I(t)$ is total received interference power at time t and G is an effective channel gain. In the rest, we will refer to this as Algorithm 1 and a power-limited version of it as Algorithm 2. It should be noted that when the target SIR is not feasible, Algorithm 1 diverges [70] (Figure (3.5)).

In [75], the number of users that achieve a target SIR is optimized. When this target is not feasible, the users whose performances are degraded by interference most are removed from the system step by step (referred to as SRA in [75]). The effect of interference on a user is measured by sum of interfering channel gains divided by user's own channel gain. Although a center decides which user is removed next, distributed power iterations similar to Eq. (3.1) are performed to decide that target SIR is not feasible (in the rest, we refer to this stepwise removal procedure as Algorithm 3).

Game theory has been a natural tool for developing distributed power control mechanisms. A main stream of questions has revolved around the choice of utility function and the quality of the resulting equilibria. In [19], a non-cooperative game is set up based on a certain utility function at the wireless nodes; when equilibrium turns out to be inefficient, price functions are introduced which result in Pareto improvements. In [53], Pareto optimal pricing that involves central processing is derived. In [18], a different central mechanism is proposed to improve the equilibrium: simultaneous reduction of target SIR values. In [70], net utilities with linear pricing led to an *admission control* mechanism.

By adjusting the linear price coefficients as a function of channel gain [3] or interference power [70], it is possible to obtain distributed power control algorithms that address important issues such as throughput efficiency, delay tolerance and near-far fairness. Perhaps pricing is most suited in a framework where the service provider (network operator) is also an actor in the game [1, 12], with a goal of maximizing its own revenue. However, pricing is not very natural in systems like wireless ad-hoc and sensor networks where agents are *cooperative* rather than *competitive*. Ironically, a level of cooperation emerges when users non-cooperatively optimize their net utilities as discussed in [70]. Each user decreases its objective with increasing interference. We investigate this cooperation emanating from the non-cooperative behaviour.

In admission-control type approaches (with or without pricing), some users are discarded from the system in order to provide certain guarantees for others. Such schemes can be characterized by *reaction curves* [6] that discontinuously fall to zero. We argue that this is not ideal for data-intensive applications, which have a certain degree of flexibility and would live with a lower than perfect SIR. Non-monotonic and continuous reaction curves may ease the inflexibility. In this context, we aim for a smooth mechanism of distributed power control via suitably designed utility functions.

Energy-efficiency has been central to the utility function in a number of recent game-theoretic power control formulations [38, 39]. A simple utility function with a power term in the denominator serves this purpose quite well, however such an approach turns out to be too simplistic for practical purposes as resulting equilibria may have very low spectral efficiency. It seems that the energy-efficiency goal should be combined with some acceptable quality objective about spectral efficiency, rate, or SIR. How this should be done requires more thought.

At this point we should recognize that as wireless channels are typically time varying, the definition of "acceptable quality" will involve time averages or a probabilistic statement on an appropriately chosen statistical channel model. One of the simplest quality metrics is outage probability. Study of distributed power control within an outage formulation is relatively recent. For example, in [47], outage probability constraints in CDMA systems with Rayleigh faded links are studied and distributed algorithms are devised to jointly control power and receiver filter coefficients. Extensive analyses of price functions and update algorithms are presented in [2].

The contribution of this chapter is to exhibit a utility function, which, when used in a non-cooperative game, allows power-efficient approximation to outage probability targets. The key to this behavior is the non-monotone and sub-homogeneous reaction curve induced by the utility function developed. This can be viewed as an alternative to pricing. We establish a sufficient condition for the existence and uniqueness of (Nash) equilibrium of our distributed algorithm. Moreover, we show how an admission control-type power control algorithm can be smoothened so that the sufficient condition is satisfied.

The next section sets up the system model, followed by the construction of our utility in §3.3. Then, definition of Nash Equilibria (NE) and their interpretation in terms of reaction curves are presented in §3.4. §3.5 is devoted to the analysis of feasibility and optimality. Smooth reduction of objectives is analyzed in §3.6. We illustrate our results with a numerical example in §3.7 and conclude the chapter in §3.8.

3.2 SYSTEM MODEL

We consider a wireless network that contains a number of interfering links $\{i = 1, \ldots, K\}$, each link corresponding to a distinct transmitter as in Fig. 3.1. We denote index set of links as $< K >$. Link i has a transmitter $t(i)$, and receiver $r(i)$. The receivers $r(i), i = 1, \ldots, K$ are not necessarily distinct. (For example, the case of all $r(i)$'s being the same models the uplink of a centralized network having a single base station or access point.)

The channel gain between $t(i)$ and $r(j)$ is $G_{ij}h_{ij}$. That is, due to a signal of power P_i transmitted by the sender of link i, the receiver of link j receives a signal (interference if $i \neq j$) of power $G_{ij}h_{ij}P_i$. The h_{ij} are fading coefficients, while the $\{G_{ij}\}$ model interference mitigation due to the specific channel allocation and coding scheme used. This model is quite general in that, by proper choice of the coefficients $\{G_{ij}\}$, it can model centralized or distributed network architectures, as well as different channel access mechanisms such as CDMA, TDMA and FDMA.

For convenience, in the rest of the chapter, a *Rayleigh-Rayleigh* [26] environment is assumed, i.e., all signal and interference terms are subject to Rayleigh fading. Accordingly, the h_{ij} are independent exponentially distributed random variables with unit variance. G_{ij} are assumed constant during the operation. For convenience, additive noise and intercell interference are modeled as Gaussian random variables with zero mean and noise variance is σ^2.

We consider *single user decoders* in the receivers. Depending on the receiver structure, cross channel gains may be suppressed by additional processing gains. We do not focus on receiver type; hence, we assume that additional gains are also represented in the channel gains. The signal to interference ratio (SIR) of user i is given as:

$$\gamma_i = \frac{G_{ii}h_{ii}P_i}{\sigma^2 + \sum_{j \neq i} G_{ji}h_{ji}P_j} \tag{3.2}$$

Each user is subject to power constraint: $P_i \leq P_{max}\ \forall i$. Wireless links undergo outage with probability O_i

$$O_i \triangleq Pr(\gamma_i < \gamma^{th})$$

where γ^{th} is threshold SIR required for communication. γ^{th} is determined by several system properties such as rate, modulation and receiver structure, which is outside

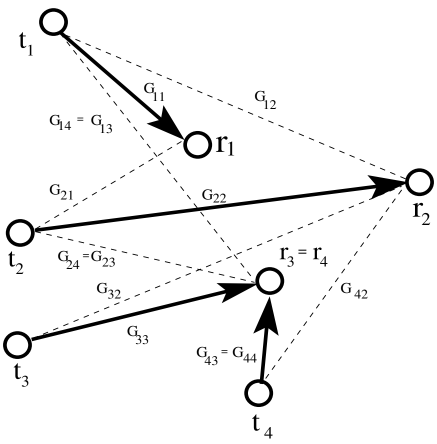

Figure 3.1: A set of interfering links

the scope of our work (e.g. see [20]). In the rest, we assume the same γ^{th} for each user[1].

The expression of O_i for the Rayleigh/Rayleigh environment is not derived here, but rather we forward reader to [47]. Let $x_{ji} = E_h(G_{ji}h_{ji}P_j) = G_{ji}P_j$ be average interference power of user j on user i and $x_{ii} = E_h(G_{ii}h_{ii}P_i) = G_{ii}P_i$ is average received signal power of user i.

$$O_i = 1 - exp(\frac{-\sigma^2\gamma^{th}}{x_{ii}})\prod_{j\neq i}\frac{1}{1+\frac{\gamma^{th}x_{ji}}{x_{ii}}} \tag{3.3}$$

In [26], outage probability in Rayleigh faded noiseless multiple access channel is upper bounded by an exponential term and this result was extended in [49] for the noisy case as follows:

$$\Pr(\gamma_i < \gamma^{th}) \leq 1 - exp(-\frac{\gamma^{th}}{SIR_i^{avg}}) \tag{3.4}$$

where

$$SIR_i^{avg} \triangleq \frac{G_{ii}P_i}{\sigma^2 + \sum_{j\neq i} G_{ji}P_j}$$

SIR_i^{avg} is average SIR with respect to all h_{ji}. In [26], it is implied that above upper bound well approximates the outage probability especially when the outage probability is below 20%. In fact, the approximation becomes very tight for both extremes.

Users have a certain tolerance to outages. We consider quality of service as utility and define it in terms of outage probability tolerances. As power control is addressed,

[1] The results straightforwardly extend to the case of different γ^{th}'s for users, which can model e.g.different receiver structures.

utility is defined as a function of average power. In the next section, the utility function will be characterized.

3.3 UTILITY FUNCTION

In game-theoretic terms, a utility function is a mapping from the action space into real numbers $u_i : \prod_{j=1}^{K} A_j \rightarrow \mathbf{R}$. In our distributed power control algorithms, the action space will contain power levels used. The actual utility of a user may depend on the type of application as well as constraints such as energy. For example, a suitable utility could be a step function (or sigmoid) of SIR for constant-rate real-time applications. On the other hand, a (function of) packet success probability divided by average power has been a popular utility function for delay-tolerant data (with ARQ) when energy is expensive [10, 53].

In [70], utility is a softened step function, which is based on target SIR. Similarly, our starting point is outage objectives. Suppose each user i has an outage probability objective O^{obj} that it wants to stay below. Being close to the outage objective will carry some utility, as long as it is not too costly in terms of power. We can capture this idea in the following simple utility function, which we shall call Utility 1:

$$u_i = \frac{exp(-\frac{SIR^{obj}}{SIR_i^{avg}})}{P_i} \tag{3.5}$$

It is shown in [47] that given an outage objective, there is an SIR^{obj} such that $SIR_i^{avg} \geq SIR^{obj}$ ensures that the outage objective is reached. From (3.4), if SIR^{obj} is chosen such that $exp^{-\frac{\gamma^{th}}{SIR^{obj}}} = 1 - O^{obj}$, the objective is automatically satisfied. This sets up a correspondence between outage and SIR objectives similar to the one in [47]. The goal of Utility 1 will be enabling users to achieve their outage objectives using minimum power. For simplicity, we assume that the target outage probability is the same for all users, though extending the results to the case of uneven outage probability targets is straightforward.

One of the key properties of the utility function in Eq. 3.5 is the quasiconcavity. This property again plays a key role in the existence of an equilibrium in the power control game.

We will devise a distributed mechanism by setting up the power control problem as a non-cooperative game and analyze the properties of the resulting equilibrium. We will first introduce the concepts of equilibrium and reaction curves. Then, we will analyze the established game.

3.4 EQUILIBRIUM IN TERMS OF REACTION CURVES

A Nash Equilibrium (NE) [6] in our setting is the following.

Definition 3.4.1 *A NE is a power allocation* $\boldsymbol{P}^* = [P_1^*, P_2^*, P_3^*, \ldots, P_K^*]$ *such that*

$$u_i(P_i^*, \boldsymbol{P}_{-i}^*) \geq u_i(P_i, \boldsymbol{P}_{-i}) \ \forall P_i \in A_i \ \forall i \in <\boldsymbol{K}>$$

where $\boldsymbol{P}_{-i} = (P_1, P_2, .., P_{i-1}, P_{i+1}, .., P_K)$.

Each user imposes optimization in its own action variable P_i. Given rival users' actions $\mathbf{P}_{-i}$, user i responds by selecting an action $\hat{P}_i$ such that $u_i(\hat{P}_i, \mathbf{P}_{-i}) \geq u_i(P_i, \mathbf{P}_{-i}) \ \forall P_i$. The set of such $\hat{P}_i$'s form the *reaction set* of user i:

$$r_i(\mathbf{P}_{-i}) = \left\{ \bar{P}_i : u_i(\bar{P}_i, \mathbf{P}_{-i}) \geq u_i(P_i, \mathbf{P}_{-i}) \ \forall P_i \in A_i \right\}$$

If the reaction set is a singleton, then it is called a *reaction curve* [6] and $\hat{P}_i$ is the *best response action* to $\mathbf{P}_{-i}$. NE finds another interpretation in terms of reaction sets. A vector of actions $(P_1^*, P_2^*, \ldots, P_K^*)$ is NE if $P_j^* \in r_j(\mathbf{P}_{-j}^*) \ \forall j$.

For Utility 1 (Eq. (3.5)), it can be shown that reaction set of each user is singleton. As action set for user i, $S_i = [0, P_{max}]$, is a compact set in $\mathbf{R}$, one can proceed with well known analysis on $\frac{\partial}{\partial P_i} u_i$. u_i has a unique maximum at either P_{max} or $\tilde{P}_i$ such that $\frac{\partial}{\partial P_i} u_i(\tilde{P}_i) = 0$. Using straightforward calculus,

$$\tilde{P}_i = SIR^{obj} \sum_{j \neq i} \frac{G_{ji}}{G_{ii}} P_j + \frac{SIR^{obj}}{G_{ii}} \sigma^2$$

From above discussion, we can state the reaction curve as:

$$r_i(\mathbf{P}_{-i}) = \min \left\{ P_{max}, \ \tilde{P}_i \right\} \tag{3.6}$$

Note that $r_i(\mathbf{P}_{-i})$ is the same as the interference function in (3.1) when P_{max} is sufficiently large. $r_i(\mathbf{P}_{-i})$ is the reaction of user i to interference power. In this context, we consider reaction curves as power update algorithms: $P_i(n+1) = r_i(\mathbf{P}_{-i}(n))\ \forall i$. Incidentally, this corresponds to user i setting its SIR to SIR^{obj} as far as its power limit allows, i.e., a clipped version of the algorithm in (3.1). In the rest, we refer to it as Algorithm 2.

The following definition from [42] will be useful.

Definition 3.4.2 *A vector power update algorithm $\boldsymbol{F}(.)$, such that $\boldsymbol{P}(t+1) = \boldsymbol{F}(\boldsymbol{P}(t))$, is monotonic and strictly subhomogeneous (MSS) if*

- $\boldsymbol{F}(P_2) > \boldsymbol{F}(P_1)\ \forall P_2 > P_1$
- $\boldsymbol{F}(\alpha P) < \alpha \boldsymbol{F}(P),\ \alpha > 1$

MSS algorithms are *standard* [71], as above properties imply positivity. It can easily be verified that the update algorithm $\mathbf{r}(.)$ is MSS when $[\mathbf{r}(.)]_i = r_i(\mathbf{P}_{-i}(n))$, the reaction curve in (3.6). In the next section, we investigate equilibria of this game.

3.5 THE POWER CONTROL GAME

In this section, we analyze the game Ω_1

$$\Omega_1 = [< K >, \{S_i\}, \{u_i\}]$$

Initially, we let the power constraints go to infinity, and present results on feasibility and optimal control for the outage probability objectives and analyze NE of Ω_1. Later, the finite P_{max} case is investigated.

3.5.1 Feasibility and Optimal Power Control

We let P_{max} be sufficiently large. For the setting with outage probability objectives, consider the set T

$$T = \left\{ [P_1, P_2, ..., P_K] \in \mathbf{R}_+^K : O_i \leq O^{obj} \right\}$$

Exact characterization of T is still an open problem. Yet, as we discussed in the utility definition, we can obtain a subset of T that is nearly equal to it (especially when O^{obj}

is small) by transforming the outage constraints into $SIR_k^{avg} \geq SIR^{obj}\ \forall k$. Therefore in the rest, we investigate the feasibility of the SIR constraints.

$SIR_k^{avg} \geq SIR^{obj}$ implies $\mathbf{P} \geq SIR^{obj} A\mathbf{P} + \mathbf{n}$ where matrix A is defined as

$$A_{mn} = \begin{cases} 0, & m = n \\ \frac{G_{nm}}{G_{mm}}, & m \neq n. \end{cases} \tag{3.7}$$

and the column vector $\mathbf{n}$ with entries $n_i = SIR^{obj} \frac{\sigma^2}{G_{ii}}$. Related to the existence of a feasible power vector, we collect important results of Perron-Frobenius theorem [56] in the following.

Theorem 3.5.1 *Let B be a square non-negative irreducible matrix. There exists a unique non-negative $\boldsymbol{x}$ that satisfies $\boldsymbol{x} = B\boldsymbol{x} + \boldsymbol{c}$, $\boldsymbol{c} > \boldsymbol{0}$ iff Perron-Frobenius eigenvalue of B, which is guaranteed to exist and to be unique, satisfies $\lambda_{PF} < 1$.*

If λ_{PF} is Perron-Frobenius eigenvalue of matrix A, $SIR^{obj} < \frac{1}{\lambda_{PF}}$ is necessary and sufficient for a feasible vector.

Consider the minimum transmit power problem:

$$min \sum_i P_i \text{ subject to } \mathbf{P} \in T$$

If T is replaced with its almost equal subset, we can obtain a near optimal solution. In that case, the power allocation that satisfies $SIR_i^{avg} = SIR^{obj}\ \forall i$ is the near optimal solution. Note that such a near optimal solution exists iff $SIR^{obj} < \frac{1}{\lambda_{PF}}$.

3.5.2 Nash Equilibrium under $P_{max} \to \infty$

Reaction sets in (3.6) for sufficiently large P_{max} become affine hyperplanes in $\mathbf{R}_+^K$ such that $SIR_i^{avg} = SIR^{obj}\ \forall i$. Therefore, NE of the game is solution of K unknown linear system:

$$G_{ii}P_i - SIR^{obj} \sum_{j \neq i} G_{ji}P_j = SIR^{obj}\sigma^2\ \forall i \tag{3.8}$$

Any solution of the system other than trivial solution is NE of Ω_1. Let I denote $K \times K$ identity matrix.

$$[I - \ SIR^{obj}\ A]\mathbf{P} = \mathbf{n}$$

Any solution of above matrix equation that has strict positive entries is NE of Ω_1. Let Perron-Frobenius eigenvalue of A be λ_{PF}. From theorem 3.5.1, if $SIR^{obj} < \frac{1}{\lambda_{PF}}$, there exists a unique NE. Otherwise, NE does not exist.

Remark 1 *If $SIR^{obj} < \frac{1}{\lambda_{PF}}$, then NE of Ω_1 is minimum total transmit power vector in the set defined by SIR constraints.*

Note that the cost term in the denominator of utility function leads to cost efficient behaviour and that further results in minimum total transmit power.

3.5.3 Finite P_{max} Case

If the power is limited by $P_{max} < \infty$, the feasible set T and corresponding set of vectors that meet SIR constraints are also constrained. In particular, there may not exist power vectors such that $SIR_i^{avg} > SIR^{obj}\ \forall i$ even if $SIR^{obj} < \frac{1}{\lambda_{PF}}$.

As in chapter 2, we can guarantee the existence of NE by quasiconcavity property of utility function and compactness, convexity of strategy set (see Theorem 1 in [53]).

Let the reaction curve be deployed as an update algorithm, $P_i(t+1) = r_i(\mathbf{P}_{-i}(t))$. As argued before, this algorithm is MSS and it will always have a unique fixed point in the finite P_{max} case [71]. As any fixed point of the algorithm is NE of the game, there exists a unique NE for the finite P_{max} case. Figure 3.2 illustrates representative reaction curves and NE.

In Fig. 3.2, it is observed that users inefficiently transmit with power P_{max} in NE when the outage probability objectives are not feasible, resulting in NE on the boundary of action set. Usually, pricing functions are introduced and net utility is optimized to avoid this inefficiency [53]. Yet, apart from bringing Pareto improvements to the network, pricing in this type of problems may have economical implications [1] and network owner can be added to the game formulation [12]. However, in networks that users cooperate (say, for a common mission) without a network owner such as ad-hoc and sensor networks, economic approaches may not make sense. In fact, the effect of pricing is basically to decrease target SIR (and even turn off transmission) as a function of interference [70]. We argue that it is more convenient for mentioned

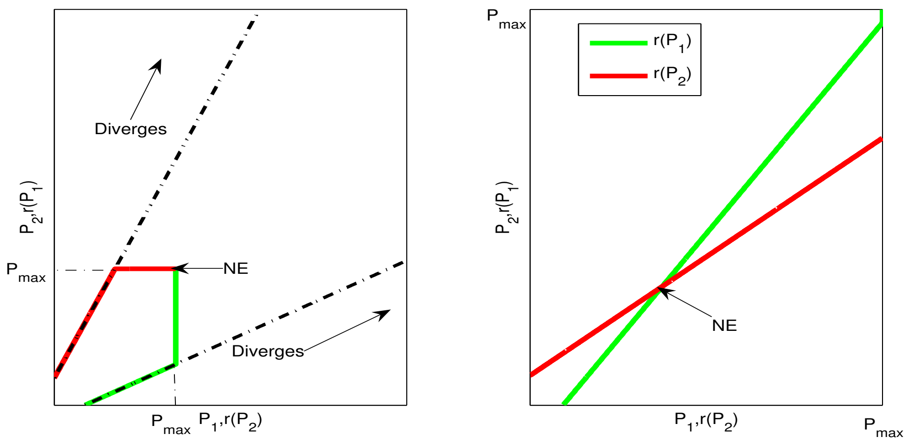

Figure 3.2: Illustration of reaction curves of two players for feasible and infeasible cases. On the left, $\lambda_{PF} > \frac{1}{SIR^{obj}}$, for the selected gain values, hence the feasible set is empty. NE is shown for $P_{max} = 1$. For $P_{max} \to \infty$, power of users diverge. On the right, $\lambda_{PF} < \frac{1}{SIR^{obj}}$

types of networks to investigate barely the decrease in target SIR with respect to interference.

With pricing, a sharp cutoff of target SIR (corresponding to turning off) will inevitably occur at some point. This is not ideal when variable bit rate applications are predominant, since they have relatively flexible demands from the network. We shall thus modify Utility 1 so that users decrease their objectives smoothly as a reaction to interference. In the next section, we analyze the game resulting from this modified utility.

