Resource Management in Multi-user Communication Systems

A THESIS

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Yao Huang Morin

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Prof. Zhi-Quan (Tom) Luo, Advisor

July 2009

Acknowledgement

I would like to take this opportunity to thank my colleagues in the University of Minnesota, Twin Cities, who have inspired me of new research ideas and encouraged me to work hard for my PhD program, in which I enjoy the challenges and hard works.

First of all, I would like to express my gratitude to my advisor Professor Zhi-Quan Luo who has provided advices and guidance without any reservation as well as financial supports for my study. I would also like to thank the dissertation committee: Professor Beth Allen, Professor Georgios B. Giannakis and Professor Mostafa Kaveh for their reviews and feedbacks on my thesis.

I feel unbelievably blessed that I am able to come to University of Minnesota and work with some extremely talented people. Specially, I greatly appreciate the opportunity I had to work with Dr. Ramy Gohary, with whose great collaboration I finished a major portion of my PhD research. I was also very happy to work with students from my research laboratory including Jinjun Xiao, Alireza Razavi, Jaymes Grossman, Meisam Razaviyayn, Randy Plate, Yingxi Liu and visiting PhD candidate Tsung-Hui Chang.

Dedication

This dissertation is dedicated to my family and friends, especially to

—the memory of my father, Shuihui Huang, who emphasized on the importance of education and hard work.

—my mother, Yuxia Lu, who loves me deeply and has supported me in every way she can.

—my husband, Andrew Morin, who has encouraged me through numerous number of phone calls and emails, who has made me strong and happy, and who loves me with all his heart.

Abstract

This dissertation aims at establishing optimal power allocation/control schemes to achieve maximal system utilities in a multi-carrier communication system, such as a digital subscriber line system and open spectrum wireless network. In such a system, users share tones to enhance the efficiency of spectrum usage due to its scarcity. However, this brings the system's intrinsic problem of inter-user interference, which has a crucial impact on communication quality. Therefore, the goal is to eliminate or diminish the impact of interference on the achievable data rate, which is a conventional measure of a user's communication quality. Based on the users' achievable data rates, system utility is defined. Hence, the goal comes to finding power allocation that can maximize the system utilities. According to different system requirements, we consider three system utilities: the weighted sum rate, "user capacity" and harmonic mean rate. For each utility function, we develop an efficient algorithm designed according to the features of the corresponding utility function.

Spectral Spectrum Balancing (SSB) aims to maximize the first goal (weighted sum rate). The algorithm partitions the N tones into three sections and efficiently determines the tones that lie in each section. Appropriate signalling structure is imposed on each section: The first section where the tones for which the crosstalk coefficients are small uses iterative water filling signalling method, the second section consists of tones with intermediate crosstalk coefficients and uses a delicate method to identify the user pairs that should share tones and Lagrangian method to allocate the power, and the third section where users suffer large crosstalk coefficients uses a dual FDMA algorithm.

While weighted sum rate is a popular measure of system utilities, we introduce "user capacity", which is a more practical goal of commercial service provider's. "User capacity" denotes the maximum number of users that can be supported by the sys-

tem, provided that each user is guaranteed a data rate that lies within a prescribed range. However, allocating power directly to approach this capacity can be quite cumbersome because it involves solving an integer programming problem which is NP-hard. In order to circumvent this difficulty, an alternate approach is proposed that is based on exploiting the fairness and per-tone convexity of the harmonic mean-rate objective. Thus an iterative scheme is proposed to approximate the harmonic mean rate objective function based on its Taylor expansion. We further exploit its convex lower bound, the dual form of which can be decomposed into several convex problems decoupled across tones. We show by broad simulation results that the algorithms we develop serve their purposes and outperform existing counterparts.

We further consider the case when a malicious jammer is presents in the system, where the jammer's goal is to minimize the total sum of the rates communicated over the network. Each user, on the other hand, allocates its power across the N tones so as to maximize the total sum rate that he/she can achieve, while treating the interference of other users and the jammer's signal as additive Gaussian noise. For this non-cooperative game, we propose a generalized version of the existing iterative water-filling algorithm whereby the users and a jammer update their power allocations in a greedy manner. We study the existence of a Nash equilibrium in this non-cooperative game as well as conditions under which the generalized iterative water-filling algorithm converges to a Nash equilibrium of the game.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This PhD dissertation focuses on resource allocation in multi-user communication systems. Our motivation is that the cost of communication hardware resources accounts for a significant portion of the cost of a communication system in practice; efficient use of resources is important to improve system profits. The resources that are most important in practice are power, time slots, and frequency bandwidth. Optimal/near optimal power allocation algorithms designed in PhD dissertation under various system requirements significantly improve system performance and efficiency. Additionally, the behaviors of the users in allocating power are also studied from a game theoretic perspective.

1.1 Multi-User Communication Systems and Dynamic Spectrum Management

Digital subscriber lines (DSL), frequency-sharing wireless networks, and broadcasting TV channels are some real-life examples of the types of multi-user communication systems considered. They have several distinct characteristics as opposed to traditional systems with fixed resource assignment, such as mobile communication systems. First, there are multiple pairs of users requesting resources from the system. Each user pair is comprised of a transmitter with limited power budget and its designated receiver wishing to receive information from the transmitter. One can think of a room full of guests, where each of individuals wants to talk to another particular individual. Second, users have full access to all communication resources (e.g. time and spectrum) so as to enhance the efficiency of the system resource usage. Returning to the party example, all conversations desire to take place simultaneously. The advantage is that everyone can talk throughout whole party instead of waiting for a segment of time. However, this sharing nature of the system results in interference among users that significantly decreases the system's performance. Using the previous example, the party may be so loud that no conversation can be clearly heard. Therefore, controlling users' power when sharing resources is crucial in the multiuser communication environment because users transmission power has critical influence on the overall communication quality.

Conventionally, communication quality is evaluated by the signal to interference and noise ratio (SINR) or signal transmission rate, both of which are directly related to the users' power. Intuitively speaking, when users are sharing channels, the transmission power of one user affects other users' SINRs or transmission rates. More precisely, the SINR and signal transmission rate of one user increase when transmission power of this user increases and it decreases when those of other users increase. So raising one user's power increases the communication quality of this user and simultaneously decreases communication quality of other users. In an analogy of the party analogy, raising one's voice can make ones conversation clearer but may disrupt others' conversations. Therefore, in order to use resources most efficiently and maintain transmission quality in a network, system designers need to carefully balance power among users. That is, system designers manage users' power and provide users in the network with satisfying quality of communication. Therefore a power control scheme is needed to serve this purpose and control the power used by each user in the network.

Recently, power allocation has drawn a vast range of research interests in the context of Dynamic Spectrum Management (DSM). Dynamic spectrum management is a set of techniques that manages users in a way that they can efficiently share a common spectrum in communication systems. The research and development of DSM is widely based on optimization theory and game theory. Historically, a fixed bandwidth is assigned to each specific service and results in inefficient utilization of the spectrum. With the help of DSM, the spectrum is shared by users, which significantly increases the spectrum's efficiency without sacrificing communication quality. DSM can be used in DSL by continuously monitoring the interference among users and dynamically allocating users' power across frequency to reduce or eliminate the interference. The use of DSM in DSL is important because DSL systems are broadly used for domestic internet service and significant interference exists because DSL phone lines are closely bound together. Other than DSL, another potential use of DSM is on open spectrum communication system which permits any device to send signals across a certain spectrum range (unlicensed bands) without permission. This system scheme facilitates mobility and efficiency, and offers an attractive solution to the under-utilized traditional licensed bands. However, it requires spectrum sharing and causes interference. A common scenario in open spectrum systems is that new wireless devices (or secondary users) are able to recognize and access idle spectrum (currently not used by primary users). Therefore, secondary users should comply with regulations so as not to disrupt the communication of primary users. One example is the unused TV broadcast channels, which are also called white spaces.

Recently, the Federal Communication Commission (FCC) has approved the use of unlicensed devices to operate on the empty analog TV channels in anticipation of the transition to digital TV. While this potentially provides users with more access to cheap spectrum, it also gives rise to interference caused by spectrum sharing among these unlicensed devices. Therefore, strict interference regulation is required in order to guarantee the quality of service provided by the unlicensed devices as well as the licensed TV broadcasters. DSM is an effective tool that providers can utilize to handle the interference regulation.

In the context of DSM, system requirements are usually formulated into objective functions of optimization problems with power constraints such as total power constraints and spectrum mask constraint. For example, in [1], [2] and [3], the weighted sum rate of the system is considered as the objective of optimization. Optimization techniques are utilized to solve the problems in an efficient way with relative low complexity. The utilization of game theory in DSM originates from the fact that the users' competitive behaviors can be effectively modeled by a non-cooperative game with the users' optimization objective as the payoff function in the game [4]. The game theoretic point of view on this traditional signal processing problem gives rise to many interesting aspects of the problems. In [5], [6], [7] and other numerous other studies, a variety of games based on the model proposed in [4] and the properties of their Nash equilibria are studied. In a typical game-theoretic formulation, each user is a player in the game and maximizes its own utility, which is usually its data rate. The users compete with each other and adjust its power allocation given the knowledge of other users' power. The data rate function is concave for the user when the others' power is fixed. So there is usually a Nash equilibrium in the game and the Nash equilibrium is unique under certain conditions. A more interesting case in the game is a system with the presence of a jammer. The disruption caused by the jammer arises the questions on not only the games and Nash equilibira but also on the stability of the system and the jammer's damage to the system. We have included a thorough study on this topic in Chapter 5.

1.2 Power Allocation and Interference Reduction

This research emphasizes power allocation over system spectrum since spectrum is a scarce and valuable resource that is currently under utilized due to conventional static tone-assignment policies. In FCC Policy Task Force Report, it is stated that "... portions of the radio spectrum are not in use for significant periods of time" and "... typical channel occupancy was less than 15%...". This inherent drawback of static tone-assignment has been a fundamental reason behind utilization of DSM in DSL systems and the emergence of unlicensed open-spectrum communication systems [8,9]. In these systems the spectrum is typically partitioned into N narrowband orthogonal tones and all users are allowed to use all the tones simultaneously. In comparison with the fixed tone-assignment policies, this setup offers significantly greater freedom in utilizing the spectrum. However, this freedom comes at the expense of a number of challenges that ought to be taken into consideration by the system designer. In particular, the inherent spectral overlap in these systems gives rise to so-called multi-user interference, which is a limiting factor for multi-user communication systems. Therefore, one of the major challenges in designing such a system is to manage users' power in a way so that users can achieve satisfying functionality and cause limited interference to other users.