3.6 SMOOTH REDUCTION OF OBJECTIVES

As argued above, feasibility of the SIR (or outage) objectives depends on the power constraint and the gain matrix (exhibited in the relationship of the Perron-Frobenius eigenvalue and the SIR objective). To illustrate, a user may not reach its objective under high interference and relatively low channel gain. When such cases occur, users reach an inefficient equilibrium in game Ω_1 and at least one of them forces the power limit. In order to mitigate this inefficiency, we propose that users smoothly decrease their objectives as a reaction to interference. In this context, we propose the objective reduction factor f_d^i for user i, which exhibits itself in the following power iterations:

$$P_i(t+1) = SIR^{tar} f_d^i(I_i(t)) \frac{I_i(t)}{G_{ii}} \tag{3.9}$$

Consider the following modification in utility function (3.5):

$$u_i^m = \frac{exp(-\frac{SIR^{obj} f_d^i(\sum_{j \neq i} G_{ji} P_j)}{SIR_i^{avg}})}{P_i} \tag{3.10}$$

We call the function in (3.10) Utility 2. Factor $f_d^i(.)$ are basically non-increasing continuous functions defined from non-negative real numbers into the range $[0, 1]$. Note that the power update algorithm in Eqn. (3.9) can be obtained if users accept Utility 2 as their net utility.

It can be shown that reaction curve of user i for Utility 2 is

$$r_i(\mathbf{P}_{-i}) = \min \left\{ P_{max},\ \breve{P}_i \right\} \tag{3.11}$$

where

$$\breve{P}_i = SIR^{obj} f_d^i(\sum_{j \neq i} G_{ji} P_j)(\sum_{j \neq i} \frac{G_{ji}}{G_{ii}} P_j + \frac{\sigma^2}{G_{ii}})$$

3.6.1 A Sufficient Condition for Existence and Uniqueness of Equilibrium

We will present a sufficient condition on functions f_d^i $\forall i$ for existence and uniqueness of equilibrium. Our analysis will make use of some results from [42]. We begin with a definition.

Definition 3.6.1 *A real valued function $\boldsymbol{f}$ is called absolutely subhomogeneous (AS) if the following inequalities hold for all $x \in \boldsymbol{R}$ and arbitrary a*

$$e^{-|a|} \boldsymbol{f}(x) \leq \boldsymbol{f}(e^a x) \leq e^{|a|} \boldsymbol{f}(x)$$

AS functions are basically smooth functions with rate of increase less than linear and rate of decrease smaller than inverse of linear. The following lemma implies that clipping an AS function yields another AS function.

Lemma 3.6.2 *If $\boldsymbol{f}(P)$ is AS, then so is* $\min\{P_{max},\ \boldsymbol{f}(P)\}$.

Proof:

For arbitrary a, we have AS function f

$$e^{-|a|}f(x) \leq f(e^a x) \leq e^{|a|}f(x)$$

Let $g(x) = min\{P_{max}, f(x)\}$ Without loss of generality, assume that a is positive. Obviously,

$$g(e^a x) = min\{P_{max}, f(e^a x)\}$$

$$\leq min\{P_{max}, e^{|a|}f(x)\}$$

$$\leq e^{|a|}min\{P_{max}, f(x)\}$$

To show that left inequality in the AS property is satisfied for $g(x)$, we need to consider four different conditions as a result of truth values of $f(x) < P_{max}$ and $f(e^a x) < P_{max}$.

For example, if $f(x) < P_{max}$ and $f(e^a x) < P_{max}$, then $g(x) = f(x)$ and $g(e^a x) = f(e^a x)$. Hence, $e^{-|a|}g(x) \leq g(e^a x)$ by AS assumption on $f(.)$. It can also be verified for other cases as follows:

If $f(x) > P_{max}$ and $f(e^a x) < P_{max}$, then $g(x) = P_{max}$ and $g(e^a x) = f(e^a x)$. So $e^{-|a|}g(x) = e^{-|a|}P_{max} \leq e^{-|a|}f(x) \leq f(e^a x) = g(e^a x)$.

If $f(x) < P_{max}$ and $f(e^a x) > P_{max}$, then $g(x) = f(x)$ and $g(e^a x) = P_{max}$. Then, $e^{-|a|}g(x) = e^{-|a|}f(x) \leq f(e^a x) \leq P_{max} = g(e^a x)$.

If $f(x) > P_{max}$ and $f(e^a x) > P_{max}$, then $g(x) = P_{max}$ and $g(e^a x) = P_{max}$. Thus, $e^{-|a|}g(x) = e^{-|a|}P_{max} \leq P_{max} = g(e^a x)$. ■

It will be useful for our purposes to collect results in [42] about AS functions and uniqueness of fixed points in the following theorem.

Theorem 3.6.3 *Let* $\boldsymbol{P}(t+1) = F(\boldsymbol{P}(t))$ *be MSS and let* $H(\boldsymbol{P}) = \phi(F(\boldsymbol{P}))$. *If vector function* ϕ *is in the form* $\phi(x_1, x_2, ..., x_K) = [\phi_1(x_1), \phi_2(x_2), ..., \phi_K(x_K)]$ *with* $\phi_i(x_i)$ *absolutely subhomogeneous, then* $\boldsymbol{P}(t+1) = H(\boldsymbol{P}(t))$ *has a unique fixed point.*

Now, let $F_i(\mathbf{P}) = SIR^{obj} \sum_{j \neq i} \frac{G_{ji}}{G_{ii}} P_j + SIR^{obj} \frac{\sigma^2}{G_{ii}}$. We mentioned earlier that vector update algorithm $F(\mathbf{P})$ is MSS. By means of some algebraic manipulations, we can express $\breve{P}_i$ in (3.11) as

$$\breve{P}_i = \phi_i(F_i(\mathbf{P}))$$

where $\phi_i(x_i) = x_i f_d^i(\frac{G_{ii}}{SIR^{obj}} x_i - \sigma^2 G_{ii})$. If we choose $f_d^i(x)$ such that $\phi_i(x_i)$ is AS, then by using above lemma and theorem, vector power update algorithm $H(\mathbf{P})$ with $H_i(\mathbf{P}) = \min\left\{P_{max},\ \breve{P}_i\right\}$ has a unique fixed point. We call the resulting vector power update scheme non-monotonic power control (NMPC) and we express it as follows:

$$P_i(t+1) = H_i(\mathbf{P}(t)) \tag{3.12}$$

3.6.2 Asynchronous Power Updates

In a real life implementation of non-monotonic power control, it may not be possible to maintain synchronization among power updates. Hence, a natural generalization of the power updates in (3.12) is the *asynchronous* version in which an arbitrary subset of users $S(t) \subset <K>$ update their power in each time slot:

$$\mathbf{P}_{S(t)}(t+1) = H(\mathbf{P}(t)) \tag{3.13}$$

The definition of asynchronous updates are similar to [7,71], hence it is not given here. Although $S(t)$ is arbitrary in each time slot, a key assumption about the asynchronous updates is that each user updates their power infinitely often. No user stops updating their power level. Each user uses past information about the power levels of users and due to infinitely often update assumption, the power levels at a specific time always become obsolete. A set of sufficient conditionfor the convergence of asynchronous algorithm in Eq. (3.13) is provided by Bertsekas and Tsitsiklis in [7], summarized in the following theorem:

Theorem 3.6.4 *Totally asynchronous fixed point iterations*

$$\boldsymbol{P}_{W(t)}(t+1) = I(\boldsymbol{P}(t))$$

where $W(t) \subset U$ is an arbitrary subset of user indices that update their power at time t, converge to the unique fixed point of the synchronous iterations if the following two conditions hold

- *If the sequence of sets* $Z(k) \subset [P_{min}, P_{max}]^K$ *such that* $\{Z(k) \supset I(Z(k-1))\}$ *satisfy* $... \subset Z(k) \subset Z(k-1)... \subset Z(1) = [P_{min}, P_{max}]^K$.

- *If there exist bounded sets* $Q_i(k) \subset \Re$ *such that the set* $Z(k)$ *can be expressed as*

$$Z(k) = Q_1(k) \times Q_2(k) \times ... \times Q_K(k)$$

In Theorem 3.6.4, the first condition is called *Synchronous Convergence Condition* and the second one is called *Box Condition*.

The synchronous convergence condition guarantees the convergence of synchronized updates, yet an update algorithm does not necessarily satisfy that condition if it is known to converge. Actually, the convergence of non-monotonic power control is based on AS property, which relies on contraction property with respect to the following metric [42]:

$$\mu(\mathbf{x}, \mathbf{y}) = max_i \, |log(x_i) - log(y_i)|$$

Hence, in the first sight, non-monotonic power updates may not satisfy the synchronous convergence condition. However, a careful investigation of the problem reveals that the synchronous convergence condition and box condition are indeed satisfied by the non-monotonic power updates.

To show this, let $Q_i(1) = [0, P_{max}]$, $Z(1) = \prod_{i=1}^{K} Q_i(1)$, $Z_i(1) = \prod_{j \neq i}^{K} Q_i(1)$ and the unique fixed point be $\mathbf{P}^*$. As AS functions cannot have the value zero [42], there exists $b_i > 0$ and corresponding $Q_i(2) = [b_i, P_{max}]$ such that $Z(2) = \prod_{i=1}^{K} Q_i(2)$ satisfies $Z(2) \supset H(Z(1))$ and $Z(2) \subset Z(1)$. It is also clear that $Z(3) \supseteq H(Z(2))$ and $Z(3) \subseteq Z(2)$ where $Z(3) = \prod_{i=1}^{K} Q_i(3)$ and $Q_i(3) = [b_i', c_i]$ with $b_i' = \min_{\mathbf{P}_{-j} \in Z_j(2)} H_i(\mathbf{P}_{-j})$ and $c_i = \max_{\mathbf{P}_{-j} \in Z_j(2)} H_i(\mathbf{P}_{-j}) \leq P_{max}$. Using the contraction property, for some $0 < \rho < 1$, $\max_i |log(H_i(\mathbf{P})) - log(P_i^*)| < \rho \max_i |log(P_i) - log(P_i^*)|$ for all $\mathbf{P} \in Z(2)$. Since $log(.)$ function is monotone increasing, then we have $Z(3) \supset H(Z(2))$ and $Z(3) \subset Z(2)$. We can show by an induction argument that $Z(k+1) \supset H(Z(k))$ and $Z(k+1) \subset Z(k)$ for $k > 3$. Hence, the synchronous convergence and the box conditions are satisfied. Asynchronous power updates in Eq. (3.13) converge to the unique fixed point.

3.6.3 Smoothing an Admission Control Type Power Control

Note that an admission control type power control scheme where transmission is stopped when the objective is not reached corresponds to an $f_d^i(.)$ as follows

$$f_d^i(x) = \begin{cases} 0, & x > B_i \\ 1, & otherwise. \end{cases}$$

where $B_i = \frac{G_i P_{max}}{SIR^{obj}} - \sigma^2$. With such f_d^i, there might not be a fixed point because of the discontinuity. In fact, $\phi_i(x_i)$ would be a sawtooth function. However, a good candidate would be least AS function [42] of the sawtooth (Fig. (3.3)). It serves as a smooth approximation to a function with abrupt changes.

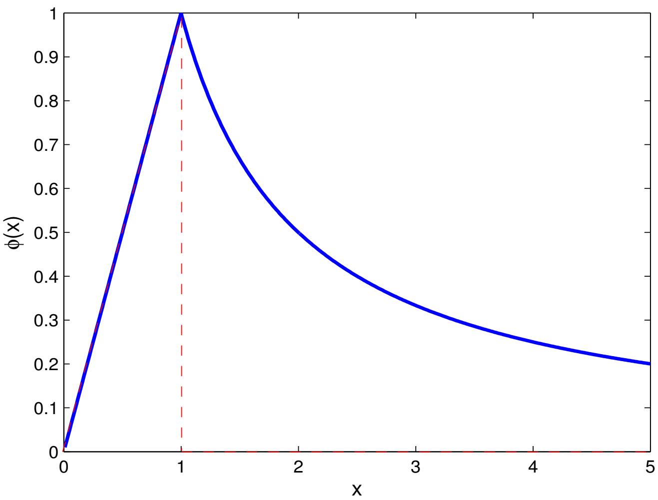

Figure 3.3: Least AS approximation for sawtooth function. $P_{max} = 1$

After some algebra, we can show that the function $f_d(.)$ that corresponds to least AS of sawtooth is expressed as

$$f_d^i(x) = \begin{cases} 1, & x < B_i \\ (\frac{B_i}{x})^2, & otherwise \end{cases} \tag{3.14}$$

With above definition of f_d^i, we now consider the game Ω_2

$$\Omega_2 = [< K >, \{S_i\}, \{u_i^m\}]$$

Ω_2 has a unique NE by construction. In fact, the modification of the utility function brought about an improvement in users' total utility when the original objectives are not feasible. They are no longer supposed to transmit with maximum power in such cases.

In Fig. 3.4, reaction curves in the simple two transmitter-receiver interference channel are shown for two different infeasible cases. When users hit the power limit without reaching the objective, they prefer to decrease the objective with the rule $f_d^i(.)$. In NE, at least one user attains objective. Note that NE is not affected by the modification if the outage objectives are feasible. Users that could not reach their objectives in the equilibrium of original game decrease objectives smoothly while others do not change their objectives and still satisfy them.

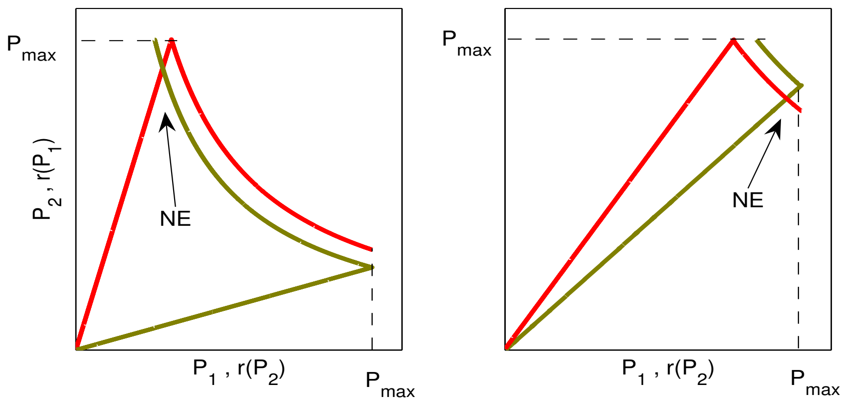

Figure 3.4: Possible reaction curves for different channel gains in a two transmitter-receiver system. Linear portion of reaction curves represent the set of power allocations in which the user attains the SIR objective. Note that the objectives are infeasible for both cases.

3.7 A NUMERICAL EXAMPLE

Actions of Utility 1 and Utility 2 will be best illustrated in comparison with Algorithm 1, Algorithm 2 and SRA [75](in the following, we call SRA Algorithm 3). To make this example as plain as possible, we use a simple setting containing only 5 transmitter-receiver pairs.

In practice, channel gains depend on many factors such as distance, antenna gains, processing gains and shadowing. Since we do not focus on a specific scenario, channel gains will be chosen arbitrarily, with direct link gains G_{ii} being considerably higher than G_{ij} $i \neq j$. In particular, we have a system of two clusters. Users 1, 2 and 3 form

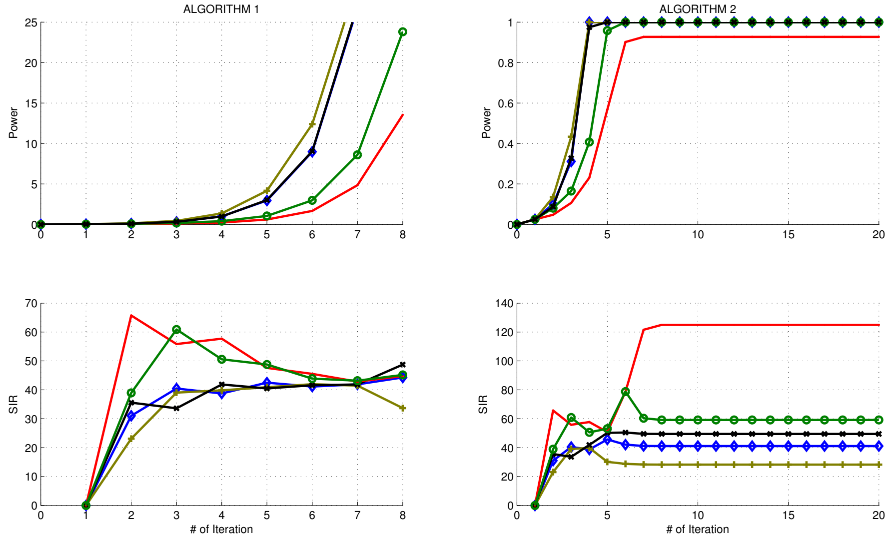

Figure 3.5: Evolutions of power and SIR for algorithms 1 and 2. Note that algorithm 1 in (3.1) diverges in just a few iterations. Algorithm 2 has an equilibrium due to finite power limit. Yet, it is inefficient as users transmit with powers nearly equal to P_{max}

the first cluster; users 4 and 5 the other. We take G_{ij} $i \neq j$ as uniformly distributed independent random variables between 0 and 0.1 within each cluster. Yet, cross gains between members of different clusters are chosen as independent uniform in the range 0 and 0.01. We let direct gains be $G_{ii} = 5$ $\forall i \in< K >$. For compatibility, noise power is $\sigma^2 = 0.001$.

We base our choice of outage threshold on reported results in [20]. Specifically, a threshold of 9 dB indicates an acceptable reliability (instantaneous bit error rate of $\approx 10^{-6}$ in AWGN). Hence, we choose $\gamma^{th} = 9$ dB. The power limit is $P_{max} = 1$. Outage probability objective is $O^{obj} = 5\%$ and consequently $SIR^{obj} \approx 21$ dB or in normal scale $SIR^{obj} \approx 125$. Note that the value of SIR^{obj} is too large for AWGN channel but it is required to guarantee an average communication quality in a Rayleigh-Rayleigh environment.

In the simulated system, Perron-Frobenius eigenvalue of matrix A is calculated as $\lambda_{PF} = 0.0239$, i.e. $SIR^{obj} > \frac{1}{\lambda_{PF}}$ and thus the system is not feasible.

Figures 3.5 and 3.6 show the evolutions of power and SIR values of five users when Algorithm 1, Algorithm 2, Algorithm 3 and MNPC are performed respectively. $f_d(.)^i$ is taken as in (3.14) for NMPC, and $L = 5$ for Algorithm 3 meaning that removal takes

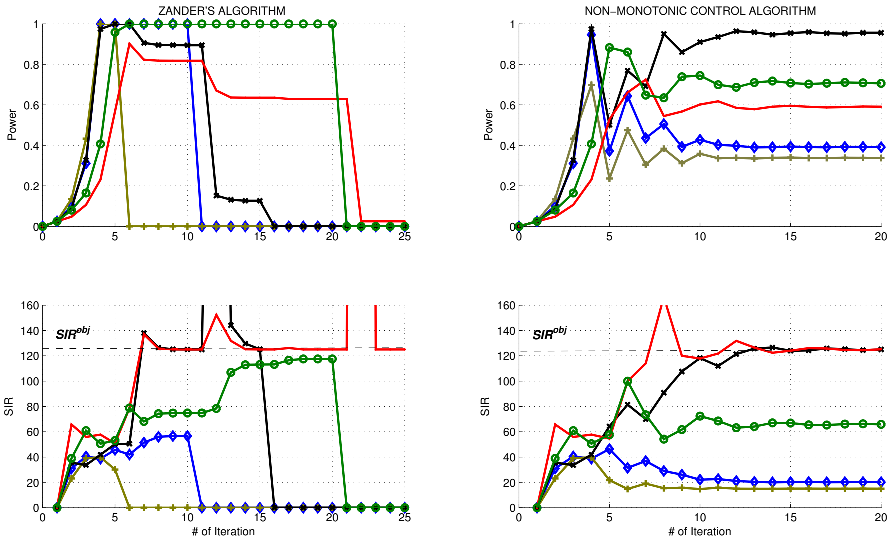

Figure 3.6: Evolutions of power and SIR values for Algorithm 3 and NMPC. Algorithm 3 removes users 2,1,3 and 5 respectively so that the objective is feasible. In NMPC, users 3 and 4 reach the objective while the others operate in a lower outage. Zig-zags are due to non-monotonic reaction curves of users.

place in every five iterations. It should be noted that Algorithm 1 and Algorithm 3 in [75] rely on infinite power assumption. However, we restrict Algorithm 3 for $P_{max} = 1$. Note that Algorithm 2 is the finite power version of Algorithm 1.

It is observed from Fig. 3.5 that two users (users 3 *and* 4) achieve the objective in NMPC while only one user (user 3) can hardly achieve the objective in Algorithm 2. In contrast, SIR values obtained by users 1 and 2 is lower in NMPC. Users 3 and 4 achieves 5% outage probability while user 5 has 10%, user 1 has 30% and user 2 has 38% outage probability in equilibrium of NMPC though they originally aimed at 5%.

NMPC converges to such an equilibrium that number of users achieving the objective increases or remains same compared to Algorithm 2. With its minimizing effect on infeasibilities, NMPC resembles Algorithm 3 in [75] (Fig 3.6). Algorithm 3 simply calculates the sum of rows of matrix A in (3.7) and at each step, the user that has the highest sum of rows is removed if there is an infeasibility. In fact, if objectives are feasible, both Algorithm 3 and NMPC operate in the minimum total transmit power vector.

NMPC boosts the use of spectrum by allocating resources of the system in favor of the user who has a potentially better channel, yet it does not remove any user. Rather, the transfer of resources takes place smoothly so that every user can survive. However, it is observed that NMPC acts on users much the same way as Algorithm 3, if direct channel qualities show great variations from one user to another. That is, NMPC *nearly removes* the user with *too* bad a channel. Intuitively, the equilibrium is in the tail of reaction curve of a bad conditioned user (Fig. 3.4). On the other hand, since Algorithm 3 relies on feasibility, NMPC may perform better than Algorithm 3 when channel gains are on the same order as in the example. It is observed that user 3 is removed by Algorithm 3 yet it achieves the SIR objective in NMPC. From SIR values in equilibrium, we see that NMPC transfers the resources from users 1 and 2 to users 3, 4 and 5.

We also note that power consumptions of users in NMPC decrease compared to algorithm 2. Algorithm 3 leads to the lowest total power among four algorithms. However, it causes a poor use of available spectrum unlike NMPC.

3.8 CONCLUSIONS

In this chapter, we presented a distributed power control mechanism based on outage objectives using non-monotonic and continuous reaction curves. A utility function that captures outage objective along with power efficiency is suggested. We investigated the feasibility of objectives and existence of NE. The inefficiency due to infeasible cases leads to a NE on the boundary of action set. In order to mitigate the inefficiency, we proposed to smoothly decrease objectives as a reaction to interference. Consequently, we obtained a new distributed algorithm, namely non-monotonic power control (NMPC), that overcomes the inefficiency by smoothly transferring resources from advantaged users to the disadvantaged ones. In the equilibrium, number of users that achieve the objective generally increases. The smoothness of this transfer is determined by $f_d(.)$ function. Hence, one possible future direction for this work may be modeling level of fairness with parameters on $f_d(.)$. One other important issue that our work leaves open is the convergence. Convergence result for NMPC is for synchronous updates. A natural extension would be asynchronous and stochastic power update schemes based on the same utility. Moreover, the deeper question to be addressed in future work is optimization of the approach in a fading channel.

CHAPTER 4

COOPERATIVE DISTRIBUTED POWER CONTROL WITH SMOOTH REDUCTION OF OBJECTIVES

4.1 INTRODUCTION

The material in this chapter is submitted and is now under review [46]. The work in this chapter has been performed as a part of the project [66].

The number of users that can be supported in a wireless network in a given bandwidth depends highly on the rate expectations of these users, their power constraints, and channel gains. In an interference network, depending on the number of users in the network, the rate expectations of users can sometimes be unrealistic. And yet, these users may prefer to lower their expectations rather than be provided no service. If the network has a mechanism through which the transmit powers of users (therefore the interference on each other) can be controlled, it may be possible to reach operating points where the number of users that are supported at a satisfactory quality of service is maximized.

There is a large body of literature on power control mechanisms. Among them, utility-based power control algorithms play an important role, as a natural method to generate a distributed algorithm. A utility based power control algorithm basically relies on users egoistically adjusting their own transmit power to maximize a “net utility”. Results of such behavior can be analyzed in depth using equilibrium concepts of economic models. It is interesting to see that the game can be set up such that the *egoistic* behavior leads to what is in effect *cooperation*.

Pricing schemes have been popular methods of cooperation in utility based control [3, 53, 70]. However, the prices are not appropriate for a fully distributed system since the coefficients of prices are determined by another agent (a center). This point is addressed in [22], where the pricing coefficients are also determined distributedly to reach the socially optimum in an ad-hoc wireless network. Actually, prices are not suitable for explaining the cooperation mechanism as it has economic implications and it necessitates price collector(s) who benefit(s) from the collected amount. In [70], it is observed that the cooperative effect of pricing is due to reduction of target SIR value as the interference increases. Hence, pricing turns out to be an implicit method to drive users lessen their rate objective as interference boosts. Assuming that users in the network voluntarily cooperate, users reduce their objectives as a reaction to interference. In this chapter, we attempt to analyze the problem directly in terms of reactions to interference.

Smoothness of power control is important when the network wants to ensure that each node stays in the system. This smoothness can be maintained by absolutely subhomogeneous [42] reactions. We can design utility functions that result in smooth reaction curves. We will analyze a game in which reaction curves are smooth. We first provide conditions for a unique Nash equilibrium. Then assuming that distributed nodes use gradient based optimization, convergence of gradient based iterative algorithms are analyzed. Lastly, the continuous time counterpart of the problem is considered and a stability condition is established for the system.

4.1.1 Related Work

Noncooperative power control has been studied in a multitude of previous work [3, 19, 53, 57] with an emphasis on mechanism design. A common point of them is the pricing as a punishment mechanism to alleviate inefficiency due to non-cooperative behaviour. SIR objective reduction implication of prices are first observed in [70] and admission control as a result of these prices are dwelled upon. Starting from the cooperation induced by objective reductions, Ozel *et.al.* [43] proposed a smooth power control scheme that allows every user to stay in the system.

A more control theoretic approach is presented in [2] where Alpcan *et.al.* analyzes a power control game based on outage probabilities in multicell wireless networks. Defining the utility as a concave function of non-outage probability, stability conditions of associated dynamical system and convergence conditions for gradient based iterative power update algorithms are developed. The main goal of this chapter is to investigate the smooth cooperative mechanism of [43] from the control theoretic perspective of [2].

The following is the organization of the rest of the chapter: Next section introduces the system model. In §4.3, specific form of utility function is given and the role of objective reduction factor is explained. In §4.4, existence and uniqueness conditions for NE are given. A gradient based distributed power control algorithm in which each user uses its own utility function is analyzed in §4.5 and the analysis of the same algorithm, carried to continuous domain is given in §4.6. Numerical illustrations are provided in §4.7. The chapter is concluded with an emphasis on main points in §4.8.

4.2 SYSTEM MODEL

In this chapter, we adopt mainly the same model as in Chapter 3. As once more depicted in Fig. 4.1, we have an ad-hoc wireless network with a number of links $\{i = 1, \ldots, K\}$, each link corresponding to a distinct transmitter. In the following, the terms *transmitter*, *user* and *link* will be used interchangeably. The set of link indices will be denoted by $< K >$.

As before, *Rayleigh-Rayleigh* [26] environment is assumed, i.e., all signal and interference terms are subject to Rayleigh fading. *Single user decoders* are used in the receivers and the signal to interference ratio (SIR) of user i is given as:

$$\gamma_i = \frac{G_{ii} h_{ii} P_i}{\sigma^2 + \sum_{j \neq i} G_{ji} h_{ji} P_j} \tag{4.1}$$

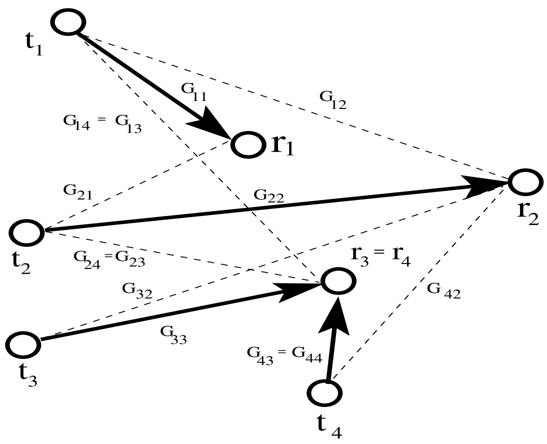

Figure 4.1: The ad-hoc wireless network model: A set of interfering links

Without loss of generality, we consider that each user is subject to the power constraint:

$$P_{min} \leq P_i \leq P_{max} \ \forall i$$

The upper bound on power models physical limitations while the lower limit can be considered as a variable system parameter (which will be discussed later in the chapter).

Average SIR with respect to all h_{ji} is SIR_i^{avg}:

$$SIR_i^{avg} \triangleq \frac{G_{ii}P_i}{I_i}$$

I_i is the average interference experienced by user i

$$I_i = \sum_{j \neq i} G_{ji}P_j + \sigma^2$$

The cross channel gains and power values determine the range of interference $I_{min}^i \leq I_i \leq I_{max}^i$ where $I_{min}^i = \sigma^2 + \sum_{j \neq i} G_{ji}P_{min}$ and $I_{max}^i = \sigma^2 + \sum_{j \neq i} G_{ji}P_{max}$. Link i has interference suppression S_i, defined by

$$S_i = \frac{G_{ii}}{\sum_{j \neq i} G_{ji}}$$

O_i is the probability that a wireless link experiences outage,

$$O_i \triangleq Pr(\gamma_i < \gamma^{th})$$

where γ^{th} is the minimum SIR required for communication. In Rayleigh-Rayleigh environment, the upper bound for O_i is given here once more:

$$\Pr(\gamma_i < \gamma^{th}) \leq 1 - exp(-\frac{\gamma^{th}}{SIR_i^{avg}}) \tag{4.2}$$

In a wireless data network, users will have a certain tolerance to outage events: through link or transport layer control mechanisms such as ARQ, they can ask for repetition of lost packets, and maintain sufficient quality of service even though the channel is not completely reliable. If utility is the eventual quality of service, one could define it in terms of the required outage probability tolerance, as a function of average transmit power. In the next section, the utility function will be made precise.

4.3 UTILTY FUNCTION AND OBJECTIVE REDUCTION FACTOR

In game-theoretic terms, a utility function is a mapping from the action space into real numbers $u_i : \prod_{j=1}^{K} A_j \rightarrow \mathbf{R}$. In our distributed power control algorithms, the action space will contain power levels used. The actual utility of a user may depend on the type of application as well as constraints such as energy. Packet success probability (or a function thereof) has been a popular choice of utility function, considering an ARQ-type link layer scheme and energy efficiency as the main concern [10, 53]. This motivates a step function or a sharply rising sigmoid of SIR as a model of utility for constant-rate real-time applications.

Following our work reported in [43], the derivation of our utility function will be based on each user having an outage probability objective, O^{obj} that it wants to stay below. For simplicity, we assume that this target outage probability is the same for all users. Utility is determined by the relative value of non-outage probability $1-O_i$ with respect to the objective $1 - O^{obj}$. Given an outage objective, we can find SIR^{obj} such that $SIR_i^{avg} \geq SIR^{obj}$ ensures that the outage objective is reached [47]. From (3.4), we choose SIR^{obj} such that $exp^{-\frac{\gamma^{th}}{SIR^{obj}}} = 1 - O^{obj}$. We start by the rather simplistic utility given below:

$$u_i = -\frac{SIR^{obj}}{SIR_i^{avg}} \tag{4.3}$$

It can be deduced by a quick investigation that the game in which users *non-cooperatively* maximize the utility in Eq. (4.3) has an inefficient NE: Each user transmits with power P_{max} in NE. Hence, a more sophisticated utility function that leads to a cooperative mechanism is required.

In [43], we proposed that node i cooperates by reducing its SIR objective with a factor f_d^i. User i's reaction to interfering power is shaped by $f_d^i(.)$. The factor $f_d^i(.)$ is basically a non-increasing continuous function defined from non-negative real numbers into the range $[0, 1]$. An example function $f_d^i(.)$ is illustrated in Fig.4.2. In general, we assume that f_d^i is first order differentiable except at some points, the set of which has finite cardinality.

The concept of objective reduction factor will be central to the development of the utility function here too, albeit with a difference: here, a unit price p_i is subtracted from utility to obtain a net utility. It will be observed later in the chapter although users reduce their objectives as interference escalates, inefficiency in NE is not alleviated unless this price term exists. The net utility function NU_i is the following:

$$NU_i(\mathbf{P}) = -\frac{SIR^{obj} f_d^i(I_i)}{SIR_i^{avg}} - p_i \tag{4.4}$$

Our net utility function employs a linear price, combined with an explicit objective reduction factor whereby SIR objectives are effectively reduced in response to increasing interference. This is akin to the approach of [3, 70] where, reduction of objectives is not explicit, but adaptation to interference is made within a price coefficient. In fact, if net utility were defined as below, the resulting equilibria would be identical:

$$NU_i = -\frac{SIR^{obj}}{SIR_i^{avg}} - \frac{1}{f_d^i(I_i)} p_i \tag{4.5}$$

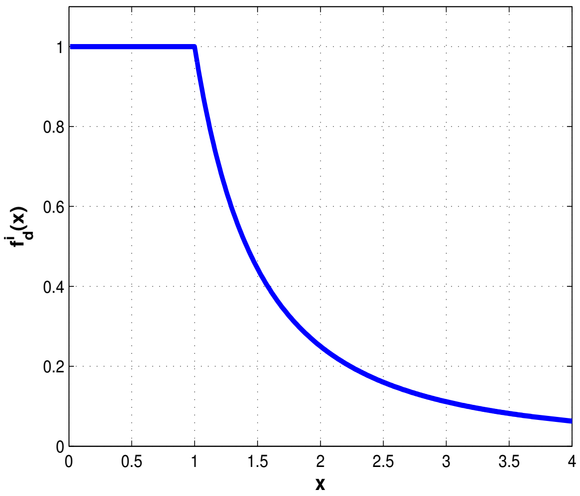

Figure 4.2: An example of objective reduction factor f_d^i

This point will be made clearer during the analysis of reaction curves in the next section. We will devise a distributed mechanism by setting up the power control problem as a non-cooperative game and analyze the properties of the resulting equilibrium. In particular, selection of f_d^i functions is key to the equilibrium and dynamics of the system. In the analysis, the following function devised from $f_d^i(.)$ plays an important role

$$H^i(x) = xf_d^i(x)$$

$H^i(x)$ is continuous and piecewise continuously differentiable. Optimum values of $H^i(x)$ will be used in the analysis, hence the following definitions are made: $H^i_{min} = \min_{x\in[I^i_{min},I^i_{max}]} H^i(x)$, $H^i_{max} = \max_{x\in[I^i_{min},I^i_{max}]} H^i(x)$, $H_{max} = \max_i H^i_{max}$ and $H_{min} = \min_i H^i_{min}$.

In our analysis, the selection of f_d^i will be such that the functions $[H_i(x)]^{1/2}$ $\forall i$ are *absolutely subhomogeneous*, which further implies that $[xf_d^i(ax)]^{1/2}$ for all positive a are *absolutely subhomogeneous*. *Absolute subhomogeneity* constitutes the basis for unique equilibrium. We will next define *absolute subhomogeneity* and point important implications of this property.

4.3.1 Absolutely Subhomogeneous Functions

In [42], Nuzman proposed *absolute subhomogeneity* for power control algorithms using a new contraction approach and a formal definition of absolute subhomogeneity is given in definition 3.6.1 of Chp. 3. Absolutely subhomogeneous (AS) functions involve an important sense of smoothness. The main characteristic of AS functions is *small rate of change.* If the function increases, the rate of increase is less than linear and if it decreases, its absolute rate of decrease is smaller than reciprocal of linear.

A collection of results in [42] about AS functions are presented in the next theorem.

Theorem 4.3.1 *Let $\boldsymbol{P}(t+1) = F(\boldsymbol{P}(t))$ be MSS and let $H(\boldsymbol{P}) = \phi(F(\boldsymbol{P}))$. If vector function ϕ is in the form $\phi(x_1, x_2, ..., x_K) = [\phi_1(x_1), \phi_2(x_2), ..., \phi_K(x_K)]$ with $\phi_i(x_i)$ absolutely subhomogeneous, then $\boldsymbol{P}(t+1) = H(\boldsymbol{P}(t))$ has a unique fixed point.*

Theorem 4.3.1 suggests that if an AS function is applied to each component of an MSS update algorithm, then there exists a unique fixed point of the new algorithm. This result justifies the choice of f_d^i such that H_i are AS (see [43]).

The definition of AS functions does not impose differentiability and even continuity. Yet, continuity will be main assumption in our analysis. We present an implication of AS property under continuity in the next lemma. The result provides a useful bound.

Lemma 4.3.2 *Let $g : \Re^+ \to \Re^+$ be a bounded continuous absolutely subhomogeneous function and let $r(x) = g^n(x)$ for some positive integer n. Subdifferential of r, $\partial r(x)$ satisfies:*

$$\partial r(x) \subset [-\frac{r(x)}{x}, \frac{r(x)}{x}]$$

Proof. Assume that $g(.)$ is differentiable at point x with a positive derivative and let $x + h = e^a x$. For $h < 0$, $a < 0$ and for $h > 0$, $a > 0$. Using the right inequality of AS definition, we have,

$$\frac{dr(x)}{dx} \leq \lim_{a \to 0} \frac{(e^{n|a|} - 1)r(x)}{(e^a - 1)x}$$

Hence, provided a positive derivative,

$$\frac{dr(x)}{dx} \leq \frac{r(x)}{x}$$

If we had a negative derivative, we can use a similar argument and left inequality of AS definition to conclude that

$$\frac{dr(x)}{dx} \geq -\frac{r(x)}{x}$$

The results straightforwardly extend for points where the derivative is undefined which can only be due to inequality of left and right limits. We use the definition of subdifferential in terms of left and right derivatives and a similar treatment of the problem yields

$$\partial r(x) \subset [-\frac{r(x)}{x}, \frac{r(x)}{x}]$$

■

4.4 EXISTENCE AND UNIQUENESS OF NASH EQUILIBRIUM

By assumed properties on reduction factor $f_d^i(.)$, the net utility function $NU_i(\mathbf{P})$ in Eq. (4.4) is continuous and piecewise continuously differentiable with respect to its arguments. In particular, the derivative of $NU_i(.)$ with respect to P_i is continuous while there may be discontinuities at some breaking points in the derivative of NU_i with respect to P_j, $j \neq i$.

The game is $\Omega = [U, \{S_i\}, \{NU_i\}]$ where $S_i = [P_{min}, P_{max}]$ and each user's optimization is

$$\max_{P_i \in S_i} NU_i(P_i, \mathbf{P}_{-i}) \tag{4.6}$$

By direct evaluation of the second derivative, we have

$$\frac{\partial^2 NU_i}{\partial P_i^2} = -2\frac{SIR^{obj} I_i f_d(I_i)}{G_{ii} P_i^3}$$

and except some points at which derivative is not defined, we have

$$\frac{\partial^2 NU_i}{\partial P_i P_j} = \frac{SIR^{obj} G_{ji}}{G_{ii} P_i^2}[I_i f_d'(I_i) + f_d(I_i)]$$

Note that NU_i is strictly concave with respect to its action variable P_i as $\frac{\partial^2 NU_i}{\partial P_i^2} < 0$. Hence, there exists a solution to each user's optimization. Moreover, the Cartesian product of action sets form a non-empty, compact and convex set, which guarantees existence of a NE (due to Theorem 4.4 p.176 in [6]).

NE can be defined in terms of reaction curves. Reaction curve (or best response strategy) of user i, $r_i(\mathbf{P}_{-i})$ is defined as

$$r_i(\mathbf{P}_{-i}) = \max_{P_i \in [P_{min}, P_{max}]} NU_i(P_i, \mathbf{P}_{-i})$$

NE point $\mathbf{P}^*$ is such a point that $\mathbf{P}_i^* = r_i(\mathbf{P}_{-i}^*)$ for all i. The explicit expression of reaction curve for our particular problem is presented in the next theorem:

Theorem 4.4.1 *For the net utility function defined in Eq. (4.4), the reaction curve is*

$$r_i(\boldsymbol{P}_{-i}) = min\{P_{max}, max\{P_{min}, [\frac{SIR^{obj}}{G_{ii}} I_i f_d(I_i)]^{1/2}\}\}$$

Proof. Given $\mathbf{P}_{-i}$, maximum of $NU_i(P_i, \mathbf{P}_{-i})$ occurs either at boundary points $P_i = P_{min}$, $P_i = P_{max}$ or at the point where the derivative wrt P_i is zero. Direct evaluation of derivative of NU_i yields:

$$\frac{\partial NU_i}{\partial P_i} = \frac{SIR^{obj} I_i f_d(I_i)}{G_{ii} P_i^2} - 1$$

$\widehat{P_i}$ that satisfies $\frac{\partial NU_i}{\partial P_i} = 0$ is

$$\widehat{P_i} = [\frac{SIR^{obj}}{G_{ii}} I_i f_d(I_i)]^{1/2}$$

Hence,

$$\max_{P_i \in [P_{min}, P_{max}]} NU_i(P_i, \mathbf{P}_{-i}) = min\{P_{max}, max\{P_{min}, \widehat{P_i}\}\}$$

■

Note that the reaction curve would be the same if the utility function were defined as in Eq. (4.5). The reaction curve in above theorem can be expressed in terms of $H^i(x)$ as follows:

$$r_i(\mathbf{P}_{-i}) = min\{P_{max}, max\{P_{min}, [\frac{SIR^{obj}}{G_{ii}} H^i(x)]^{1/2}\}\}$$

In order to guarantee an inner NE, $\frac{\partial NU_i}{\partial P_i} > 0$ at $P_i = P_{min}$ and $\frac{\partial NU_i}{\partial P_i} < 0$ at $P_i = P_{max}$. As NE is the intersection of reaction curves, these necessary boundary conditions can be expressed in terms of reaction curves. In order not to have an equilibrium on the boundary, we must have

$$P_{min} < r_i(\mathbf{P}_{-i}) < P_{max}$$

Putting the value,

$$P_{min} < [\frac{SIR^{obj}}{G_{ii}} I_i f_d(I_i)]^{1/2} < P_{max}$$

The following two boundary conditions are assumed to hold,

Assumption 1

$$\frac{SIR^{obj} H^i_{min}}{G_{ii} P^2_{min}} > 1 \quad and \quad \frac{SIR^{obj} H^i_{max}}{G_{ii} P^2_{max}} < 1$$

We state our findings about NE of the game Ω in the following theorem

Theorem 4.4.2 *If boundary conditions in assumption 1 hold and if $[H_i(x)]^{1/2}$ are AS $\forall i$, then the game Ω has a unique inner NE.*

Proof. The existence of NE is guaranteed. Uniqueness result follows from Theorem 4.3.1. It is well known that $P_i(t+1) = eI_i(t) =: F_i(\mathbf{P}(t))$ where e is a positive constant, is an MSS update algorithm. After a few algebraic manipulations, one can show that the reaction curve in Th. 4.4.1 can be expressed as

$$r_i(\mathbf{P}_{-i}) = min\{P_{max}, max\{P_{min}, \widehat{P}_i\}\}$$

where $\widehat{P}_i = vH_i(\frac{F_i(\mathbf{P}(t))}{c})$ for some $v > 0$. Finally, AS functions are closed under clipping operations (see Lemma 1 in [43]). Boundary conditions in assumption 1 assure that NE is not on the boundary. ■

4.5 ITERATIVE POWER UPDATES AND CONVERGENCE

In this section, we will investigate certain distributed power iterations of users that capture important properties of real applications. In particular, users are assumed to use a gradient based algorithm to optimize the net utility. We will first investigate convergence of synchronous updates and establish sufficient conditions for convergence. Then we will show that the same conditions are also sufficient for asynchronous updates. Lastly, we will make an error sensitivity analysis and investigate convergence of corresponding random number sequences.

Users update their power with step size Δ as follows:

$$P_i(n+1) = P_i(n) + \Delta \frac{\partial NU_i(\mathbf{P})}{\partial P_i} \tag{4.7}$$

We consider the update algorithm $I(\mathbf{P})$ with $I_i(\mathbf{P}(n)) = P_i(n+1)$. In our analysis, we will follow the framework in [2].