In order to determine the power that each user allocates to the N tones, the system is typically managed by a system designer that has information about the channels and users' power. In a typical scenario, each user has a limited power budget and the system designer wishes to determine the power that each user ought to allocate to each tone in order for the system to achieve the maximum system utilities [10] (such as individual users' rates, weighted sum rates, harmonic mean rates, min rate utilities), the choice of which is tailored to specific system requirements. We study maximum sum rate and harmonic mean rate. The reason for choosing these two utilities is, briefly speaking, that they accommodate maximum system throughput and user fairness, which are the most common system requirements.

1.3 Sum Rate Optimization

The problem of finding power-allocations that maximize the sum-rate was shown in [11] to be NP-hard. This problem can be solved exactly using the optimal spectrum balancing (OSB) algorithm developed in [2]. However, the computational cost of this algorithm is quite prohibitive, which makes it suitable only for systems with small numbers of users and tones. A less complex algorithm that can be used for more practical systems is the so-called autonomous spectrum balancing (ASB) algorithm [3]. This algorithm provides an approximate solution to the power allocation problem and requires side information that may be difficult to acquire. Both the OSB and the ASB algorithms share common drawbacks. For instance, neither algorithm can be readily tailored to optimize alternative design objectives, nor to incorporate other design constraints. Moreover, neither algorithm takes fairness into consideration. In particular, both OSB and ASB tend to allocate power in such a way that favors stronger users over weaker ones. There are other popular decentralized algorithms that locally maximize rate objectives such as IWFA [4] and SCALE [12], with which we will compare our proposed algorithms in the simulation section.

We first develop a computationally efficient algorithm for approaching the maximum sum-rate of DSL systems based on its spectral features. Unlike currently available algorithms, the proposed algorithm partitions the tones into sections and imposes a signalling structure on each section. This signalling structure facilitates the optimization of the power allocations in each section. We begin by recalling a result from [13]. This result implies that if the crosstalk coefficients between users exceed a certain threshold on some tones, then these tones ought to be operated in a frequency division multiple access (FDMA) mode in order to approach the maximum sum-rate of the DSL system. That is, none of these tones ought to be occupied by more than one user. We also note that if the crosstalk coefficients between users are close to zero on some tones, then, from a sum-rate perspective, it may be beneficial for the users to share these tones. These observations suggest that, in a general DSL system, one can use the crosstalk coefficients to partition the tones into sections and utilize a different optimization technique on each section. In particular, we propose to partition the tones into three sections, where the membership of a tone in one of the sections depends on whether crosstalk coefficients on this tone lie below, in between or above two thresholds. In this disseratation, we determine these thresholds by using a quasi-bisection optimization technique for relaxing the (somewhat stringent) thresholds provided in [13].

Assuming that each user allocates a certain power budget to each section of tones, we deploy an optimization algorithm that is suitable for maximizing the sum-rate for the signalling structure in each section. In particular, for the section in which the users suffer strong cross-talks, we use the FDMA sum-rate maximization algorithm developed in [13]. This algorithm is computationally efficient due to the fact that it exploits the FDMA structure to decouple the Lagrange dual formulation across tones. Now, for the section in which the crosstalk coefficients are close to zero, we use the classical iterative water-filling algorithm (IWFA) [4]. In each iteration of this algorithm, each user updates its power allocation so as to water-fill [14] on the noise-plus-interference levels observed in the previous iteration. Finally, we consider the section of tones in which the crosstalk coefficients assume intermediate values that are neither large enough to operate the tones in an FDMA mode nor small enough to operate the tones in an IWFA mode. In this section, we propose using the Lagrange dual algorithm described in [15] along with a scheme that refines the partition of the users. This algorithm iteratively updates the primal and the dual variables using a standard gradient ascent algorithm. As we will see, the fact that this algorithm generates the dual solutions as well as the primal ones can be very useful for our power allocation algorithm. Finally, we develop a greedy technique for determining the power that each user allocates to each of the three sections. In particular, in each iteration of this technique, a sensitivity analysis is used to determine the section of tones that yields the highest sum-rate gain for a given power increment. This sensitivity analysis utilizes the Lagrange dual variables which are readily obtained from the algorithms deployed in the three sections; viz., the FDMA sum-rate maximization algorithm, the IWFA algorithm and the Lagrange dual algorithm.

1.4 User Capacity and Harmonic Mean Objective

In the second part of the research of power allocation, we direct our attention to the problem of maximizing the number of users that can be accommodated by a system where users are categorized into groups according to the quality-of-service that they purchase from the system provider. The reason for considering this utility function is that the more users a system accommodates the more revenue it brings. However, this design objective is integer-valued and hence generally difficult to handle directly. As an alternative, we consider a design problem in which we maximize the harmonic mean of the users' rates.

The objective of maximizing harmonic mean rate possesses two desirable features: first, from the users' perspective, this objective is known to be fairer than maximizing the sum-rate [11]; second, for single-tone systems, maximizing this objective can be cast as a convex optimization problem for which the global solution can be obtained efficiently. The first feature renders the harmonic mean a natural design objective for maximizing the number of users. This is because for one to be able to compare the number of users that can be supported by two systems, one ought to guarantee that the users obtain the same service in both systems. The second feature, on the other hand, enables us to design an efficient algorithm that exploits the per-tone convexity to provide power allocations that yield relatively high harmonic mean-rates. In particular, we begin by providing a lower bound on the harmonic mean-rate, which results in a per-user per-tone harmonic mean formulation. Using the dual form, we decompose the problem of maximizing the lower bound into several convex optimization problems. These problems are not coupled across tones, and hence result in low design complexity. This feature renders this algorithm attractive for practical application in systems with a large number of users and tones. Furthermore, in developing this algorithm we show how to incorporate different quality-of-service levels. With the quality-of-service guaranteed, we run an outer (quasi-bisection) algorithm for maximizing the number of users. In particular, for every number of users (with associated crosstalk coefficients), we solve a feasibility problem which serves as an indicator to whether this number of users can be supported by the system with the prescribed quality-of-service levels.