4.5.1 Analysis of Synchronous Updates

Let *unique* NE of the game be $\mathbf{P}^*$. A set of functions for analysis, $c_i(.) : [0,1] \rightarrow \Re$ are defined as follows:

$$c_i(\tau) \triangleq \tau P_i + (1-\tau)P_i^* + \Delta\phi_i(\tau\mathbf{P} + (1-\tau)\mathbf{P}^*) \tag{4.8}$$

where

$$\phi_i(\mathbf{P}) = \frac{\partial NU_i(\mathbf{P})}{\partial P_i}$$

We will reach a sufficient condition so that the algorithm has the contraction property under max norm.

$$\|I(\mathbf{P}) - I(\mathbf{P}^*)\| \leq \rho \, \|\mathbf{P} - \mathbf{P}^*\|$$

We have the following bound

$$|P_i - P_i^*| = |c_i(1) - c_i(0)| = |\int_0^1 \frac{\mathrm{d}c_i(\tau)}{\mathrm{d}\tau} d\tau| \leq \int_0^1 |\frac{\mathrm{d}c_i(\tau)}{\mathrm{d}\tau}\mathrm{d}\tau| \leq \max_{\tau\in[0,1]} |\frac{\mathrm{d}c_i(\tau)}{\mathrm{d}\tau}|$$

Hence, in order to reach the condition, we will bound $|\frac{\mathrm{d}c_i(\tau)}{\mathrm{d}\tau}|$.

$$|\frac{dc_i(\tau)}{d\tau}| = |(1 + \Delta\frac{\partial\phi_i}{\partial P_i})(P_i - P_i^*) + \Delta\sum_{j\neq i}\frac{\partial\phi_i}{\partial P_j}(P_j - P_j^*)| \tag{4.9}$$

By triangle inequality,

$$|\frac{dc_i(\tau)}{d\tau}| \leq |(1+\Delta\frac{\partial\phi_i}{\partial P_i})||P_i - P_i^*| + \Delta\sum_{j\neq i}|\frac{\partial\phi_i}{\partial P_j}||P_j - P_j^*| \tag{4.10}$$

which, in turn implies

$$|\frac{dc_i(\tau)}{d\tau}| \leq (1+\Delta(\frac{\partial\phi_i}{\partial P_i} + \sum_{j\neq i}|\frac{\partial\phi_i}{\partial P_j}|))\,\|\mathbf{P}-\mathbf{P}^*\| \tag{4.11}$$

provided that

$$\Delta|\frac{\partial\phi_i(\mathbf{P})}{\partial P_i}| < 1$$

If $\frac{\partial\phi_i}{\partial P_i}+\sum_{j\neq i}|\frac{\partial\phi_i}{\partial P_j}| < 0$, then the contraction property is guaranteed. We use Lemma 4.3.2 to bound $|\frac{\partial\phi_i}{\partial P_j}|$ for $j \neq i$ and eliminate common terms. The following condition is sufficient for $\rho < 1$

$$\frac{\sum_{j\neq i} G_{ji}}{2} < \frac{H^i_{min}}{P_{max}}$$

provided that

$$\Delta < \frac{\min_i G_{ii} P^3_{min}}{2SIR^{obj} H_{max}}$$

Note that we can obtain a bound on the stepsize Δ by considering the boundary conditions in assumption 1. We state our findings in the following theorem.

Theorem 4.5.1 *The synchronous power update algorithm in Eq. (4.7) converges to the unique NE* $\boldsymbol{P}^*$ *if*

$$\frac{H^i_{min}}{I^i_{max}} > \frac{1}{2}\ \forall i$$

and

$$\Delta < \frac{\min_i G_{ii} P^3_{min}}{2\max_i G_{ii} P^2_{max}}$$

4.5.2 Extension to Asynchronous Updates

The convergence conditions of Bertsekas and Tsitsiklis [7] for totally asynchronous updates, stated in the Th. 3.6.4 in Chp. 3, will be the main guide to extend the convergence result of previous section for asynchronous updates. In Theorem 3.6.4, the first condition is called *Synchronous Convergence Condition* and the second one is called *Box Condition*. Since the conditions are based on contraction property in max norm, asynchronous version of power updates in Eq. (4.7) also converges to unique inner NE, $\mathbf{P}^*$.

Let $\delta(t) := \|I(\mathbf{P}(t)) - \mathbf{P}^*\|$ for $t \geq 1$. Define $S_i(t) = [\mathbf{P}_i^* - \delta(t), \mathbf{P}_i^* + \delta(t)]$. Then the Cartesian product of $S_i(t)$

$$S(t) = S_1(t) \times ... \times S_K(t)$$

$S(t)$ satisfy the synchronous convergence condition by definition of max norm and due to contraction property. Hence, under the same conditions given in Th. 4.7, totally asynchronous version of gradient based power updates in Eq. (4.7) converge to the unique inner NE by Theorem 3.6.4.

4.5.3 Error Sensitivity and Conditions for Almost Sure Convergence

In the above analysis, we assumed that users have perfect information before processing and they can calculate the derivative perfectly. But, errors are involved before and during processing in a real life application of the algorithm. In order to illustrate how the algorithm would operate in a lossy environment, we assume that the error terms are coalesced into a single random variable as a linear factor to the update algorithm as follows:

$$\widetilde{P}_i(n+1) = \widetilde{P}_i(n) + \zeta_i(n)\Delta\frac{\partial NU_i(\widetilde{\mathbf{P}})}{\partial P_i} =: I_i(\widetilde{\mathbf{P}}, \zeta_i) \qquad (4.12)$$

In the notation of the update algorithm, $\widetilde{P}_i(n)$ is the random process representing i^{th} user's power level and $\zeta_i(n)$ is the error random process. $\zeta_i(n)$ are assumed *i.i.d* uniform in the interval $[1-\xi, 1+\xi]$ for all n and i. We will follow a similar analysis to the one presented for synchronous updates. In this case, the norm is

$$\|\mathbf{x}\| = \max_i E(|x_i|)$$

where $\mathbf{x}$ is the vector of random variables and $E(.)$ is the expected value. Using the bound obtained for deterministic P_i and ζ_i values, we have the following:

$$E(\left|I_i(\widetilde{\mathbf{P}}, \zeta_i) - \mathbf{P}_i^*\right|) \leq E\left(\left|1 + \zeta_i \Delta \frac{\partial \phi_i}{\partial P_i}\right| |\widetilde{P}_i - P_i^*| + \zeta_i \Delta \sum_{j \neq i} \left|\frac{\partial \phi_i}{\partial P_j}\right| |\widetilde{P}_j - P_j^*|\right)$$

Assume that $\Delta(1+\xi)B_i^1 < 1$ where B_i^1 is a lower bound on $\left|\frac{\partial \phi_i}{\partial P_i}\right|$ in $[P_{min}, P_{max}]$. Independence of $\zeta_i(n)$ for all n implies independence of $x_i(n)$ and $\zeta_i(n)$ for all n, which in turn follows:

$$E(|I_i(\mathbf{P}, \zeta_i) - \mathbf{P}_i^*|) \leq \left(1 + E(\zeta_i)\Delta(-B_i^1 + (K-1)B_i^2)\right) \|\mathbf{P} - \mathbf{P}^*\|$$

where B_i^2 is an upper bound on $\left|\frac{\partial \phi_i}{\partial P_j}\right|$ $\forall j \neq i$. In order to satisfy contraction property,

$$-B_i^1 + (K-1)B_i^2 < 0$$

One can show that the bounds can be selected as

$$B_i^1 = \frac{2SIR^{obj} H^i_{min}}{G_{ii} P_{max}^3}$$

$$B_i^2 = \frac{SIR^{obj} \max_j G_{ji}}{G_{ii} P_{min}^2}$$

Following similar steps to derive a sufficient condition as in the case for synchronous updates, and eliminating common terms in B_i^1 and B_i^2, a sufficient condition for contraction property is:

$$\frac{P_{max}^3}{P_{min}^2} < \frac{2H^i_{min}}{(K-1)\max_j G_{ji}} \ \forall i$$

Random power updates defined in Eq. (4.12) converge to $\mathbf{P}^*$ *almost surely* under derived conditions. In order to show this, we will make use of a similar technique as in [2]. Using Markov inequality first and then definition of max norm, we have:

$$\begin{aligned} \sum_{t=1}^{\infty} Pr(|P_k(t) - \mathbf{P}_k^*| > \epsilon) \leq \sum_{t=1}^{\infty} \frac{E(|P_k(t) - \mathbf{P}_k^*|)}{\epsilon} \quad &\leq \quad \frac{1}{\epsilon} \sum_{t=1}^{\infty} \|\mathbf{P}(t) - \mathbf{P}^*\| \\ &\leq \quad \frac{1}{\epsilon} \sum_{t=1}^{\infty} \rho^t \|\mathbf{P}(0) - \mathbf{P}^*\| \\ &\leq \quad \frac{\|\mathbf{P}(0) - \mathbf{P}^*\|}{\epsilon(1-\rho)} \end{aligned}$$

for arbirtary $\epsilon > 0$ and $\|\mathbf{P}(0) - \mathbf{P}^*\|$ is a non-negative constant. $Pr(A)$ represents the probability of event A and the last inequality follows from contraction property (that is guaranteed by the conditions). The sequence of partial sums is bounded above and thus convergent. From Borel-Cantelli lemma, the random update scheme in Eq. (4.12) converges a.s. to the unique NE.

We collect overall results of the analysis in the following theorem:

Theorem 4.5.2 *The stochastic power updates in Eq. (4.12) converges to the unique NE $\boldsymbol{P}^*$ if the following two conditions hold:*

$$\frac{P_{max}^3}{P_{min}^2} < \frac{2H_{min}^i}{(K-1)\max_j G_{ji}} \ \forall i$$

$$\Delta < \frac{\min_i G_{ii} P_{min}^3}{2(1+\xi)\max_i G_{ii} P_{max}^2}$$

4.6 POWER UPDATES AS A CONTINUOUS TIME DYNAMICAL SYSTEM

Consider now the iterative power updates of previous section with small slot duration and step size Δ tending to zero. In this case, we obtain continuous time counterpart of unilateral utility maximization of users in Eq. (4.6). Mathematically, we now have a dynamical system that is governed by the following differential equations:

$$\dot{P_i} = \frac{\partial NU_i}{\partial P_i} := \phi_i(\mathbf{P}) \ \forall i \tag{4.13}$$

We will investigate the stability of the system in Eq. (4.13). In particular, the notion of stability will be *global asymptotic stability in the sense of Lyapunov* [28]. NE is the candidate to be a stable point of the system as it is a solution of the nonlinear system $\frac{\partial NU_i}{\partial P_i} = 0 \ \forall i$ (as an inner NE is guaranteed). We will find sufficient condition(s) on the stability of NE in defined dynamical system .

It should be noted that $\phi_i(\mathbf{P})$ is differentiable almost everywhere and thus it is Lipschitz continuous [28] although it is not differentiable at a set of points where $(f_d^i)'(.)$

is possibly discontinuous. We assume $(f_d^i)'(.)$ is not continuous at finite number of points $\{d_{ji}\}$ for all i. As the surfaces $I_i = d_{ji}$ form a set of Lebesgue measure zero in $[P_{min}, P_{max}]^K$, $\phi_i(\mathbf{P})$ is differentiable almost everywhere. Hence, it is legitimate to apply Lyapunov Theory.

Boundary conditions in assumption 1 immediately guarantee that the power values $P_i(t)$ evolve in such a way that they bounce back to inside of $[P_{min}, P_{max}]$ whenever $P_i(t) = P_{min}$ or $P_i(t) = P_{max}$. Hence, the power levels of users always reside inside the set $[P_{min}, P_{max}]$.

A proper Lyapunov function is to be guessed now in order to establish the sufficient conditions. A candidate Lyapunov function $V : \Re^K \to \Re$ for the system in Eq. (4.13) is as follows

$$V(\mathbf{P}) = \sum_{i=1}^{K} \phi_i^2(\mathbf{P}) \tag{4.14}$$

Note that $V(\mathbf{P})$ is actually defined in $[P_{min}, P_{max}]^M$. Since NE is inner and unique, $\phi_i(\mathbf{P}) = 0 \ \forall i$ if and only if $\mathbf{P} = \mathbf{P}^*$. Hence, $V(\mathbf{P})$ is positive for all $\mathbf{P}$ other than NE. The following condition is sufficient for stability of NE, if satisfied at points other than NE and whenever the derivative is well defined,

$$\dot{V}(\mathbf{P}) < 0$$

We begin the analysis by calculating $\dot{V}(\mathbf{P})$ and then bounding it

$$\dot{V}(\mathbf{P}) = \sum_{i=1}^{K} -2\frac{SIR^{obj}}{G_{ii}P_i^2}\alpha_i\phi_i^2 + \sum_{i=1}^{K}\sum_{j\neq i} 2\frac{SIR^{obj}}{G_{ii}P_i^2}\beta_{ij}\phi_i\phi_j$$

where

$$\alpha_i := \frac{2f_d^i(I_i)I_i}{P_i}$$

$$\beta_{ij} := G_{ij}[I_i\frac{\mathrm{d}f_d^i(I_i)}{\mathrm{d}I_i} + f_d^i(I_i)]$$

Using the identities $\beta_{ij} \leq |\beta_{ij}|$ and $\phi_i^2 + \phi_j^2 - 2\phi_i\phi_j \geq 0$, and lemma 4.3.2, we have

$$\dot{V}(\mathbf{P}) \leq \left(\sum_{m=1}^{K} \frac{SIR^{obj}}{G_{mm}P_m^2}\right)(-2\min_k \alpha_k + K\max_{ij}|\beta_{ij}|)\sum_{i=1}^{K}\phi_i^2$$

We reach the following result.

Theorem 4.6.1 *A sufficient condition for global asymptotic stability of the system in (4.13) is*

$$\frac{H_{min}}{P_{max}\max_{i,j\neq i} G_{ji}} > \frac{K}{4}$$

Note that the interference characteristic plays a key role in the stability condition of the system. Besides, it is inconclusive about the stability of the system if the condition in Th. 4.6.1 is not satisfied.

4.7 NUMERICAL STUDY

In this section, we will provide examples to make the operation of the algorithms and resultant equilibria clearer to the reader. We will first explain the grounds for selection of the specific objective reduction factor used in the examples. Then, numerical results will be shown.

4.7.1 Selection of Objective Reduction Factor

The functional shape of objective reduction factor f_d^i reflects user's attitude towards utilizing network resources. If the application running on the user requires constant performance throughout the operation, then f_d^i is set to 1 for all interference levels, and this reflects a rather non-cooperative behaviour. On the other hand, the user behaves in cooperation with other users of the network if it decreases its objective as the interference blows up.

Various user behaviour can be addressed by defining different f_d^i. In our work, we assume that f_d^i has a specific generic form with piecewise definitions in three regions. In particular, users are allowed to be non-cooperative up to a certain interference level L_1^i. After this level, users cooperate by means of decreasing their objectives. The rate of decrease is bound to satisfy absolute subhomogeneity of $[xf_d^i(x)]^{1/2}$. An additional constraint comes from boundary conditions. One can show that the condition on

H^i_{max} turns to a condition on the non-cooperative interference region

$$L^i_1 < \frac{G_{ii}P^2_{max}}{SIR^{obj}}$$

User i reacts to interference level higher than L^i_1 by decreasing its objective. The decreasing region of f^i_d is constrained by absolute subhomogeneity assumption of $[xf^i_d(x)]^{1/2}$. An optimal way of cooperation under AS assumption can be described using least AS upper bound [42] of the reaction that decreases to zero sharply at interference level L_i. One can show that the decreasing region of least AS upper bound of sharp decrease is proportional to $1/x^3$. The proportionality constant is determined such that the function is continuous.

A further constraint on the decreasing region is proposed as each user must stay in the system. User's reaction to interference is imposed to be higher than its reaction to the minimum interference it experiences.

$$H^i_{min} = I^i_{min}$$

The region decreasing with proportion to $\frac{1}{x^3}$ stops when the reaction level reaches $H^i(I_i) = I_{min}$. The interference level I_i at which $H^i(I_i) = I_{min}$ is L^i_2. f^i_d decreases proportional to $1/x$ and $H^i(I_i) = I_{min}$ for all $I_i > L^2_i$. The general form of f^i_d is as follows:

$$f^i_d(x) = \begin{cases} 1, & I^i_{min} < x < L^i_1 \\ (\frac{L^i_1}{x})^3, & L^i_1 < x < L^i_2 \\ \frac{(L^i_1)^3}{(L^i_2)^2}\frac{1}{x}, & L^i_2 < x < I^i_{max}. \end{cases} \tag{4.15}$$

We allow the non-cooperative region to be as large as possible, hence $L^i_1 = \frac{G_{ii}P^2_{max}}{SIR^{obj}}$. Note that L^i_1 are set proportional to the direct channel gain G_{ii}. This way, users having higher direct channel gains favor the network resources more and thus throughput efficiency of the network is supported. For continuity, $L^i_2 = \frac{L^i_1}{(I^i_{min})^{1/3}}$. Introduction of the third region in definition of f^i_d brings a notion of fairness to the control algorithm as it maintains users' remaining in the system. Hence, by defining f^i_d as in (4.15), users react to interference with a balance between throughput efficiency and fairness.

There may be exceptional cases that require modification in the general form of f_d^i given above. If $I_{min}^i > L_1^i$, then f_d^i would have only the $\frac{1}{x}$ decreasing region and the reaction would be a constant for all interference levels. In this extreme case, though user i transmits with $P_{min} + \upsilon$, it is not allowed to control its power and is in a sense discarded from the system[1]. Another exception occurs if $I_{min}^i < L_1^i$ but $I_{max}^i < L_2^i$. In this case, $\frac{1}{x}$ decreasing region does not appear. Note that the imposed condition $H_{min}^i = I_{min}^i$ is satisfied in both cases. Possible reaction curves in several different cases are depicted in Fig. 4.3.

Putting $H_{min}^i = I_{min}^i$, the boundary condition becomes $I_{min}^i > \frac{G_{ii}P_{min}^2}{SIR^{obj}}$, which in turn requires

$$P_{min} < \frac{SIR^{obj}}{S_i}$$

Sufficient condition for synchronous convergence in Th. 4.5.1 becomes

$$\frac{P_{max}}{P_{min}} < 2 + \frac{2\sigma^2}{\sum_{j \neq i} G_{ji} P_{min}}$$

Similarly, global asymptotic stability condition in Th. 4.5.1 turns to

$$\frac{P_{max}}{P_{min}} < 4 \frac{\min_i \sum_{j \neq i} G_{ji}}{K \max_{i, j \neq i} G_{ji}}$$

4.7.2 Simulation Setting and Results

We will simulate an ad-hoc network with 5 transmitter-receiver pairs. Direct and cross channel gains vary with Rayleigh fading. As the system model can represent a wide range of communication scenarios, we do not specify physical models about link gains and multiple access schemes used. Keeping the generality of the model, the average channel gains G_{ij} are determined arbitrarily with $G_{ii} \gg G_{ij}$ for all i.

The threshold SINR is set to $\gamma^{th} = 9$ dB for all experiments. The particular value of the threshold is chosen since it guarantees a certain communication quality[2]. Users'

[1] Note that $0 < \upsilon \ll P_{min}$. This makes sure that the equilibrium does not lie on the boundary

[2] 9 dB threshold ensures BER less than 10^{-6} for QPSK modulation and BER is in acceptable levels for higher order modulations. See [20].

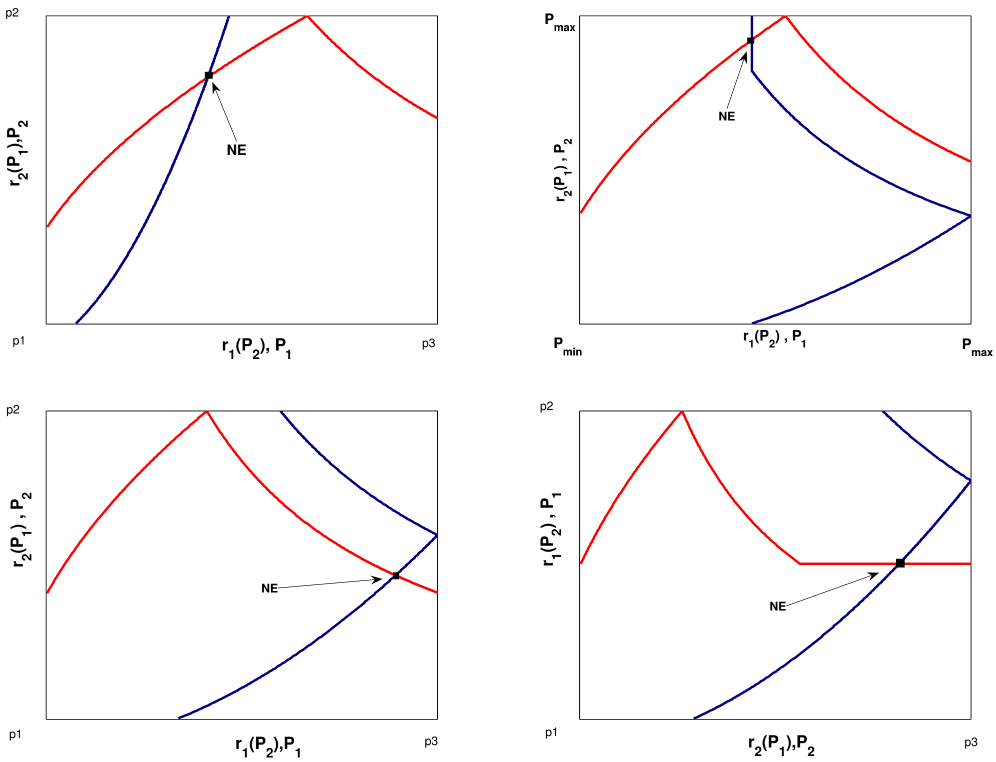

Figure 4.3: Possible Reaction Curves

goals are to stay below an outage probability of $O^{obj} = 10\%$, which corresponds to $SIR^{obj} = 18.7$ dB or $SIR^{obj} = 75$ in normal scale. As in previous chapter, the objective SIR level is too high compared to the SIR levels needed in AWGN channel. The reasons of this are the assumed time variations in the channel and the fact that users aim to have an average communication quality in the varying environment.[3]

In the simulated system, users are capable of a suppression level of around $S_i = 15$ dB. Note that the exact values of channel gains and noise power are not important in terms of operation. G_{ii} takes arbitrary value in the interval $[1, 2]$ for all i and exact values of G_{ij} are chosen randomly (say distributed uniformly in a relatively small interval than its mean value) so that around 15 dB suppression is maintained. Throughout the simulations, we set maximum transmit power as $P_{max} = 1$ and noise

[3] Average SIR required to satisfy outage objective depends on the statistics of fading. Under Ricean and Nakagami fading models, required average SIR turns out to be less if judiciously compared to Rayleigh fading. See [48].

power as $\sigma^2 = 0.001$.

In the simulated setting, $\max_i \sum_{j \neq i} G_{ji} = 0.034$, $\min_i \sum_{j \neq i} G_{ji} = 0.024$, $\max_{ij} G_{ji} = 0.01$, $\max_i G_{ii} = 1.5$ and $\min_i G_{ii} = 1$. Interference suppression capability S_i ranges from 30 to 60. We fix $P_{max} = 1$ for all experiments.

In the first experiment, we investigate the NE by means of convergence of iterating reaction curves. Perron-Frobenius eigenvalue of the system is $\lambda_{PF} = 0.0154$, which implies that $SIR^{obj} = 75$ is not feasible. Actually, users can achieve a common SIR level of at most $\frac{1}{\lambda_{PF}} = 65$ in normal scale. Transmitters optimize their utilities in each iteration given the *perfect* information of other users' power actions. In particular, users have instantaneous access to the *perfect* information of congestion level at the intended receiver and process this information to adjust its power level. Minimum power is taken as $P_{min} = 0.3$. Evolutions of power level of transmitters and achieved SIR level in their corresponding receivers are shown in Fig. 4.4. Convergence is observed in about 15 iterations.

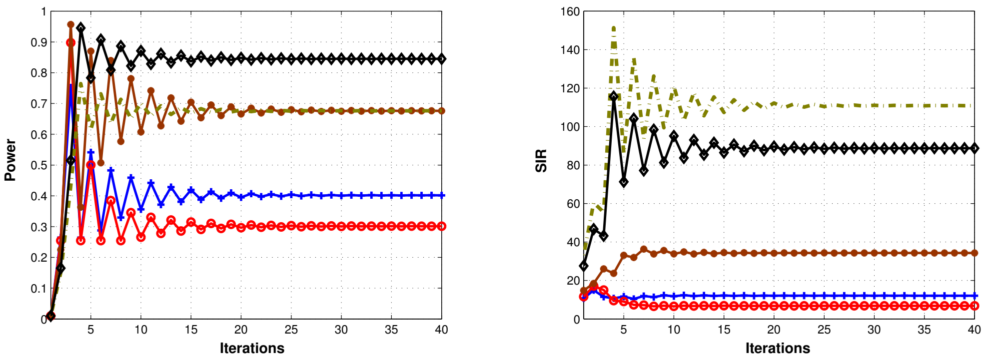

Figure 4.4: Evolutions of power and SIR are illustrated for $P_{min} = 0.3$. Users have perfect knowledge of congestion level at their intended receiver and synchronously react by changing their power levels.

In the next experiment, synchronous iterative power update algorithm in Eq. (4.7) is simulated. Perron-Frobenius eigenvalue of the system is $\lambda_{PF} = 0.0227$. $\frac{1}{\lambda_{PF}} = 65$ in normal scale and $SIR^{obj} = 75$ is not feasible. Choices of P_{min} and step size Δ determine the convergence of the update algorithm. In particular, sufficient conditions in Th. 4.5.1 and boundary conditions reveal an interval for P_{min}:

$$\frac{1}{2} - \frac{\sigma^2}{\sum_{j \neq i} G_{ji}} < P_{min} < \frac{SIR^{obj}}{S_i}$$

It should be noted that interfering channel gains play an important role in the choice of P_{min} while direct channel gains and objective SIR^{obj} do not appear in the constraints. Minimum power is taken as $P_{min} = 0.5$. The step size is chosen as $\Delta = 0.05$ so that (along with the chosen value for P_{min}) the synchronous convergence conditions given in Th. 4.5.1 are satisfied. Evolutions of power and SIR are illustrated in Fig. 4.5. Convergence is observed in about 20 steps.

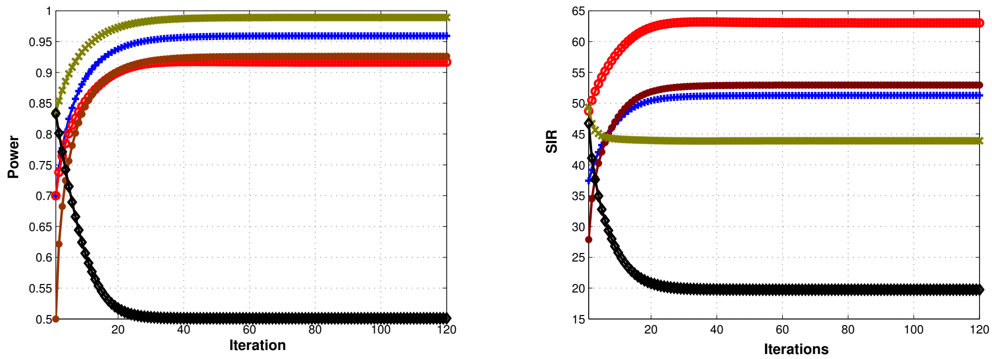

Figure 4.5: Evolutions of Power and SIR in synchronous update algorithm of Eq. (4.7) with $P_{min} = 0.5$ and $\Delta = 0.05$.

Now, in the same system as in previous experiment, we change the minimum power to $P_{min} = 0.3$ and $\Delta = 0.01$, in which case P_{min} violates the convergence condition in Th. 4.5.1. However, the update algorithm is observed to converge in Fig. 4.6. Convergence is relatively slow in this case due to smaller step size. It can be concluded that the convergence results in Th. 4.5.1 are not necessary conditions though they provide a useful guideline for selection of the parameters. Moreover, it turns out that the conditions obtained for step size are necessary for invariance of iterative power update algorithms.

Finally, we address error sensitivity of the synchronous power update scheme. System is not feasible as Perron-Frobenius eigenvalue of the system is $\lambda_{PF} = 0.0207$. $\frac{1}{\lambda_{PF}} = 49$ in normal scale. $P_{min} = 0.5$, $\Delta = 0.05$ and $\xi = 0.5$ are the parameters taken for the

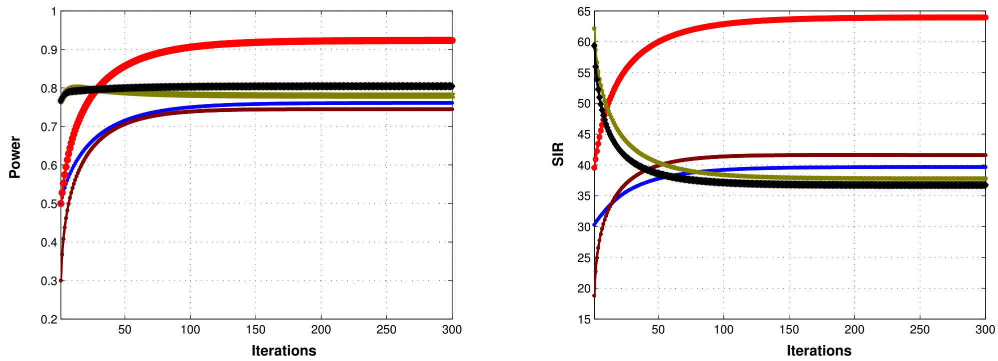

Figure 4.6: Evolutions of Power and SIR in synchronous update algorithm of Eq. (4.7) with $P_{min} = 0.3$ and $\Delta = 0.01$.

simulation. Note that the step size Δ satisfies the convergence condition in Th. 4.5.2, but P_{min} does not satisfy the respective sufficient condition. Yet, in this case too, the erroneous updates converge, as can be observed in Fig. 4.7.

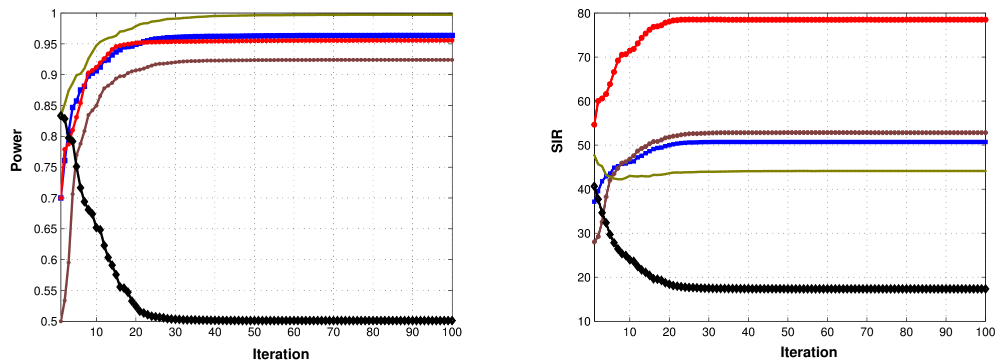

Figure 4.7: Evolutions of Power and SIR in synchronous update algorithm of Eq. (4.12) with erroneous feedback information. $P_{min} = 0.5$, $\xi = 0.5$ and $\Delta = 0.05$.

4.8 CONCLUSION

In this chapter, we presented a game theoretic analysis of utility based distributed power control with smooth reactions to interference. The utility function has been defined as a measure of relative performance with respect to rate objectives. Users decrease their rate objectives as a reaction to interference, exhibited in the definition of net utility by means of an objective reduction factor. This is inherently an egoistic cooperation mechanism which serves to improve overall network performance. Each user's reaction to interference is assumed to be absolutely subhomogeneous so that a unique Nash equilibrium of the game is guaranteed. Iterative power update algorithms based on gradient optimization techniques have been investigated and convergence conditions have been derived. Finally, power updates are considered as the result of a continuous time dynamic system, which is a limiting case for small step size. Sufficient conditions for stability of the system have been established. Numerical results verify the sufficiency of the convergence conditions of iterative algorithms, while cases are shown exhibiting these conditions are not necessary for convergence.

CHAPTER 5

A RESOURCE ALLOCATION APPROACH TO BUFFER SHARING

5.1 INTRODUCTION

The work in this chapter has been performed as a part of the project [66].

Memory is a limited resource in communication devices. Although the communication, computation and memory capabilities continuously increase, with the advance of 3G and broadband wireless MAN, there is also a substantial increase in the demand for bandwidth and memory. For example a typical WiMax base station serves a metropolitan area, where hundreds of users are demanding high speed multimedia applications. Therefore buffer management schemes have to be devised in order to get maximum performance with limited memory space. Besides, even if there is unlimited buffer space, limiting the space used for a communication session can be used in order to limit the delay.

Sharing the limited buffer resources among multiple communication sessions is a problem that previously attracted interest in the context of shared-memory switches [23] and wireline networks [25]. There are two main and opposite methods of buffer management, which are complete sharing and complete partitioning. In complete sharing (CS), the complete memory pool can be used by all sessions. This provides a degree of flexibility and increases the utilization factor of the buffer. On the other hand, a high-rate session can completely occupy the memory space and cause dropping of some other low-rate sessions with strict delay requirements. Even if there are no strict delay requirements, this "hogging" of the buffer by a subset of the sessions will have detrimental effects on total throughput, and in a wireless system, limit the benefit from *multiuser diversity*.

The hogging caused by the CS policy, i.e. buffer being occupied by only a subset of the sessions, is therefore detrimental in terms of multiuser diversity, and hence limits overall rate much below system capacity. The opposite extreme, Complete Partitioning (CP) of the buffer, where each session is given equal amount of dedicated memory space, could potentially solve this problem though it performs poorly for low loads. To find a middle point in between CP and CS, it is possible to determine the dedicated memory space according to the arrival rates and service capacities of users. To illustrate, a larger space can be dedicated for a user with high arrival rate for same service condition or smaller space is given to the user with lower service rate for identical arrival conditions. This provides a QoS for the delay constrained users and also reveals a novel trade-off between buffer utilization and multiuser diversity.

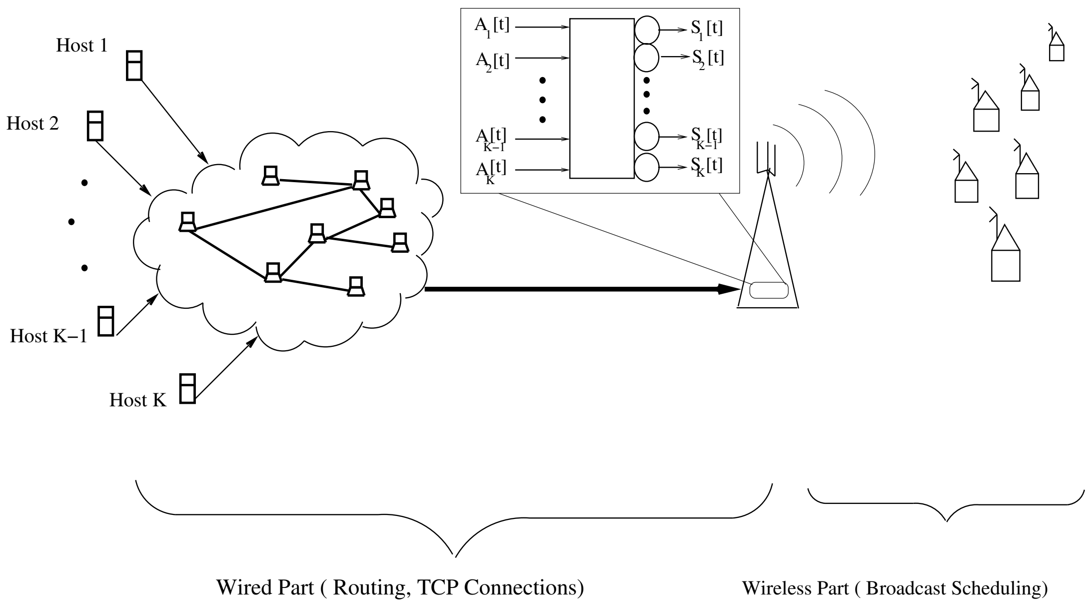

Figure 5.1: A network with a wireless last mile: Multiple sessions sharing a finite buffer space are taking scheduled service in a wireless downlink.

Complete partitioning has clearer advantage over sharing under unbalanced and high load [23]. Indeed high load is a very interesting case to consider. A multiuser downlink may work in the overloaded regime for several reasons. Such a system typically serves various uncoordinated users, as in fixed wireless [50] Internet access, as well as in cellular systems. It is to be expected that sessions initiated by various user applications do not have correct estimates of the bandwidth, that is to say transmission rate, available to them, as the total number of sessions is dynamic, as well as the channel itself. Under such uncertainties, operating close to instability may be preferable to occasionally idling and not fully utilizing the tight wireless resource, as consequent packet drops may be tolerated by higher-layer mechanisms (such as TCP). That is, perhaps the unstable regime is a practical reality in wireless systems.

Buffer partitioning would both give individual throughput guarantees to low rate users and also exploit multiuser diversity by keeping waiting rooms for individual sessions. Throughput maximization has been addressed in a large body of literature: from sum-rate maximizing schemes [31] to proportional-fair schedulers [68], and utility-based rate allocations [37], there is a plethora of results about achieving various degrees of complexity and fairness. It is well known that stability or 100% throughput requires taking full advantage of server capacity and minimizing idleness. In the specific case of a wireless node, the server is the wireless channel, and achieving its full capacity region requires exploiting multiuser diversity, i.e., benefiting from the increasing likelihood of a user having a very good channel condition as the number of users increases [64]. This of course requires there being a sufficient number of packets in the queues at any time [17]. In an infinite-buffer system, there are well-known mechanisms for achieving 100% throughput, for example by a maximum-weight-matching (MaxWeight) between queue states and channel states at any given time.

Throughput maximization problem in a finite shared buffer scenario requires handling the problems of buffer management and scheduling jointly. Optimal buffer management and scheduling in terms of throughput in the finite-buffer case is still an open problem, and the suboptimality of MaxWeight was established in [55] for the same problem in NxN switches. It should also be noted that MaxWeight requires making rate allocation decisions based on joint queue and channel state information. We believe that being able to separate the rate allocation problem from the buffer management problem carries practical value, as the former is traditionally in the physical layer and it can be cumbersome to keep physical layer algorithms informed about queue state for especially high speed operation.

Optimal buffer *partitioning* can be applied in order to maximize throughput under high load and limited buffer memory. It may be claimed that optimal buffer management is obsolete, since higher layer mechanisms adjust arrival rate so that the system becomes stable. However, this is not completely true because depending on the network structure, response time of TCP can be on the order of seconds [29, 35], which makes order of thousand time slots. Considering the highly varying wireless channel, the system can easily become overloaded between congestion window updates. Optimal buffer management can be used together with higher layer mechanisms in order to better utilize wireless resources. As an example, consider the network in Figure 5.1 where there is a network terminated by a wireless last hop. The buffers at the wireless transmitter will need to have a sufficient number of packets to be able to exploit multiuser diversity and operate at a timescale determined by the rate of the wireless channel. The queue lengths here could be capped at the optimal partitioning levels. The TCP's that work end to end could be responsible for satisfying a long-term rate requirement to ensure that the right number of packets is maintained.

5.1.1 Contributions

We mainly asks two questions in this chapter: (1) Given arrival and service rates, how should we partition a finite buffer among users to maximize total throughput? (2) How close can we get to maximum throughput if we let scheduling be done without regard to queue state, and use optimally partitioned buffer for the resulting service rates? In answer to the first question, an optimal iterative algorithm for allocating buffer space to queues based on their arrival and service rates will be derived. To address the second question, let us describe further the model in consideration. We consider a downlink multiuser system where N independent packet arrival processes are separately queued to be sent by a single transmitter over a wireless channel which can be described as a stationary stochastic process. The service model of course depends on how the data streams are multiplexed to be transmitted. We consider two main channel allocation mechanisms

- Case 1. Fixed channel allocation (e.g. an FDMA system with orthogonal channels, experiencing outage and possibly correlated fading.)
- Case 2. Channel-Aware Dynamic Scheduling (e.g. selecting user(s) with good channel states at each scheduling interval. Subcases based on TDMA, MIMO and OFDMA will be considered.)

Note that in all of the above cases, we can effectively associate a server with each queue, which may take vacations (such as when the corresponding queue is not selected for service, due to its channel condition). These vacations will be considered part of service time of the corresponding head-of-line packet that needs to wait for service during that time. As a result, service times are not independent across users. However, they will be independent of the arrival process and the queue state of the same queue. This is consistent with scheduling mechanisms that are queue state-blind.

5.2 OPTIMAL BUFFER PARTITIONING

In a multiple arrival multiple service finite buffer system, effective buffer management can provide considerably good throughput performance [13, 25]. While optimality would require a dynamic allocation of buffer space among queues in general, static allocation, referred to as *partitioning*, was observed to perform very well (and is perhaps optimal) for unbalanced and high loads [25]. It is shown in [13] that optimal policy for two user balanced high load case is equally partitioning the available buffer for users. We will introduce framework in [13] and discuss implications of the framework for unbalanced load case in the following subsection.

5.2.1 Partitioning for Unbalanced Load

Assume arrival and service processes are Poisson so that multi-dimensional Markov chains are the tool of analysis [13]. In a system of N users with a total pool of B buffers, the feasible set Ψ is

$$\Psi = \left\{ \mathbf{m} = (m_1, m_2, ..., m_N),\ m_i \in N^+ : \sum_{i=1}^{N} m_i \leq B \right\}$$

A policy Ω is defined as a subset of the feasible set. Applying a policy means that the states in that subset are allowed and remaining states in the feasible set are not allowed. Management is done through dropping packets to prevent occurrence of the states outside the policy region.

In the sequel, we assume policies are coordinate convex. By definition, a finite set of integer n-tuples is coordinate convex if $((x_1-1)^+, (x_2-1)^+, ..., (x_n-1)^+)$ is in that set whenever $(x_1, x_2, ..., x_n)$ is in that set. If Ω is coordinate convex, then the packets can not be dropped once they are allowed to wait in the buffer. It is possible to allow push-out type policies in general [9]. However, push-out type policies are hard to implement as it requires extra read-write operations on the memory, which degrades the speed of operation. Hence, we restrict ourselves to the coordinate convex policies.

Optimal coordinate convex policy for N user case is not known [13]. However, some observations about the geometry of the optimal policy can be made using the framework of [13]. Any subset ω of Ψ is removable from Ω if $\Omega|\omega$ is coordinate convex. Similarly, ω' is annexable to Ω if $\Omega \cup \omega'$ is a policy. We will use following two theorems from [13]:

Theorem 5.2.1 *Let ω be removable from Ω and ω' be annexable to Ω. Then*

1. $id(\omega) <,=,> id(\Omega) \Longleftrightarrow id(\Omega|\omega) >,=,< id(\Omega)$

2. $id(\omega') <,=,> id(\Omega) \iff id(\Omega \cup \omega') <,=,> id(\Omega)$

Here, $id(.)$ stands for idleness probability. But, it is defined in a more general sense. id is a function that takes each subset ω of Ψ to the interval $[0,1]$. Let ρ be defined as the load (arrival rate / service rate) and w_j be defined as set of points in ω that has j^{th} entry zero.

$$id(\omega) = \frac{\sum_{j=1}^{N} \sum_{(i_1,i_2,\ldots,i_N)\in\omega_j} \prod_{k=1}^{N} \rho_k^{i_k}}{\sum_{(i_1,i_2,\ldots,i_N)\in\Omega} \prod_{j=1}^{N} \rho_j^{i_j}}$$

Optimal policy is defined as the policy that minimizes idleness probability. There may be many optimal policies but there is a unique maximal policy, the one which has highest cardinality.

Theorem 5.2.2 *If Ω is maximal, then*

1. *ω removable from Ω implies $id(\omega) \leq id(\Omega)$*
2. *ω' annexable to Ω implies $id(\omega') > id(\Omega)$*

From [13], we have for 2-user case, maximal policy Ω^* is in the following form:

$$\Omega^* = \{(x_1, x_2) \in I^2 | x_1 \leq m_1, x_2 \leq m_2, x_1 + x_2 \leq B\}$$

where $m_1 + m_2 \geq B$ and $I = \{1, 2, ..., B\}$.

We are interested in the evolution of Ω^* as ρ_2 increases and ρ_1 is fixed. Although partitioning may not be optimal in general, as the loads become unbalanced there is a tendency of the optimal policy for partitioning. We will present two observations concerning this issue.