In contrast to solving harmonic mean function per-user per-tone previously described, we advance our research by considering the problem of maximizing the regular per user harmonic mean function over all tones; this is known to obtain fairer data rates to users while allowing the system designer to provide the quality of service guaranteed to system users. Nonetheless, the harmonic mean rate objective function is NP-hard in general. Therefore we should resort to an efficient approximation algorithm to solve this problem for practical use. We utilize the first order Taylor expansion as well as its convex lower bound. Based on the Taylor expansion, we make a connection between the weighted sum rate and the per-user harmonic mean, and draw insights based on the solution of harmonic mean to assist us to design a novel algorithm. This algorithm exploits the per-tone convexity to optimize the originally non-decomposable and computationally expensive harmonic mean objective function and provide power allocation strategies that yield close to optimal harmonic mean rate.

1.5 Open Spectrum System and Jamming

As mentioned in previous sections, the system we consider allows multiple users to access the shared spectrum simultaneously and freely. This feature renders these systems susceptible to antagonistic behavior of potential jammers, who may be interested in reducing the utility of the entire system.[1] For example, a jammer may be able to 'listen' to the users' transmissions, and then subsequently updates its power allocation across tones in order to reduce the total sum rate communicated over the network. As such, the procedure of both the users and the jammer can be represented as a non-cooperative game [16] in which players are interested in maximizing their individual utilities in a selfish fashion. Since the impact of the jammer's signal can be deleterious to the overall system performance, our goal is to study Nash equilibrium of this sum rate game and subsequently the jammer's effect on the achievable system

[1] In this dissertation, the sum rate of each user across tones will be referred to as the utility of the user, and the sum of utilities of all users will be referred to as the system utility.

utility. We also wish to analyze the convergence behavior of a generalized version of IWFA for this non-cooperative game in which users and the jammer sequentially update their power allocations in a greedy manner to maximize their respective utilities. The distributed nature of IWFA makes it an attractive power allocation strategy not only in open wireless communications but also in multi-carrier communications such as DSL.

Convergence analysis of the IWFA algorithm were developed in [5] and [11] for synchronous systems and for asynchronous systems in [17] and [18]. While IWFA is a popular decentralized power allocation algorithm that individual users may use to locally maximize their rate objectives, other decentralized strategies were proposed in [6, 19–21] for various utility functions in the jammer-free case. In the presence of a jammer, single-user systems in which the jammer's goal is to minimize the mutual information of the 'legitimate' user were considered in [7] and [22]. The jammer's impact was also studied in [23] and [24] for multi-user single-tone communication systems in which the users' utilities are not directly related to rate maximization.

The remainder of the dissertation consists of five sections: In Section 2, we introduce the multi-user communication system and formulation of the data rate and system utilities. In Chapter 3, we study sum rate system utility and develop a power allocation scheme based on structure feature of DSL system. Chapter 4 focuses on harmonic mean rate utility and its advantages on user fairness as well as number of users served in a system. In Chapter 5, we study the system with a jammer from the game theoretical perspective based on [25]. Chapter 6 concludes this dissertation.

Chapter 2

System Model and Problem Statement

2.1 Multi-user Communication Systems

In this section we define the fundamental system model in this dissertation. The system we consider is a multi-user communication system in which N tones are shared by K users. Here a 'user' refers to a transmitter and receiver pair that attempts to communicate. (In a DSL system, a 'user' may refer to a central office (CO) or a remote terminal (RT) that transmits data to a modem at the subscriber's end.)

Let h_{jk}^n be the complex channel gain between the transmitter of User j and the receiver of User k on the n-th tone, where $n \in \mathcal{N} \triangleq \{1, \ldots, N\}$ and $j, k \in \mathcal{K} \triangleq \{1, \ldots, K\}$. In this notation h_{kk}^n denotes the channel gain between the transmitter of the k-th user and its intended receiver and $\alpha_{jk}^n \triangleq |h_{jk}^n|^2/|h_{kk}^n|^2$ denotes the normalized channel gain or crosstalk coefficients. Let s_k^n be the power allocated by User k to the n-th tone. Thus User k's signal power and interference power can be written as $|h_{k,k}^n|^2 s_k^n$ and $\sum_{j \neq k} |h_{j,k}^n|^2 s_j^n$, respectively (see Figure 2.1 for an illustration of signal

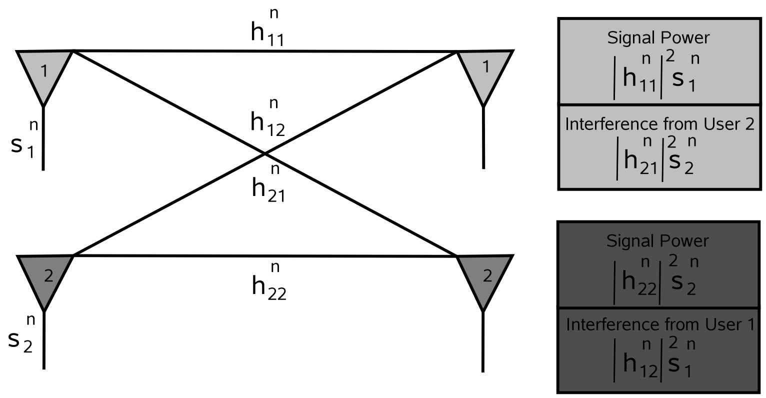


Figure 2.1: Interference Channel System

and interference power in a 2-user communication system). Assuming that each user uses Gaussian signalling and that every user can only decode its intended messages, the maximum rate that User $k \in \mathcal{K}$ can achieve on the n-th tone is given by [14]