First observation is that m_2 is non-increasing as ρ_2 increases. Let for some fixed ρ_1^*, ρ_2^*, the maximal policy has parameters m_1^*, m_2^* and $\omega' = \{(0, m_2^*+1), (1, m_2^*+1), ..., (B - m_2^* - 1, m_2^*+1)\}$. Obviously, ω' is annexable to Ω^*. We denote $\Omega^*(\rho_1^*, \rho_2)$ as the same region only the loads changed. Using theorem 5.2.2 of Foschini:

$$id(\omega') > id(\Omega^*(\rho_1^*, \rho_2^*)) \tag{5.1}$$

If we attempt to write the explicit expression for $id(\omega')$, we have:

$$id(\omega') = \frac{\rho_1^0 \rho_2^{m_2^*+1}}{\sum_{j=1}^{B-m_2^*-1} \rho_1^j \rho_2^{m_2^*+1}} = \frac{1}{\sum_{j=0}^{B-m_2^*-1} \rho_1^j} \tag{5.2}$$

That is, $id(\omega')$ does not vary with ρ_2. Since, idleness probability decreases as any one of the loads increase, we have for all $\rho_2 > \rho_2^*$

$$id(\omega') > id(\Omega^*(\rho_1^*, \rho_2)) \tag{5.3}$$

Hence, from the two theorems of Foschini, m_2 can not increase as ρ_2 increases.

The second observation is that m_2 decreases after some point as ρ_2 increases. This is due to the fact that idleness probability decreases as load is increased. Now, let $\omega'' = \{(0, m_2^*), (1, m_2^*), ..., (B - m_2^*, m_2^*)\}$. It is clear that ω'' is removable from $\Omega^*(\rho_1^*, \rho_2^*)$. Using the second theorem, we immediately have:

$$id(\omega'') \leq id(\Omega^*) \tag{5.4}$$

Similar to previous analysis, $id(\omega'')$ does not vary with ρ_2. Hence, after some point in ρ_2, $id(\omega'') > id(\Omega^*(\rho_1^*, \rho_2))$, which means that removing ω'' leads to better throughput using again both of the theorems cited from [13].

As conjectured in [13], it is possible to extend the result for multiple user case. Actually, the benefit of buffer partitioning for unbalanced load are discussed in [32, 69] under different data and flow models. Hence, we will resort to buffer partitioning as the management policy. Each user is assigned a portion of the available buffer budget which fits to a resource allocation setting. In the next section, we will formulate buffer partitioning as a resource allocation problem and develop an efficient algorithm to find throughput-optimal buffer partitioning.

5.2.2 Partitioning as a Resource Allocation Problem

In an M/G/1/m system, throughput is a monotone increasing and concave function of both arrival rate and buffer space for a fixed service process. Let $T(\lambda_i, m)$ be the throughput at arrival rate λ_i with waiting room for m packets. Let ΔT_i denote the increase in throughput provided by an increase of one in buffer space in queue i.

$$\Delta T_i(m) = T(\lambda_i, m) - T(\lambda_i, m-1) \tag{5.5}$$

Adding one more waiting room to the system always increases the throughput [4, 76], i.e $\Delta T_i(m) > 0$. Moreover, the concavity implies diminishing returns, i.e $\Delta T_i(m+1) < \Delta T_i(m)\ \forall m$.

In the rest, we use the shorthand $T_i(m)$ to mean $T(\lambda_i, m)$. The buffer allocation that maximizes total throughput is a solution to the following optimization problem:

$$\max \sum_{i=1}^{N} T_i(m_i)\ s.t.\ \mathbf{m} \in \Psi \tag{5.6}$$

Using the monotonicity and concavity of throughput function, the following algorithm finds an optimal buffer allocation.

5.2.3 Algorithm for Maximum Total Throughput

As no user will be denied service in our model[1], we must allocate a buffer space of at least one to each user. To divide up the rest of the buffer space, we shall now propose an iterative mechanism. The server has a total of B rooms available for N users. A pseudo code of the algorithm for reaching an optimal buffer allocation is as follows:

1. Initially, allocate each user one waiting room, i.e. $m_i = 1\ \forall i$
2. Calculate $\Delta T_i(m_i)$ for any i that has not been calculated.
3. For $j \in \arg\max_i\ \Delta T_i(m_i)$, $m_j = m_j + 1$
4. If $\sum_i m_i = B$ then stop, otherwise go to step 2.

The algorithm greedily allocates buffers to users that yield maximum increase at each step. We next discuss optimality and complexity of the algorithm.

[1] Combining buffer allocation with admission or flow control [5] is very interesting yet outside the scope of this thesis work.

5.2.3.1 Optimality

An equivalent problem to the optimization problem stated in Eq. (5.6) was studied in [58] and a similar algorithm was proposed for optimal resource allocation. Optimality can be established using a contradiction argument. Assume that there is an allocation that yields strictly greater sum of throughputs than that found by the algorithm. Note that procedure of the algorithm is equivalent to finding $B-N$ highest numbers among all $N(B-N)$ possible $\Delta T_i(m_i)$ $i = 1, 2, ..., N$ and $\mathbf{m} \in \Psi$. The resultant sum throughput is sum of the $B-N$ highest $\Delta T_i(m_i)$ plus the sum of throughputs for $m_i = 1$. Hence, existence of assumed allocation contradicts with the maximality of $B-N$ numbers selected by the algorithm.

5.2.3.2 Complexity

The first recursion of second step involves $\Theta(N)$ time units. Then $B-N$ recursions take place with a sorting based operation in each recursion. Hence, the time complexity of the algorithm is $\Theta(N + (B-N)logN)$ units. Furthermore, it is argued in [41] that another algorithm with $\Theta(N^2(log(B-N))^2)$ time complexity can be proposed. This algorithm is based on Lagrange multipliers and it may be less complex for $B \gg N$.

Next, we consider the application of optimal buffer allocation to several settings modeling the cases in Section 5.1.

5.3 APPLICATIONS

We start with a simple general setting similar to the one in [25] where queues are assumed $M/M/1/m_k$. Then, we investigate a model of the FDMA downlink, where each frequency band is characterized by an outage probability, corresponding to Case 1 in Section 5.1. We state and solve joint buffer allocation and channel assignment problem. We then turn to a setting where users or groups of users share the channel in time, which models Case 2. This time, the buffer allocation problem is solved for parallel $M/G/1/m_k$ queues.

5.3.1 Parallel Channels as M/M/1/m_k Queues

This is a simple model for users being given separate orthogonal channels, arrivals are Poisson, packet lengths are exponential, and each channel has a (possibly different) constant transmission rate.

Average throughput T and packet drop probability P_d of the $M/M/1/m$ queue [16] are expressed in the following equations

$$T(\lambda, \rho, m) = \lambda(1 - \frac{(1-\rho)\rho^m}{1-\rho^{m+1}}) \tag{5.7}$$

$$P_d(\rho, m) = \frac{(1-\rho)\rho^m}{1-\rho^{m+1}} \tag{5.8}$$

The optimal buffer allocation can yield a considerable increase in the total throughput (see figure 5.2). Note that the percentage increase in the throughput becomes higher as more users share the available buffer space. This is due to monotone decreasing property of $\Delta T_i(m)$.

5.3.2 FDMA with Channel Outage

Consider a frequency division multiple access (FDMA) multi-user downlink. There are N users, and a frequency band will be allocated to each user. Each frequency band exhibits "outage" at random times, that is, the SNR dips below a level that can support the (fixed) code rate being used. On every channel, the outage periods are IID across time, and the starting times of outage events form a renewal process. We also assume that both outage durations and the periods between outage are long enough for hundreds of packet transmissions, such that in each state, the queue reaches

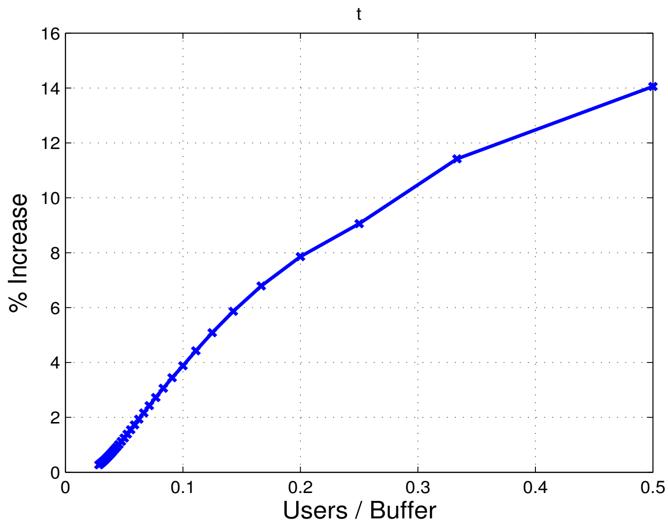

Figure 5.2: The percentage increase in total throughput v.s. users per buffer. Optimal buffer allocation is compared to even buffer allocation in parallel M/M/1/m_i system with a total buffer of $B = 3500$. There are two classes of users. 25% of users belong to first class and the remaining belongs to the second. $\rho_1 = 11\rho_2 = 1.1$. As more users share the buffers, buffer allocation yields higher increase in throughput.

steady-state. The probability of outage depends on the frequency band used, and not on which user is using this channel[2].

Suppose each user is allocated one of the bands in an arbitrary fashion. Since the queue reaches steady-state in both outage and non-outage, each user's queue is an $M/M/1/m_i$ queue during non-outage, and is full (contains exactly m_i packets) during outage. Specifically, let queue j be served in frequency band i, whose outage probability is p_i^{out}. At steady state in outage, the queue is full, so the stationary probability of drop in outage is 1. In the non-outage case the packet drop probability is $P_d(\rho, m)$. The average packet drop probability and the throughput are therefore:

$$P_d^{avg}(\lambda, p^{out}, m) = (1 - p^{out})P_d(\lambda, m) + p^{out} \tag{5.9}$$

$$\begin{aligned} T(\lambda, p^{out}, m) &= \lambda[1 - P_d^{avg}(\lambda, p^{out}, m)] &(5.10)\\ &= (1 - p^{out})\lambda[1 - P_d(\lambda, m)] &(5.11)\end{aligned}$$

[2] The channel statistics not depending on user (and hence receiver location) may correspond, for example, to the case when the receivers are geographically clustered far away from the base station.

The algorithm to find optimal buffer allocation can be applied with a slight modification in this case.

$$\Delta T_i^{out}(m_i, p_i^{out}) = (1 - p_i^{out})\lambda_i[P_d(\lambda_i, m_i) - P_d(\lambda_i, m_i + 1)] \tag{5.12}$$

Under these assumptions, the introduction of outage channel to the problem brings forth a new dimension in terms of optimization: Assigning the channels to users for optimal total throughput. Channels with outage probabilities $p_1, p_2, \ldots, p_N$ are matched to the users in a one-to-one fashion.

Problem 1 *Given λ_i and available channels' outage probabilities p_i, maximize $\sum_i (1 - p_{\pi(i)}) T_i(\lambda_i, m_i)$ subject to $\sum_i m_i = M$ and $m_i \geq 1$ and π is any permutation of $i = 1, 2, ..., N$.*

The problem is an extended version of the resource allocation problem discussed in section 5.2 with a combinatoric dimension. Hence, the resource allocation problem is revisited in more detail. Since the resource parameter is discrete, analysis becomes harder. Thus, we resort to investigate the continuous counterpart of the problem; i.e.buffer parameter m is assumed continuous. Then, we will discuss applicability of the analysis to the discrete problem.

Two monotone positive real functions are *monotone disuniting* if their difference diverges to infinity. Note that monotone functions have well-defined inverse functions. In our analysis, we will use the same idea for inverses and we introduce *monotone inverse disuniting functions.*

Definition 5.3.1 Monotone Inverse Disuniting Functions *The pair of functions f_1 and f_2 are said to be monotone inverse disuniting if*

1. *$f_1 : \Re^+ \rightarrow I_1$ and $f_2 : \Re^+ \rightarrow I_2$, $I_1, I_2 \subset \Re^+$ are monotone increasing with $f_1(x) > f_2(x)$ $\forall x \in \Re^+$.*

2. $\forall y_1, y_2 \in I_1 \cap I_2$, $y_1 > y_2 \Rightarrow$
$(f_2^{-1}(y_1) - f_1^{-1}(y_1)) > (f_2^{-1}(y_2) - f_1^{-1}(y_2))$

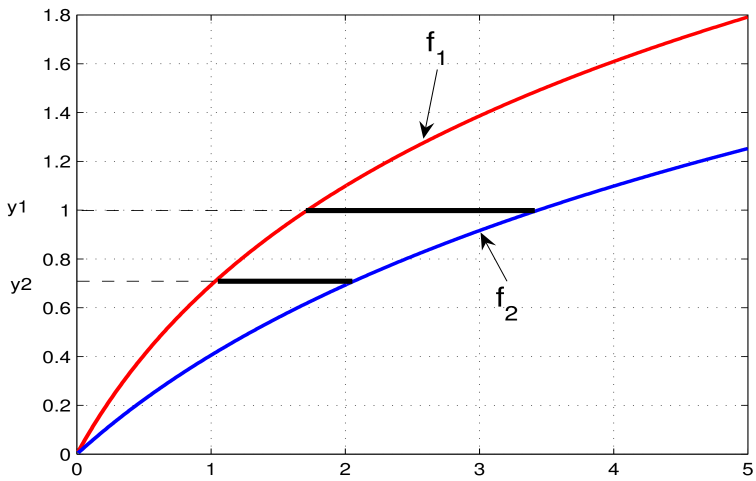

Figure 5.3: Monotone Inverse Disuniting Functions. The difference increases as y is increased

Now, we can state the result on the joint optimization problem:

Theorem 5.3.2 *Let M be a positive constant and*

$$S \triangleq \{(x_1, x_2) : x_1 + x_2 \leq M, x_1 \geq 1, x_2 \geq 1\}$$

If f_1, f_2 are monotone inverse disuniting and $\alpha_1 > \alpha_2 > 0$,

$$\max_{\boldsymbol{x} \in S}\{\alpha_1 f_1(x_1) + \alpha_2 f_2(x_2)\} > \max_{\boldsymbol{x} \in S}\{\alpha_1 f_2(x_2) + \alpha_2 f_1(x_1)\}$$

Proof. Let $Z = \max_{\mathbf{x} \in S}\{\alpha_1 f_2(x_2) + \alpha_2 f_1(x_1)\}$, $\mathbf{x}^* = (x_1^*, x_2^*) = \arg\max_{x \in S}\{\alpha_1 f_2(x_2) + \alpha_2 f_1(x_1)\}$. It is enough to show that there exists some $(x_1^{**}, x_2^{**}) \in \mathcal{S}$ such that $\alpha_1 f_1(x_1^{**}) + \alpha_2 f_2(x_2^{**}) > Z$. To show this, we will consider two cases:

1. Assume $f_1(x_1^*) \geq f_2(x_2^*)$. Then setting $x_1^{**} = x_1^*$ and $x_2^{**} = x_2^*$ and exchanging the channels, $\alpha_1 f_1(x_1^{**}) + \alpha_2 f_2(x_2^{**}) > Z$.

2. Assume now $f_1(x_1^*) < f_2(x_2^*)$. Let's exchange the channels and define $x_1^{***} = f_1^{-1}(f_2(x_2^*))$ and $x_2^{***} = f_2^{-1}(f_1(x_1^*))$. Note that by definition we have $\alpha_1 f_1(x_1^{***}) + \alpha_2 f_2(x_2^{***}) = Z$. The same throughput is achieved with total buffer $X^{***} = f_1^{-1}(f_2(x_2^*)) + f_2^{-1}(f_1(x_1^*))$. In the previous allocation, total buffer was $X^* = x_1^* + x_2^* = f_1^{-1}(f_1(x_1^*)) + f_2^{-1}(f_2(x_2^*))$. Because of the monotone disuniting property (and for $f_1(x_1^*) < f_2(x_2^*)$), we have $f_2^{-1}(f_2(x_2^*)) - f_1^{-1}(f_2(x_2^*)) > f_2^{-1}(f_1(x_1^*)) - f_1^{-1}(f_1(x_1^*))$. After rearranging we get, $f_2^{-1}(f_2(x_2^*)) + f_1^{-1}(f_1(x_1^*)) > f_2^{-1}(f_1(x_1^*)) + f_1^{-1}(f_2(x_2^*))$. This means that $X^{***} < X^*$. The same throughput is achieved with smaller buffer memory. Hence, there exists some allocation $(x_1^{**}, x_2^{**}) \in S$ such that $\alpha_1 f_1(x_1^{**}) + \alpha_2 f_2(x_2^{**}) > Z$.

■

Corollary 1 *For $\alpha_1 > \alpha_2 > ... > \alpha_K > 0$, and (f_i, f_j) $\forall i < j$ are monotone inverse disuniting, permutation π^* that solves the joint optimization problem*

$$\max_{\pi, \boldsymbol{x} \in S} \alpha_{\pi(i)} f_i(x_i)$$

is the identity permutation $\pi^*(i) = i$

Proof. Assume another permutation $\pi'(i) \neq i$ solves the joint optimization problem. There exists at least two indices i_1, i_2 such that $i_1 < i_2$ and $\pi'(i_1) > \pi'(i_2)$ so that $\alpha_{\pi'(i_1)} < \alpha_{\pi'(i_2)}$. If above theorem is applied to these two indices, it is deduced that another permutation π'' with $\pi''(i_1) = \pi'(i_2)$ and $\pi''(i_2) = \pi'(i_1)$ yields better, which is a contradiction. Hence, the identity permutation $\pi^*(i) = i$ yields the joint optimal. ■

Lemma 5.3.3 *For $\lambda_1 > \lambda_2$, let $f_i(m) = T(\lambda_i, m)$ $i = 1, 2$ as in Eqn 5.10. f_1 and f_2 are monotone inverse disuniting with $f_1(m) > f_2(m)$ $\forall m \in \Re^+$.*

Proof. The derivative w.r.t. m is $\frac{-\rho^{m+1}(1-\rho)\ln\rho}{(1-\rho^{m+1})^2}$, which is always positive. The derivative w.r.t. ρ is $\frac{1+m\rho^{m+1}-(m+1)\rho^m}{(1-\rho^{m+1})^2}$, which is also greater than zero (The nominator of the derivative is a convex function with minimum of zero). Therefore the first condition is satisfied.

As for the second condition, after some rearrangement, we get $f_i^{-1}(y) = \frac{\ln\left(\frac{\rho_i - y}{\rho_i(1-y)}\right)}{\ln \rho_i}$. Let's define $F_{21}(y) = f_2^{-1}(y) - f_1^{-1}(y)$.

$$F_{21}(y) = \frac{\ln\left(\frac{\rho_2 - y}{\rho_2(1-y)}\right)}{\ln \rho_2} - \frac{\ln\left(\frac{\rho_1 - y}{\rho_1(1-y)}\right)}{\ln \rho_1} \tag{5.13}$$

$$F'_{21}(y) = (\frac{1}{\rho_1 - y})\frac{1}{\ln \rho_1} - (\frac{1}{1-y})\frac{1}{\ln \rho_1} - (\frac{1}{\rho_2 - y})\frac{1}{\ln \rho_2} + (\frac{1}{1-y})\frac{1}{\ln \rho_2} \tag{5.14}$$

Collecting common terms once more, we get,

$$F'_{21}(y) = \frac{1}{\ln \rho_2}\left(\frac{\rho_2 - 1}{(\rho_2 - y)(1-y)}\right) + \frac{1}{\ln \rho_1}\left(\frac{\rho_1 - 1}{(\rho_1 - y)(1-y)}\right) \tag{5.15}$$

We know that $y < 1$ and $y < \rho_1, \rho_2$, therefore we need to check for the positivity of the terms $\frac{\rho_i - 1}{\ln \rho_i}, i = 1, 2$. For both of the cases $\rho_i > 1$ and $\rho_i < 1$, it is positive therefore the inverse difference function $F_{21}(y)$ is increasing in y. Hence, the pair of functions are monotone inverse disuniting. ■

Theorem 5.3.4 *Suppose* $\lambda_1 > \lambda_2 > ... > \lambda_K$ *and* $p_1^{out} \leq p_2^{out} \leq ... \leq p_K^{out}$. *Optimal channel allocation that solves problem 1 is* $\pi^*(i) = i$.

Proof: The result immediately follows from above theorem and lemma. ■

In our actual problem, the arguments of functions f_1 and f_2 in theorem 5.3.2 are integers. We can still let the functions be defined for positive real numbers but the optimization is performed over integers. Then, the steps in the proof can be applied the same way in general. But there is an exceptional case in which monotone inverse disuniting property may not be sufficient. Let $f_1^{-1}(f_2(x_2^*)) = I_1 + d_1$ and $f_2^{-1}(f_1(x_1^*)) = I_2 + d_2$ such that I_i and d_i for $i = 1, 2$ are integer and fractional parts of the corresponding numbers. If $d_1 < 0.5$, $d_2 > 0.5$, $I_1 + I_2 = B - 1$ and $d_1 + d_2 < 1$, then a resource of amount $1 - (d_1 + d_2)$ is available but integer arguments can not be obtained by adding that amount. So, one has to decrease one of the arguments and increase the other. Adding the remaining *fractional* resource by decreasing one of the arguments and increasing the other may not yield better total throughput.

Note that Theorem 5.3.4 implies a separation between the problems of buffer allocation and channel assignment in this case: The optimal solution is a best-channel highest-arrival rate allocation, i.e., channel assignment is based on arrival rate but not on queue (buffer) state. It is of interest whether this separation can be carried on to more general multiplexers.

5.3.3 User Selection in a Time-Varying Channel

Now, we generalize our service model to cover Case 2 in the Introduction. Here, a user, or a subset of users, is selected at each scheduling time, based on their (combined) achievable rate at that time. Note that the selection (or scheduling) decision does not respect the instantaneous queue state.

For simplicity, we start with analyzing a user scheduler where only a packet of one user is selected at a time in a multiuser wireless downlink. Selection of user is based on channel state and the scheduling is performed at the end of service of a packet. Packet lengths are assumed constant. The achievable rate of any user is drawn from the same distribution, independently. Let this rate be R. The random variable $R \in \{1, 2, \ldots, r_{\max}\}$ is described by some probability mass function $p_R(r)$. We will assume that users have symmetric channels; more explicitly, their channel gains $h_i(t)$ are independent memoryless random processes with the same statistics. Accordingly, if the packet of user i has been selected for service, the service time of a packet will be $1/R_i$ where R_i has the same distribution as R: $R_i \sim R$.

With memoryless arrivals, the individual user packet queues can be viewed as M/G/1/m_i systems. We will use Gelenbe's approximate expression [36] for M/G/1/m_i packet drop probability P_d.

$$P_d(\lambda, \mu, m) = \frac{\lambda(\mu - \lambda)e^{-2\frac{(\mu-\lambda)(m-1)}{\lambda a^2+\mu s^2}}}{\mu^2 - \lambda^2 e^{-2\frac{(\mu-\lambda)(m-1)}{\lambda a^2+\mu s^2}}} \tag{5.16}$$

where $a = \frac{Var(T_a)}{E(T_a)^2}$ and $s = \frac{Var(T_s)}{E(T_s)^2}$. Note that $a = 1$ for Poisson arrivals. Throughput can be expressed in terms of P_d as follows:

$$T(\lambda, \mu, m) = \lambda(1 - P_d(\lambda, \mu, m)) \tag{5.17}$$

It can easily be verified that throughput in (5.17) is monotone increasing with both λ and m. Hence, the incremental buffer allocation algorithm also solves the throughput maximization problem here.

Recall that scheduling is done without regard to queue backlog information. We, now, consider the scheduling that takes queue states into account. It is of interest to ask whether a separation of scheduling and buffer allocation holds here as in the parallel channel case. Though the general problem is too complicated, we attempt to solve the simplest version of the problem next.

Consider the buffer management problem in which service is maintained by user selection over independent channels with two possible states, 0 or 1 (Fig. 5.4). Only one packet of a user is served if that user is selected. Note that when both channels are in state 0, service is not provided for any user. We again resort to judicious buffer partitioning. Assume that channels are in state 0 with probability p_0 (symmetric

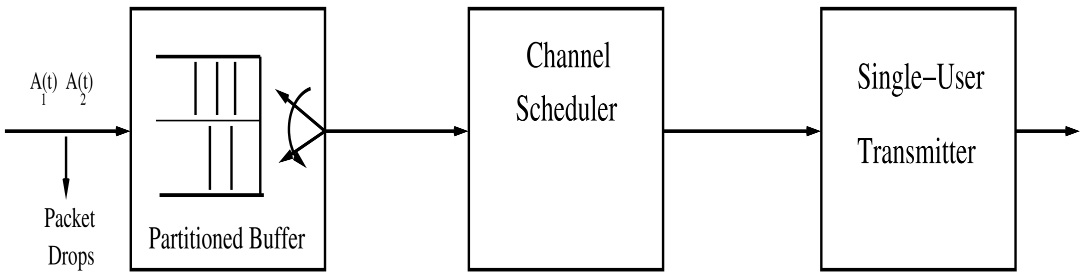

Figure 5.4: The model for two user joint buffer management and user scheduling in a time-varying channel

channels) and the arrivals are Bernoulli with probability (rate) λ_1 and λ_2. Without loss of generality, let $\lambda_1 > \lambda_2$. It is sure that user 1 is scheduled when channel 1 is in state 1 and channel 2 is in state 0 and user 2 is scheduled for the reverse case. However, when both channels are in state 1, scheduler should select one of them. We impose that scheduler selects user 1 with probability a given both channels are in state 1. Accordingly, service probabilities μ_1 and μ_2 are expressed as follows:

$$\mu_1 = p_0(1 - p_0) + a(1 - p_0)^2 \tag{5.18}$$

$$\mu_2 = p_0(1 - p_0) + (1 - a)(1 - p_0)^2 \tag{5.19}$$

Throughput expression is the same as that for $M/M/1/m$:

$$T = \lambda(1 - \frac{(\lambda/\mu)^m(1 - (\lambda/\mu))}{1 - (\lambda/\mu)^{m+1}}) \tag{5.20}$$

Assuming that B buffers are available and treating the buffer m_i as a continuous variable, the joint buffer partitioning and scheduling problem is stated as follows:

$$\max\ T_1(\lambda_1, \mu_1, m_1) + T_2(\lambda_2, \mu_2, m_2) \tag{5.21}$$

subject to

$$m_1 \geq 1, m_2 \geq 1, m_1 + m_2 \leq B, \mu_1 + \mu_2 \leq C \mu_1 \geq c_1, \mu_2 \geq c_2$$

where $C = 2p_0(1 - p_0)$, $c_1 = c_2 = p_0(1 - p_0)$.

Note that throughput in Eq. (5.20) is monotone increasing and concave with respect to both m and μ. The constraint set is a compact and convex region, hence existence of an optimum is guaranteed. Actually, uniqueness of the optimum value is also guaranteed as KKT conditions are satisfied.

A first result about the joint optimization follows for $\lambda_1 = \lambda_2$. In this case, $m_1^* = m_2^* = M/2$ and $a = 0.5$. The exact characterization of the joint optimal seems to be intractable but some trends can be understood via numerical observations. Numerical results show that if $\lambda_1 > \frac{1}{p_0}\lambda_2$, then $a^* = 1$ (see Fig. 5.5). That is, if the average arrival rate of user 1 is higher than a ratio times the average arrival rate of user 2, then joint optimal scheduling is *always* scheduling the higher arrival rate user. For the arrival rates in between $\lambda_2 < \lambda_1 < \frac{1}{p_0}\lambda_2$, $0.5 \leq a^* \leq 1$. Actually, a^* experiences a sharp transition from 0.5 to 1 if available buffer B is relatively small. However, a smooth transition occurs if B is large. See Fig. 5.6.

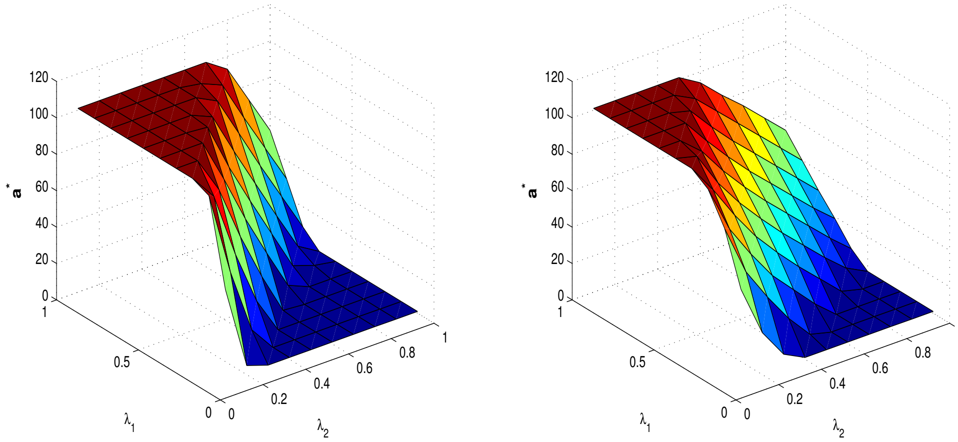

Figure 5.5: The joint optimal scheduling probability a^* in (%) is plotted with respect to (λ_1, λ_2) plane. $B = 20$ buffers per user. $P_0 = 0.6$ and $P_0 = 0.4$ on the left and right plots respectively.

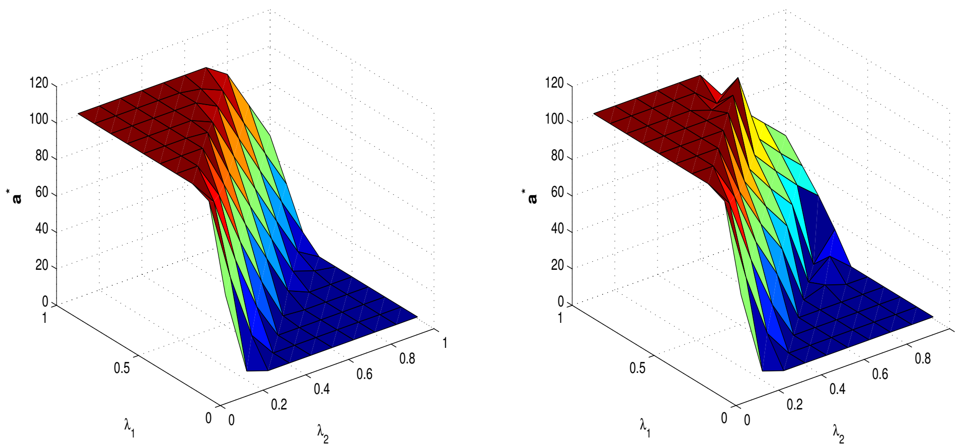

Figure 5.6: The joint optimal scheduling probability a^* in (%) is plotted with respect to (λ_1, λ_2) plane. $P_0 = 20$. $B = 20$ and $B = 5$ on the left and right plots respectively.

5.4 SIMULATIONS

In this section, we will compare throughput performances of several joint buffer management and scheduling policies by means of simulations. In particular, queue aware and queue blind scheduling with and without buffer partitioning will be compared in a multiple state wireless downlink channel.

We address joint user scheduling and buffer management in the simulations. Simulated user scheduling mechanisms are MaxWeight (MW), Max. Channel (MC) and Time Division Multiplexing (TDM). MaxWeight scheduling calculates the product of backlog and rate at each slot and selects the user that has the maximum of the products. Note that MW scheduler relies on cross layer operation of link layer and physical layer as it necessitates instantaneous backlog and channel state information. Max. Channel selects the user that has the best rate. Due to discrete nature of the rates, there may be ties, i.e. best rate can be achieved by more than one user. We assume that the scheduler has the information of arrival rate λ and the user with higher λ is selected in case of ties. This way, MC scheduler does not process the instantaneous backlog information but rather first order statistic of the arrival process. Hence, the cross layer operation in MW is not observed in MC. TDM scheduling, the simplest of the three, is basically the round robin scheduling of users.

As buffer management policies, complete sharing (CS), equal partitioning (EP) and optimal partitioning (OP) schemes are considered. CS policy allows each user to be accommodated if there is an available space. On the other hand, EP reserves equal buffer spaces for each user. OP policy is the one proposed in Section 5.2 with Gelenbe's expression used in the packet drop probability. The service rate and second order statistic of the service process is assumed to be known to the buffer manager. Note that packet service times are not independent in this simulation setting as packet length is assumed fixed. However, the approximate expression of Gelenbe is still employed. By means of extensive MATLAB simulations, Gelenbe's formula is shown to still approximate the drop probability.

One may argue that the proposed scheduling and buffer partitioning assumptions rely mainly on the idealistic assumption of Poisson arrivals. In the simulations, we will also examine the performance of the joint policies for bursty arrivals.

5.4.1 Simulation Setting and Results

Wireless channel of users is modeled as four-state independent discrete random variables. In each state, a rate is achievable. Assuming fixed length packets, we normalize the rate with packet length. For channel state i, i packets can be sent in each slot, $i \in \{0, 1, 2, 3\}$.

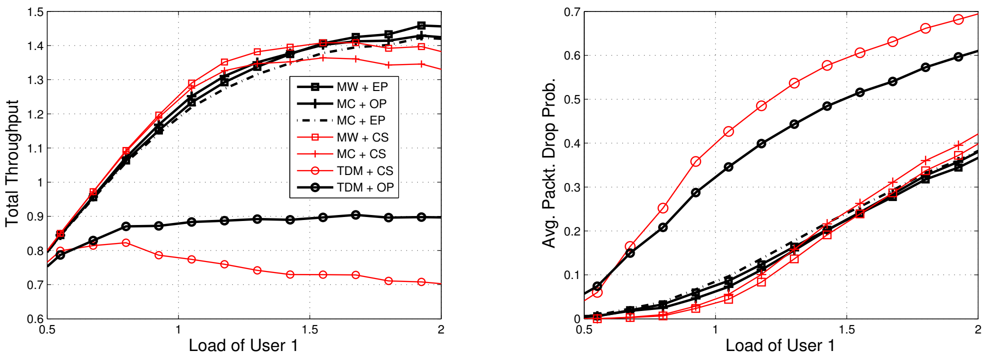

Figure 5.7: N=2 and B=5 buffers per user. x-axis represents load of user 1. $\rho_2 = 0.3$

We simulated several joint buffer management and user scheduling policies and compared their total throughput and average packet drop probability performances using MATLAB. In each experiment, 10^6 time slots are simulated. In the beginning of each slot, packets are flushed from the buffer due to the rate allocated in the previous slot and then arrivals are accepted if the management policy allows. In Figs. 5.7 and 5.8, 2 user system with λ_2 fixed and λ_1 is varied in the x-axis of the plot. Channel service capacity is $\mu = 0.35$ pckts/slot/user. Loads and throughput are normalized according to μ. $\rho_2 = 0.3$ in Fig. 5.7 and $\rho_2 = 0.6$ in Fig. 5.8. The multiuser diversity gain is observed when TDM and MC scheduling are compared. Throughput of MW scheduling with CS or EP is observed to outperform the others but the MC scheduling with OP performs quite close to MW. CS policy in each scheduling has decreasing throughput after some load level though partitioning retains its performance. Next,

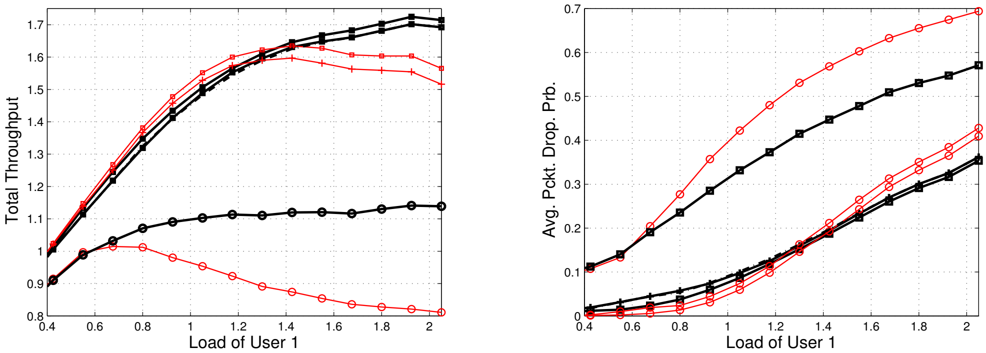

Figure 5.8: N=2 and B=5 buffers per user. Load of user 1 is changing in the x-axis while $\rho_2 = 0.6$

we experiment the joint policies in a 5 user downlink with unbalanced loads. The results are shown in Fig. 5.9. MW scheduling clearly outperforms the others. MC + OP policy comes after the MW. The advantage of optimal partitioning is observed.

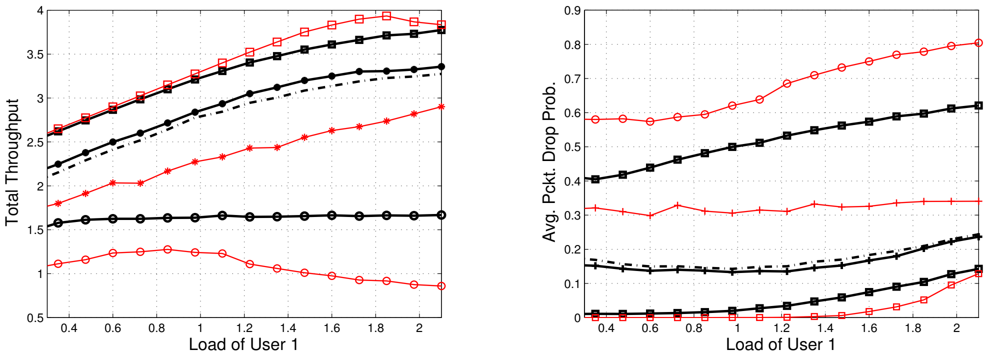

Figure 5.9: N=5 and B=5 buffer per user. A realistic unbalanced load regime: x-axis represents load of user 1 and $\rho_2 = 0.2$ $\rho_3 = 0.8$, $\rho_4 = 0.3$, $\rho_5 = 0.9$

In the last experiment, we examine the performance of the policies under bursty arrivals. In particular, Markovian Modulated Poisson Arrivals (MMPA) are assumed. Throughput and packet drop probability of the joint policies are shown in Fig. 5.10. Similar trends are observed as the Poisson case. The load level that starts to decrease in CS policy is observed to be lower for bursty arrivals.

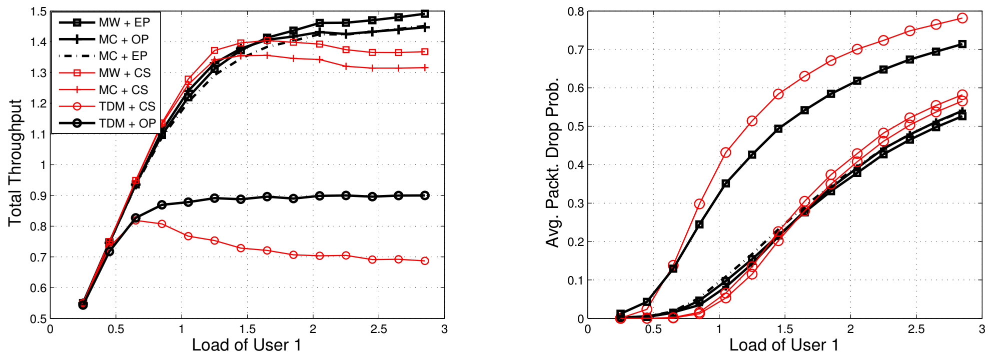

Figure 5.10: The performance of buffer partitioning under MMPA with modulating Markov chain transition probabilities equal to 0.5. Arrival rate is 0.2λ and 1.8λ according to the state of the Markov chain.

5.5 CONCLUSION

In this chapter, we proposed an efficient suboptimal buffer management method and showed several applications of the method to different communication scenarios. In particular, partitioning buffers judicuously for unbalanced load according to the arrival and service statistics is an effective method to boost throughput performance. We resort to partitioning as buffer management policy. Solution of optimal buffer partitioning is provided with an efficient algorithm. Then, buffer partitioning is considered jointly with user scheduling and channel assignment problems. It is numerically shown that using first order statistic of the arrival process along with buffer partitioning can provide good performance improvement.

CHAPTER 6

CONCLUSIONS AND FUTURE DIRECTIONS

In this thesis, we analyzed power control and buffer management and proposed distributed and efficient algorithms that can be applicable for different scenarios in wireless networks. We made use of game theory in the analysis of distributed power control while tools of queueing theory have been employed to analyze the buffer management problem in wireless networks.

Chapter 2 addresses energy efficient distributed power control in a wireless network with multiple access point system. Considering multiple access points as a single access point with multiple antennae distributed in space, users are allowed to multiplex their data to access points. Using a game theoretic approach, utility is set as *bits per Joule* and user strategies are vectors of powers in each access point direction. We show that users transmit to a single access point in the equilibrium point of the game. This way, usual single access point transmission gains another meaning: It is an energy-efficient strategy for users themselves. Moreover, the tradeoff between energy and spectral efficency is shown via modifying the utility function. It may be interesting to apply same ideas to the same multiple access points but considered as a macrodiversity system in this case. Message of a user is decoded jointly by using observations of all access points. As the diversity gain is observed in the exponent of bit error ratio, a quick investigation of the problem reveals that the sigmoidal assumption of efficiency function holds also for this case and the multiple access point system reduces to a single access point system.

In Chapter 3, a *smooth* distributed power control mechanism is developed. In general, equilibrium of utility based distributed power control is inefficient and externalities such as prices are required to punish users for their selfish behaviour. Sometimes this punishment may be so high that some users may be discarded from the system. We developed a new externality to the utility function based on reducing objectives as interference escalates. This reduction is rather *smooth* so that every user stays in the system unlike admission control. Apart from previous work, we analyzed the problem in terms of reactions to interference and we established sufficient conditions for unique equilibrium in terms of functional form of reactions to interference. Actually our new approach provides analysis of distributed actions from a practical perspective as implementation of the algorithm just requires the reactions to interference in the form of a lookup table on each node.

Notion of smooth reactions to interference is investigated from a control theoretic perspective in Chp. 4. Defining utility as a function of SIR objectives, we introduce the new externality to the utility function and net utility is obtained. Under absolute subhomogeneity, unique Nash equilibrium is guaranteed. Then we investigated convergence of gradient based power iterations and extended the analysis for asynchronous and erroneous cases. Finally, a continuous time dynamical system as a limiting case of iterative power updates has been analyzed.

We leave stability analysis of discrete time system open, as one of future directions of the current thesis work. Although the stability of continuous time equivalent system is more general, it may be interesting to investigate the stability of the discrete time system. Since there are not straightforward techniques to establish stability for general discrete time systems, the particular problem at hand may yield a simple analysis. One other issue that has not been addressed in the thesis is the convergence rates of iterative update algorithms. Although numerical results gave some feeling about the order of convergence rates, there is no analysis presented in the thesis. Hence, explicit analysis of convergence rates can be another future direction. The performance of algorithms can also be analyzed taking the variation of slowly fading component of channel gains into account so that the effect of possibly slow convergence rate and relatively faster channel variation is revealed. To address the dynamic channel in the future work, the notion of smoothness can be analyzed with extended game theoretic formulations. Strategy set of users can be the transmit power function in each joint channel state as was done in [33]. Finally, convergence rate of the update algorithms can also be examined in asynchronous case under certain delay models.

In chapter 5, management of finite buffer resource in a wireless downlink is addressed. Since partitioning available buffers judiciously for each user proves to be a good policy for high throughput performance, we attempted to solve the optimal buffer partitioning problem as that of a resource allocation. We proposed a fast algorithm for optimum partitioning. It requires the first order statistic of arrival process, first and second order statistic of the service process, which can be obtained by appropriate filtering. Then, we presented several applications of the algorithm. First, we applied the algorithm to parallel $M/M/1/m_k$ queues and illustrated that a considerable throughput increase is obtained. Then, buffer partitioning is considered jointly with channel assignment problem and a separation result is obtained: assign higher arrival rate user to better channel. Finally, the buffer partitioning problem is formulated jointly with user scheduling. By means of extensive simulations, buffer partitioning is shown to provide a good merit as a separate link layer mechanism that can operate without regard to instantaneous values of varying physical layer parameters. Although complex cross layer mechanisms such as MaxWeight scheduling achieves higher throughput performance, our algorithm operates on disjoint layer basis and can achieve comparable performance to MaxWeight. It is concluded that the benefit of cross layer design should be revised under finite buffer assumption.