$$R_k^n(s_1^n, \ldots, s_K^n) = \log\Big(1 + \frac{s_k^n}{\sigma_k^n + \sum_{j \neq k} \alpha_{jk}^n s_j^n}\Big), \tag{2.1.1}$$

where $\sigma_k^n \triangleq N_0/|h_{kk}^n|^2$ denotes the normalized noise variance observed by User k on the n-th tone, and N_0 is the variance of the background Gaussian noise. Now, the total rate R_k of User k is given by $\sum_{n=1}^N R_k^n(s_1^n, \ldots, s_K^n)$, and it is the maximum rate that can be reliably transmitted over the communication channel [14]. In the next section, we will define our system utilites based on this achievable rate.

2.2 System Utilities

There are several commonly used choices of system utility functions based on the achievable data rate defined in (3.1.1):

(i) Weighted sum of the rates of all tones and users: $\sum_{k=1}^K \sum_{n=1}^N R_k^n$;

(ii) Geometric mean of the rates of users' individual rates: $\left(\prod_{k=1}^{K} R_k\right)^{\frac{1}{K}}$;

(iii) Harmonic mean of the rates of users' individual rates: $K\left(\sum_{k=1}^{K} (R_k)^{-1}\right)^{-1}$;

(iv) Minimum of the the rates of users' individual rates: $\min_{1\le k\le K} R_k$.

In this dissertation, we focus on the utility function (i) and (iii). In addition to the above utilities, we define a new utility function in this dissertation called 'user capacity', which is the maximum number of users a system can support with a prescribed minimum rate. In Chapter 4, we provide details on how to interpret this utility and solve optimization problems with this utility function as objective.

In the following, we will focus on sum rate utility, maximum number of users utility and harmonic mean rate utility to accommodate different system requirements. We aim to provide suitable utility functions and the corresponding algorithms for specific system requirements. For example, sum rate utility is chosen when the system designer requires maximum system throughput (Section 3), maximum number of users utility is used when the system designer desires to serve as many users as possible with the prescribed minimum rate (Chapter 4), and harmonic mean carefully balances both aspects and provides weaker users fairer rates compared to sum rate and other utility functions (Chapter 4). Furthermore, in Chapter 5, we study the behaviors of users with individual rate utility. That is, each of the users has individual rate $\sum_{n=1}^{N} R_k^n$ as its utility function and aims at maximizing its utility selfishly. We also study this case with the presence of a malicious jammer in terms of the system performance and its stability.

Chapter 3

Power Allocation for Sum Rate Objective Function

3.1 System Model and Problem Formulation

Consider a DSL communication system (Figure 3.1 shows a DSL system topology), in which N tones are shared by K users. Let h_{jk}^n be the complex channel gain between the transmitter of User j and the receiver of User k on the n-th tone, where $n \in \mathcal{N} \triangleq \{1, \ldots, N\}$ and $j, k \in \mathcal{K} \triangleq \{1, \ldots, K\}$. In this notation h_{kk}^n denotes the channel gain between the transmitter of the k-th user and its intended receiver. Let the crosstalk coefficient from User j to User k on the n-th tone be denoted by $\alpha_{jk}^n \triangleq |h_{jk}^n|^2 / |h_{kk}^n|^2$, and let s_k^n be the power allocated by User k to the n-th tone. Assuming that each user uses Gaussian signalling and that every user can only decode its intended messages, the maximum rate that User $k \in \mathcal{K}$ can achieve on the n-th tone is given by (see [14]).

$$R_k^n(s_1^n, \ldots, s_K^n) = \log\Big(1 + \frac{s_k^n}{\Gamma\big(\sigma_k^n + \sum_{j \neq k} \alpha_{jk}^n s_j^n\big)}\Big), \tag{3.1.1}$$

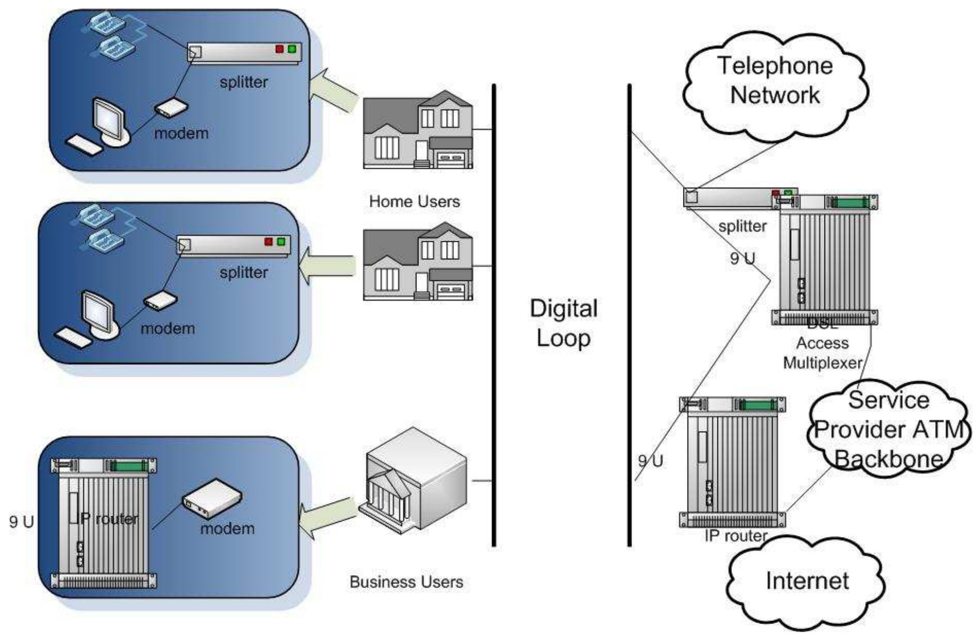


Figure 3.1: A DSL System Topology

where $\sigma_k^n \triangleq N_0/|h_{kk}^n|^2$ denotes the normalized noise variance observed by User k on the n-th tone, N_0 is the variance of the background Gaussian noise, and Γ is the so-called capacity gap, which is typically used to account for the non-Gaussianity of the signalling constellations used in practice [3].

Now, the system designer's goal is to find the power allocation that maximizes the over all sum-rate $\sum_{k=1}^{K}\sum_{n=1}^{N} R_k^n(s_1^n,\ldots,s_K^n)$, provided that the total and per-tone powers utilized by each user do not exceed certain thresholds. In addition, the system designer may wish to enforce a bit-cap B_k^n in order to ensure that the rates communicated on each tone can be supported by commercial modulators [2]. Using these constraints, it can be shown that the power allocation problem can be formulated as

$$
\begin{aligned}
\max \quad & \textstyle\sum_{k=1}^{K}\sum_{n=1}^{N} R_k^n(s_1^n,\ldots,s_K^n), && \text{(3.1.2a)}\\
\text{subject to} \quad & \textstyle\sum_{n=1}^{N} s_k^n \le P_k,\ \forall\, k, && \text{(3.1.2b)}\\
& 0 \le s_k^n \le S_{\max,k}^n,\ \forall\, k,n, && \text{(3.1.2c)}\\
& \left(2^{B_k^n}-1\right)^{-1} s_k^n - \textstyle\sum_{j\neq k}\alpha_{jk}^n s_j^n \le \sigma_k^n,\ \forall\, k,n, && \text{(3.1.2d)}
\end{aligned}
$$

where P_k, $S_{\max,k}^n$ and B_k^n are the total power budget, the spectral mask and the bit-cap of User k on the n-th tone, respectively. Note that (3.1.2) is not a convex optimization problem even though the constraints (3.1.2b), (3.1.2c) and (3.1.2d). are linear. The non-convexity of the problem comes from the fact that the objective function is non-convex. We will further examine the properties of the optimization problem and propose an efficient algorithm to solve this problem.

3.2 Problem Decomposition

Solving (3.1.2) directly is known to be NP-hard [26], which makes the task of finding a global optimal solution rather formidable even for relatively small systems. As an alternative, we propose to use inherent features of the optimal solution [1] in order to decompose (3.1.2) into three subproblems that are relatively easy to solve. The major reason that makes partitioning the problem into three subproblems possible is that DSL system possesses a very special spectrum feature, that is, the cross-talks are small when frequency is low and large when frequency is high. Figure 3.2 shows the cross talks versus frequency, where ”1” presents low frequency area, ”2” represents medium frequency area, and ”3” presents high frequency area. This relation between frequency and cross-talks leads to our method of partitioning spectrum and treating

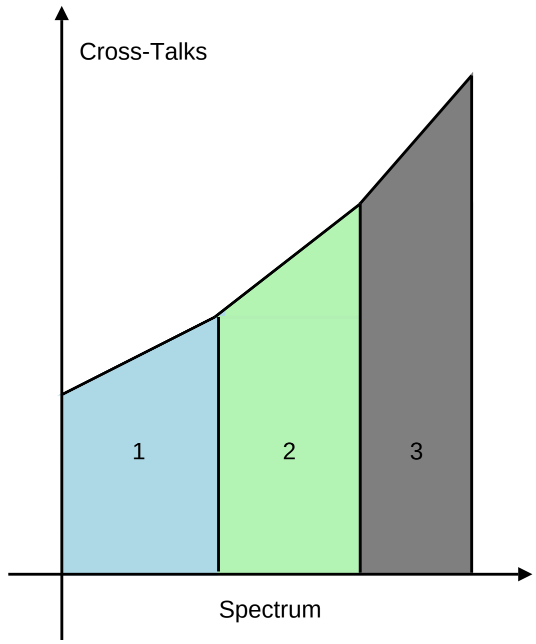


Figure 3.2: DSL System Spectrum

low, medium and high frequency bands differently. In the next sections we will describe our partitioning methodology.