Because of practicality and comfort it brings to our lives, wireless applications are becoming more popular day by day. Especially wireless last mile systems are deployed in offices or in cellular communication. Data arrives from a wired source and is served by a wireless link in the wireless last mile. Time varying nature of wireless links makes the wireless last mile systems vulnerable to congestion. Although TCP has an inherent congestion control mechanism, reaction time of the source to congestion can be high. The optimal buffer partitioning algorithm proposed in Chp. 5 is well suited to operate in the transient time between reactions to congestion. It can be applied as a supporting mechanism for TCP. Buffer partitioning can also be considered together with Active Queue Management, which are possible future extensions of the work. One can also elaborate on the queueing theoretic side of the problem. The use of buffer partitioning in several different arrival and service statistics can be investigated. For example, multiple sessions with bursty MAP arrivals and deterministic service can be analyzed. Moreover, results related to polling systems from queueing theory can be used to analyze maximum channel wireless scheduling.

REFERENCES

[1] D. Acemoglu, A. Ozdaglar, and R. Srikant. The marginal user principle for resource allocation in wireless networks. volume 2, pages 1544–1549, Dec. 2004.

[2] Tansu Alpcan, Tamer Basar, and S Dey. A power control game based on outage probabilites for multicell wireless data networks. *IEEE Transactions on Wireless Communications*, 5(4):890–899, April 2006.

[3] Tansu Alpcan, Tamer Basar, R Srikant, and Eitan Altman. Power control in CDMA as a non-cooperative game. *Kluwer Academic Publishers, Wireless Networks*, 8:659–670, 2002.

[4] E. Altman and A. Jean-Marrie. The loss process of messages in an m/m/1/k queue. *Proceedings of INFOCOM 94. Networking for Global Communications*, pages 1191–1198, 1994.

[5] Lachlan L. H. Andrew, Stephen V. Hanly, and Rami G. Mukhtar. Active queue management for fair resource allocation in wireless networks. *IEEE Transactions on Mobile Computing*, 7(2):231–246, February 2008.

[6] Tamer Basar and Geert Jan Olsder. *Dynamic Noncooperative Game Theory.* SIAM, Philedelphia, PA, 1999.

[7] D. Bertsekas and J. N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods.* Prentice Hall, Upper Saddle River, NJ, 1989.

[8] S. Buzzi and H.V. Poor. Noncooperative games for spreading code optimization, power control and receiver design in wireless data networks. In *eprint [Online], http://www.arxiv.org.*

[9] Isreal Cidon, Leonidas Georgiadis, Roch Guerin, and Asad Kamisy. Optimal buffer sharing. *IEEE Journal on Selected Areas in Communications*, 13(7):1229–1240, September 1995.

[10] S Dey and J Evans. Optimal power control in wireless data networks with outage-based utility guarantees. In *Proc. 2003 42nd IEEE Conference on Decision and Control*, volume 1, pages 570–575, December 2003.

[11] D. Falamori, N. Mandayam, and D. Goodman. *A New Framework for Power Control in Wireless Data Networks: Games, Utility, and Pricing*, volume Volume 524. Springer US, Online, 1999.

[12] Nan Feng, Siun-Chuon Mau, and Narayan B. Mandayam. Joint network-centric and user-centric radio resource management in a multicell system. *IEEE Transactions on Communications*, 53(7):1114–1118, July 2005.

[13] G. Foschini and B. Gopinath. Sharing memory optimally. *IEEE Transactions on Communications*, 31(3):352–360, March 1983.

[14] G. J. Foschini and Z. Miljanic. A simple distributed autonomous power control algorithm and its convergence. *IEEE Transactions on Vehicular Technology*, 42(4):641–646, November 1993.

[15] Drew Fudenberg and Jean Tirole. *Game Theory*. The MIT Press, Cambridge, MA, 1991.

[16] Robert G. Gallager. *Discrete Stochastic Processes*. Kluwer Academic Publishers, Boston/Dordrecht/London, 1996.

[17] Leonidas Georgiadis, Michael J. Neely, and Leandros Tassiulas. *Resource Allocation and Cross Layer Control in Wireless Networks*. Now Publishers Inc, Boston-Delft, 2006.

[18] David Goodman and Narayan Mandayam. Network assisted power control for wireless data. *Mobile Networks and Applications*, 6:1022–1026, 2001.

[19] David J. Goodman and Narayan B. Mandayam. Power control for wireless data. *IEEE Personal Communications Magazine*, 7:48–54, April 2000.

[20] K A Hamid. On the reliability of outage models. In *Proc. 2004 IEEE 60th Vehicular Technology Conference*, volume 6, pages 4349–4351, September 2004.

[21] S.V. Hanly. Capacity and power control in spread spectrum macrodiversity radio networks. *IEEE Transactions on Communications,*, 44(2):247–256, Feb 1996.

[22] Jianwei Huang, R.A. Berry, and M.L. Honig. Distributed interference compensation for wireless networks. *IEEE Journal on Selected Areas in Communications*, 24(5):1074–1084, May 2006.

[23] Marek I Irland. Buffer management in packet switch. *IEEE Transactions on Communications*, COM-26(3):328–337, March 1978.

[24] H Ji and C. Huang. Non-cooperative uplink power control in cellular radio systems. *Springer Wireless Networks*, 4(4):233–240, June 1998.

[25] Farouk Kamoun and Leonard Kleinrock. Analysis of shared finite storage in a computer network node environment under general traffic conditions. *IEEE Trans. Commun*, 28(7):992–1003, July 1980.

[26] Sunil Kandukuri and Stephan Boyd. Optimal power control in interference limited fading wireless channels with outage probability specifications. 1(1):46–55, January 2002.

[27] Abdin Keshavarzian, Elif Uysal-Biyikoglu, D Lal, and K Chintalapudi. From experience with indoor wireless networks: A link quality metric that captures channel memory. *IEEE Communications Letters*, 11(9):729–731, September 2007.

[28] H K Khalil. *Non-linear Systems.* Prentice Hall, Upper Saddle River, NJ, 1996.

[29] Hyogon Kim, Changhoon Han, and Inhye Kang. Reducing tcp response time in face of wireless uplink losses. In *Vehicular Technology Conference, 2001. VTC 2001 Fall. IEEE VTS 54th*, volume 1, pages 262–266 vol.1, 2001.

[30] S. Kittipiyakul and Tara Javidi. Resource allocation in OFDMA with time-varying channel and bursty arrivals. *IEEE Communications Letters*, 11(9):708–710, September 2007.

[31] R. Knopp and P. A. Humblet. Information capacity and power control in single-cell multiuser communications. In *Proc. Int. Conf. Commun.'02*, volume 1, pages 331–335, Seattle, USA, June 1995.

[32] S. Krishnan, A.K. Choudhury, and F.M. Chiussi. Dynamic partitioning: a mechanism for shared memory management. In *INFOCOM '99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 1, pages 144–152 vol.1, Mar 1999.

[33] Lifeng Lai and Hesham El Gamal. The water-filling game in fading multiple access channels. *IEEE Transactions on Information Theory*, 54(5):2110–2122, May 2008.

[34] Dhananjay Lal, A. Manjeshwar, F. Herrmann, E. Uysal-Biyikoglu, and A. Keshavarzian. Measurement and characterization of link quality metrics in energy constrained wireless sensor networks. In *IEEE GLOBECOM 03*, volume 1, pages 446–452, Dec. 2003.

[35] Douglas J. Leith, Lachlan L. H. Andrew, Tom Quetchenbach, Robert N. Shorten, and Kfir Lavi. Experimental evaluation of delay/loss based TCP congestion control algorithms. In *[Online],* $http://www.hamilton.ie/net/delay_tests_final.pdf$.

[36] G Li and Hui Liu. Dynamic resource allocation with finite buffer constraint in broadband ofdma networks. In *Proc. IEEE Wireless Communications and Networking Conference(WCNC) '03*, volume 2, pages 1037–1042, March 2003.

[37] P. Liu, R.A. Berry, and M.L. Honig. A fluid analysis of a utility-based wireless scheduling policy. *IEEE Transactions on Information Theory*, 52(7):2872–2889, July 2006.

[38] F. Meshkati, M. Chiang, H. V. Poor, and S. C. Schwartz. A game-theoretic approach to energy-efficient power control in multicarrier CDMA systems. *IEEE Journal on Selected Areas in Communications*, 24(6):1115–1129, June 2006.

[39] F. Meshkati, A. J. Goldsmith, H. V. Poor, and S. C. Schwartz. A game-theoretic approach to energy-efficient modulation in CDMA networks with delay QoS constraints. *IEEE Journal on Selected Areas in Communications*, 24(6):1069–1078, August 2007.

[40] F. Meshkati, H. V. Poor, S. C. Schwartz, and N. B. Mandayam. An energy-efficient approach to power control and receiver design in wireless data networks. *IEEE Transactions on Communications*, 53(11):1185–1194, November 2005.

[41] T. Ibaraki N. Katoh and H. Mine. A polynomial time algorithm for the resource allocation problem with a convex objective function. *Operations Research*, 30(5):449–455, May 1979.

[42] Carl J Nuzman. Contraction approach to power control with non-monotonic applications. In *Proc. IEEE Global Telecommunications Conference (Globecom)*, November 2007.

[43] Omur Ozel and Elif Uysal-Biyikoglu. Distributed power control using non-monotonic reaction curves. In *Proc. 2009 Intl. Conf. on Game Theory for Networks*, pages 60–67, May 2009.

[44] Omur Ozel and Elif Uysal-Biyikoglu. Power control games in wireless networks. In *IEEE 17th Signal Processing and Communications Applications Conference*, pages 496–499, April 2009.

[45] Omur Ozel and Elif Uysal-Biyikoglu. Distributed control of transmit power for energy efficiency in wireless networks with multiple access point. submitted.

[46] Omur Ozel and Elif Uysal-Biyikoglu. A power control game with smooth reduction of SINR objectives. submitted.

[47] John Papandriopoulos, J Evans, and S Dey. Optimal power control for rayleigh-faded multiuser systems with outage constraints. *IEEE Transactions on Wireless Communications*, 4(6):2705–2715, November 2005.

[48] John Papandriopoulos, J Evans, and S Dey. Outage-based optimal power control for generalized multiuser fading channels. 54(4):693–703, April 2006.

[49] John Papandriopoulos, Jamie Evans, and Subhrakanti Dey. Optimal power control in CDMA networks with constraints on outage probability. In *in Proc. WiOpt03, INRIA Sophia Antipolis*, pages 279–284, 2003.

[50] A. Paulraj, R. Nabar, and D. Gore. *Introduction to Space Time Wireless Communications.* Cambridge University Press, New York, USA, first edition, 2003.

[51] V. Rodriguez. An analytical foundation for resource management in wireless communication. In *Proceedings of the IEEE Global Telecommunications Conference (Globecom)*, pages 898–902, San Francisco, CA, USA, December 2003.

[52] V. Rodriguez. Robust modeling and analysis for wireless data resource management. In *Proc. 2003 IEEE Wireless Comm. Netw. Conf. New Orleans LA*, pages 717–722, March 2003.

[53] Cem U Saraydar, Narayan B Mandayan, and David J. Goodman. Efficient power control via pricing in wireless data networks. *IEEE Transactions on Communications*, 50(2):291–303, February 2002.

[54] C.U. Saraydar, N.B. Mandayam, and D.J. Goodman. Pricing and power control in a multicell wireless data network. *IEEE Journal on Selected Areas in Communications*, 19(10):1883–1892, October 2001.

[55] S. Sarkar. Optimum scheduling and memory management in input queued switches with finite buffer space. *IEEE Transactions on Information Theory*, 50(12):3197–3220, Dec. 2004.

[56] E Senata. *Non-Negative Matrices and Markov Chain*. Springer -Verlag, New York, 1981.

[57] V. Shah, N.B. Mandayam, and D. J. Goodman. Power control for wireless data based on utility and pricing. In *The Ninth IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1427–1432, September 1998.

[58] W. Shih. A new application of incremental analysis in resource allocations. *Operations Research*, 25(4):587–597, December 1974.

[59] G. Song, Y. Li, Jr. L. J. Cimini, and H Zheng. Joint channel-aware and queue-aware data scheduling in multiple shared wireless channels. In *Proc. IEEE Wireless Communications and Networking Conference(WCNC) '03*, volume 3, pages 1039–1044, March 2004.

[60] Vivek Srivastava, James Neel, Allen B. McKenzie, Rekha Menon, Luiz A. DaSilva, James E. Hicks, Jefferey H. Reed, and Robert P. Gilles. Using game theory to analyze wireless ad hoc networks. *IEEE Communications Surveys and Tutorials*, 7(4):46–56, Fourth Quarter 2005.

[61] Junhua Tang, Gang Feng, Siew Chee-Kheong, and Liren Zhang. Providing differentiated services over shared wireless downlink through buffer management. *IEEE Transactions on Vehicular Technology*, 57(1):548–555, January 2008.

[62] D. Tse and P. Viswanath. *Fundamentals of Wireless Communications*. Cambridge University Press, New York, first edition, 2005.

[63] David Tse. Optimal power allocation over parallel Gaussian broadcast channels. In *[Online], http://degas.eecs.berkeley.edu/dtse/pub.html*.

[64] David Tse and Pramod Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, Cambridge, 2005.

[65] E. Uysal-Biyikoglu. Bounds and observations on the transmission of data packets with minimum energy. *IEEE Communications Letters*, 10(10):707–709, Oct. 2006.

[66] E. Uysal-Biyikoglu. Design of minimum-energy high performance wireless communication networks: Inter-layer optimization and algorithms. project sponsored by TUBITAK, No:106E119, 2006.

[67] S. Verdu. On channel capacity per unit cost. *IEEE Transactions on Information Theory*, 36(5):1019–1030, Sep 1990.

[68] P. Viswanath, D. Tse, and R. Laroia. Opportunistic beamforming using dumb antennas. *IEEE Transactions on Information Theory*, 48(6), June 2002.

[69] Guo-Liang Wu and J.W. Mark. A buffer allocation scheme for atm networks: complete sharing based on virtual partition. *IEEE/ACM Transactions on Networking*, 3(6):660–670, Dec 1995.

[70] Mingbo Xiao, Ness B. Shroff, and Edwin K. P. Chong. A utility-based power control scheme in cellular wireless systems. *IEEE/ACM Transactions on Networking*, 11(2):210–221, April 2003.

[71] R. D. Yates. A framework for uplink power control in cellular radio systems. *IEEE Journal on Selected Areas in Communications*, 13(7):1341–1347, September 1995.

[72] R. D. Yates and C. Y. Huang. Integrated power control and base station assignment. *IEEE Transactions on Vehicular Technology*, 44(3):638–644, August 1995.

[73] Edmund Yeh and Aaron S. Cohen. Information theory, queueing, and resource allocation in multi-user fading communications. In *Proc. of the 2004 Conference on Information Sciences and Systems*, pages 1396–1401, Princeton, NJ, USA, March 2004.

[74] E.M. Yeh and A.S. Cohen. Throughput and delay optimal resource allocation in multiaccess fading channels. *Information Theory, 2003. Proceedings. IEEE International Symposium on*, pages 245–245, June-4 July 2003.

[75] Jens Zander. Distributed cochannel interference control for cellular radio systems. *IEEE Transactions on Vehicular Technology*, 41(3):305–311, August 1992.

[76] Serhan Ziya. On the relationships among traffic load, capacity, and throughput for the m/m/1/m, m/g/1/m-ps, and m/g/c/c queues. *IEEE Transactions on Automatic Control*. (to appear).

APPENDIX A

It will be shown that

$$\frac{f(\widehat{h}_{kb_k^*}p_k^*)}{p_k^*} \geq \frac{\sum_b f(\gamma_{kb})}{\sum_b p_{kb}} \; \forall [p_{11}\; p_{12} \cdots p_{1M}] \in S_k \tag{.1}$$

Let $W = [0, P_{max}]$. Without loss of generality, assume $k = 1$ (i.e. consider the first user) and $b_k^* = 1$ (i.e. for the first user the maximum $\widehat{h}_{bk}$ parameter is obtained with base station 1). Hence,

$$\widehat{h}_{11} \geq \widehat{h}_{1b} \; \forall b \in \{1, 2, \cdots, M\} \tag{.2}$$

There are two cases to consider:

1. $p_{11}^* = \frac{\gamma^*}{\widehat{h}_{11}}$
2. $p_{11}^* = P_{max}$

Assume the first case. Note that $\gamma_{1b} = \gamma_{1b}(p_{11},\; p_{12}, \cdots, p_{1M})$ is a function of user 1's power strategy vector $[p_{11}\; p_{12} \cdots p_{1M}]$ given the other p_{ij}. As for $b = 1$, we have $\forall p_{11} \in W$

$$\frac{f(\gamma^*)}{p_{11}^*} \geq \frac{f[\gamma_{11}(p_{11},\; p_{12} = 0,\; p_{13} = 0,\; \cdots, p_{1M} = 0)]}{p_{11}} \tag{.3}$$

Then the inequality follows:

$$\frac{f(\gamma^*)}{p_{11}^*} \geq \frac{f[\gamma_{11}(p_{11},\; p_{12},\; p_{13},\; \cdots, p_{1M})]}{p_{11}} \tag{.4}$$

$$\forall [p_{11}\; p_{12} \cdots p_{1M}] \in S_1$$

Consider for $b = 2$, maximization is at either $p_{12} = P_{max}$ or $p_{12} = \frac{\gamma^*}{\widehat{h}_{12}}$. For the latter case, the procedure is similar to previous one. In the former case, at $p_{12} = P_{max}$, then

$$\frac{f(\gamma_{12}(p_{11} = 0,\; p_{12} = P_{max},\; p_{13} = 0,\; \cdots, p_{1M} = 0)}{P_{max}}$$

$$\leq \frac{f(\gamma^*)}{p'_{12}} \leq \frac{f(\gamma^*)}{p^*_{11}}$$

for some $p'_{12} \geq P_{max}$ such that $p'_{12} = \frac{\gamma^*}{\hat{h}_{21}}$, then it is obvious that $p'_{12} \geq P_{max} \geq p^*_{11}$. Hence it follows $\forall p_{12} \in W$:

$$\frac{f(\gamma^*)}{p^*_{11}} \geq \frac{f[\gamma_{12}(p_{11}=0,\ p_{12},\ p_{13}=0,\ \cdots, p_{1M}=0)]}{p_{12}} \tag{.5}$$

We immediately see that:

$$\frac{f(\gamma^*)}{p^*_{11}} \geq \frac{f[\gamma_{12}(p_{11},\ p_{12},\ p_{13},\ \cdots, p_{1M})]}{p_{12}} \tag{.6}$$

$$\forall [p_{11}\ p_{12} \cdots p_{1M}] \in S_1$$

By proceeding similarly for other base stations, the inequalities obtained in .4 and .6 can be generalized :

$$\frac{f(\gamma^*)}{p^*_{11}} \geq \frac{f[\gamma_{1b}(p_{11},\ p_{12},\ p_{13},\ \cdots, p_{1M})]}{p_{1b}} \tag{.7}$$

$$\forall [p_{11}\ p_{12} \cdots p_{1M}] \in W^M and\ \forall b \in \{1, 2, \cdots, M\}$$

Converting the inequalities to

$$\frac{p_{1b}}{p^*_{11}} \geq \frac{f[\gamma_{1b}(p_{11},\ p_{12},\ p_{13},\ \cdots, p_{1M})]}{f(\gamma^*)} \tag{.8}$$

$$\forall [p_{11}\ p_{12} \cdots p_{1M}] \in S_1$$

and summing over b, we obtain:

$$\frac{\sum_b p_{1b}}{p^*_{11}} \geq \frac{\sum_b f[\gamma_{1b}(p_{11},\ p_{12},\ p_{13},\ \cdots, p_{1M})]}{f(\gamma^*)} \tag{.9}$$

$$\forall [p_{11}\ p_{12} \cdots p_{1M}] \in S_1$$

Converting once more, we obtain the desired result.

In the second case, $p^*_{11} = P_{max}$. Using the maximality assumption of $\hat{h}_{11}$, we observe that $p^*_{1b} = P_{max}\ \forall b \in \{1, 2, \cdots, M\}$.

$$\frac{f(\hat{h}_{1b} p^*_{1b})}{p^*_{1b}} = \frac{f(\hat{h}_{1b} P_{max})}{P_{max}}\ \forall b \in \{1, 2, \cdots, M\}\,. \tag{.10}$$

Again using the maximality of $\hat{h}_{11}$, we reach Eqn. .7. Hence, the desired result follows for the second case. ■

APPENDIX B

Best response $r_k(p_k(t), P_{-k}(t)) = \min\{p_i(t), P_{max}\}$ where $p_i(t)$ is determined in two steps as follows:

$$b_i^*(t) = \arg\max_{b \in B} \frac{G_p\ h_b p_i(t)}{\sum_{j \neq i} h_b p_j(t) + \sigma_b^2}$$

$$p_i(t) = \frac{\gamma^*(\sum_{j \neq i} h_{b_i^*(t)j} p_j(t) + \sigma_{b_i^*}^2)}{G_p\ h_{b_i^*(t)i}}$$

It is sufficient to show that the update algorithm $I(.)$ with $I_i = p_i(t)$ is a standard algorithm.

1. The positivity of the algorithm is obvious: for all $p(t) \geq 0$, $p_i(t+1) > 0, i = 1, 2, ..., K$.

2. The monotonicity of the algorithm follows from the fact that if $p^x(t) > p^y(t)$ then $\tilde{h}^x_{b_i^*(t)i}(t) < \tilde{h}^y_{b_i^*(t)i}(t)$ for all i. Note that the base station selections for $p_i^x(t)$ and $p_i^y(t)$ may be different. That is, in general $b_i^{x*}(t+1) \neq b_i^{y*}(t+1)$ for all i. Yet the inequality always holds because of the monotone decreasing property of $\tilde{h}_{b_i^*(t)i}(t)$ with $p(t)$. $p_i^x(t+1) > p_i^y(t+1)$ for all i. Hence $p^x(t+1) > p^y(t+1)$.

3. The scalability of the algorithm follows with a similar reasoning. Let $p^x(t) = \alpha p^y(t)$ for some $\alpha > 1$. For a fixed base station (i.e. $b_i^*(t)$ is same for all t), scaling the powers with α leads to $p_i^1(t+1) < \alpha p_i^2(t+1)$. As the minimum power requiring base station is selected, this inequality is again satisfied for all i. Hence, $p^x(t+1) < p^y(t+1)$. We conclude that the power update algorithm is standard. ∎