3.2.1 FDMA-operated Tones

Let $\mathcal{F} \subseteq \mathcal{N}$ be the set of tones for which

$$\alpha_{jk}^n \alpha_{kj}^n \geq \tfrac{1}{4} - \delta_1, \quad \forall\, j \neq k \in \mathcal{K}, \tag{3.2.3}$$

where $\delta_1 \in [0, \frac{1}{4}]$ is a designed parameter to be determined. Let $P_{\mathcal{F},k}$ be the power allocated by User $k \in \mathcal{K}$ to the tones in $\mathcal{F}$. In order to understand the role of δ_1, we note that one of the key results in [1], implies that if $\delta_1 = 0$, then all the tones in $\mathcal{F}$ must be operated in an FDMA mode in order for the maximum sum-rate to be approached. However, since this condition is only sufficient, in some scenarios it may be too stringent and higher sum-rates can be obtained if more tones are operated in an FDMA mode. Hence, by operating the tones in $\mathcal{F}$ in an FDMA mode, δ_1 can be regarded as a parameter for relaxing the condition in [1].[1] For a given $P_{\mathcal{F},k}$, the algorithm in [27] can be used to find power allocations that maximize the sum-rate achieved on the tones in $\mathcal{F}$. This algorithm exploits the FDMA structure to decouple the power allocated across tones by assigning each tone to the user with the "so-called" highest 'shadow rate'. For a target precision of ϵ_1, the complexity of this algorithm can be shown to be $\mathcal{O}(K^2 \log^2 \epsilon_1)$ [1].

3.2.2 IWFA-operated Tones

For this section, let $\mathcal{W} \subseteq \mathcal{N} - \mathcal{F}$ be the set of tones for which

$$\alpha_{jk}^n \alpha_{kj}^n \leq \delta_2, \quad \forall\, j \neq k \in \mathcal{K}, \tag{3.2.4}$$

where $\delta_2 < \frac{1}{4} - \delta_1$ is a parameter that plays a role similar to the one played by δ_1 in (3.2.3). Let $P_{\mathcal{W},k}$ be the power that User $k \in \mathcal{K}$ assigns the tones in $\mathcal{W}$. Now, if for the tones in $\mathcal{W}$, $\alpha_{jk}^n = \alpha_{kj}^n = 0$, then δ_2 can be set equal to zero. Since in this case the users are completely decoupled, the maximum sum-rate can be achieved by classic water-filling [14]. However, in practice it is rarely the case that the cross-talk coefficients are exactly equal to zero, and $\mathcal{W}$ will contain those tones with small, but positive, cross-talk coefficients that satisfy (3.2.4) with sufficiently small δ_2. In this case, the maximum sum-rate on $\mathcal{W}$ can be approached by using the iterative, instead of the classic, water-filling algorithm (IWFA) [4]. The complexity of this algorithm is $\mathcal{O}(KN \log^2 \epsilon_1)$. Similar to δ_1, the value of δ_2 should be adjusted in order to maximize the sum-rate of the DSL system.

[1] Note that while (3.2.3) gives a sufficient FDMA optimality condition for any $K \geq 2$, a tighter condition can be used for $K = 2$; see [1].

3.2.3 Tones with Unstructured Signalling

Finally, let $\mathcal{M} \subseteq \mathcal{N} - (\mathcal{F} \bigcup \mathcal{W})$ be the set of tones on which the crosstalk coefficients do not satisfy either (3.2.3) or (3.2.4), and let $P_{\mathcal{M},k}$ be the power allocated by User $k \in \mathcal{K}$ to the tones in $\mathcal{M}$. Since the crosstalk coefficients on $\mathcal{M}$ assume intermediate values, the optimal signalling structure on $\mathcal{M}$ does not necessarily resemble any of the standard signalling patterns. Hence, we settle for power allocations that are locally sum-rate optimal. Such allocations can be found using the standard the primal-dual updates algorithm described in [15]. Similar to IWFA, the complexity of this algorithm is $\mathcal{O}(KN \log^2 \epsilon_1)$. In Section 3.5, we propose an advanced method to deal with this case with the help of the approximate max-clique algorithm to pair up the users that are compatible and solve the problem in a pairwise case.

3.3 Power Budget Partitioning

In the previous section a framework for partitioning the N tones into $\mathcal{F}, \mathcal{W}$, and $\mathcal{M}$ sections was presented. The powers allocated by any User $k \in \mathcal{K}$ to these sections are $P_{\mathcal{F},k}$, $P_{\mathcal{W},k}$, and $P_{\mathcal{M},k}$, respectively, where for (3.1.2b) to be satisfied, we must have[2]

$$P_{\mathcal{F},k} + P_{\mathcal{W},k} + P_{\mathcal{M},k} = P_k. \tag{3.3.5}$$

Our goal now is to find locally optimal $P_{\mathcal{F},k}$, $P_{\mathcal{W},k}$, and $P_{\mathcal{M},k}$ that enable the maximum sum-rate of the DSL system to be approached. In order to do that, we begin by introducing the following definitions. Let the k-th entry of $\mathbf{\Delta}_i \triangleq [\Delta_{i,1} \cdots \Delta_{i,K}]$ be an additional power by which User k increments $P_{\mathcal{F},k}$ and $P_{\mathcal{W},k}$ for $i = 1$ and 2, respectively. With $\Delta_{i,k}$ defined as such, satisfying the power constraint in (3.3.5),

[2] It can be seen that if each user occupies at least one tone in $\mathcal{F}$ the optimal solution of (3.1.2) must satisfy (3.1.2b) with equality.

implies that the decrement of $P_{\mathcal{M},k}$ is $\Delta_{3,k} = -\sum_{i=1}^{2} \Delta_{i,k}$. For given $P_{\mathcal{F},k}$, $P_{\mathcal{W},k}$, and $P_{\mathcal{M},k}$, let $g_1(\mathbf{\Delta}_1)$ be the (locally) optimal sum-rate on the $\mathcal{F}$ tones for a power increment $\mathbf{\Delta}_1$. That is, for every (sufficiently small) $\mathbf{\Delta}_1$, $g_1(\mathbf{\Delta}_1)$ is the solution of

$$\begin{aligned} \max \quad & \textstyle\sum_{n\in\mathcal{F}} \sum_{k\in\mathcal{K}} R_k^n, && (3.3.6a) \\ \text{subject to} \quad & \text{(3.1.2c) and (3.1.2d)}, \quad \forall\, n \in \mathcal{F}, && (3.3.6b) \\ & \textstyle\sum_{n\in\mathcal{F}} s_k^n - P_{\mathcal{F},k} = \Delta_{1,k}, \quad \forall\, k \in \mathcal{K}. && (3.3.6c) \end{aligned}$$

Similarly, one can define $g_2(\mathbf{\Delta}_2)$ and $g_3(\mathbf{\Delta}_3)$ for the sum-rates on the $\mathcal{W}$ and $\mathcal{M}$ tones, respectively. We note that, apart from the constraint in (3.3.5), the optimization problems that correspond to the functions $\{g_i(\cdot)\}_{i=1}^3$ are decoupled. Moreover, each of these problems can be solved efficiently using the techniques outlined in Section 3.2.

For brevity, we will focus on $g_1(\mathbf{\Delta}_1)$, and the analysis for the other two functions follows similar paths. Let us consider the Lagrange dual form of (3.3.6). For this dual let k-th entry of $\boldsymbol{\lambda}_1(\mathbf{\Delta}_1) \in \mathbb{R}^K$ be the Lagrange dual variable that corresponds to the k-th constraint in (3.3.6c). Now, using the sensitivity theorem in [15, Proposition 3.2.2], we have

$$\nabla_{\mathbf{\Delta}_1} g_1(\mathbf{\Delta}_1) = -\boldsymbol{\lambda}_1(\mathbf{\Delta}_1). \tag{3.3.7}$$

This implies that the k-th entry in $\boldsymbol{\lambda}_1(\mathbf{\Delta}_1)$ can be used to quantify the increase in the sum-rate of User k on $\mathcal{F}$ that corresponds to a power increment of $\Delta_{1,k}$. Using a similar observation, the Lagrange dual vectors $\boldsymbol{\lambda}_i(\mathbf{\Delta}_i)$, $i = 2, 3$ can be used to quantify the additional sum-rate that each user can obtain by increasing its power budget by a small $\Delta_{i,k}$, $i = 2, 3$, on the $\mathcal{W}$ and $\mathcal{M}$ tones, respectively.

Now, for any initial power partition for which (3.3.5) is satisfied and power increment vectors $\{\mathbf{\Delta}_i\}_{i=1}^3$, a (local) maximum of total sum-rate that can be achieved on the $\mathcal{F}$, $\mathcal{W}$ and $\mathcal{M}$ tones is given by $\sum_{i=1}^3 g_i(\mathbf{\Delta}_i)$, and the the sum-rate increase that corresponds to increment vectors $\mathbf{\Delta}_1$ and $\mathbf{\Delta}_2$ is

$$\nabla_{\mathbf{\Delta}_1}\Big(\sum_{i=1}^{3} g_i(\mathbf{\Delta}_1)\Big) = \boldsymbol{\lambda}_3 - \boldsymbol{\lambda}_1, \quad \text{and} \tag{3.3.8a}$$

$$\nabla_{\mathbf{\Delta}_2}\Big(\sum_{i=1}^{3} g_i(\mathbf{\Delta}_2)\Big) = \boldsymbol{\lambda}_3 - \boldsymbol{\lambda}_2, \tag{3.3.8b}$$

respectively. Notice that in writing (3.3.8) we have used the fact that for the power partitions perturbed by $\mathbf{\Delta}_i,\ i = 1, 2, 3$, to satisfy (3.3.5), $\sum_{i=1}^{3} \mathbf{\Delta}_i = 0$.

Using (3.3.8), we can now use a standard gradient ascent algorithm to find power partitions that yields (locally) maximum total sum-rate. In order to do that, let the k-th entry of $\mathbf{P}_{\mathcal{F}}^{(\nu)}$, $\mathbf{P}_{\mathcal{W}}^{(\nu)}$ and $\mathbf{P}_{\mathcal{M}}^{(\nu)} \in \mathbb{R}_{+}^{K}$, be the ν-th iterates of the power partitioning of the k-th user on the $\mathcal{F}, \mathcal{W}$ and $\mathcal{M}$ tones, respectively, and let $\boldsymbol{\lambda}_i^{(\nu)},\ i = 1, 2, 3$ be the Lagrange dual vectors generated by the algorithms outlined in Section 3.2; viz, FDMA power allocation algorithm described in [27], the IWFA algorithm [4], and the primal-dual updates algorithm [15], respectively. The steepest ascent algorithm for updating the power partitions can now be expressed as

$$\mathbf{P}_{\mathcal{F}}^{(\nu+1)} = \mathbf{P}_{\mathcal{F}}^{(\nu)} + \mu_1(\boldsymbol{\lambda}_3 - \boldsymbol{\lambda}_1), \tag{3.3.9a}$$

$$\mathbf{P}_{\mathcal{W}}^{(\nu+1)} = \mathbf{P}_{\mathcal{W}}^{(\nu)} + \mu_2(\boldsymbol{\lambda}_3 - \boldsymbol{\lambda}_2), \tag{3.3.9b}$$

$$\mathbf{P}_{\mathcal{M}}^{(\nu+1)} = \mathbf{P} - \mathbf{P}_{\mathcal{F}}^{(\nu+1)} - \mathbf{P}_{\mathcal{W}}^{(\nu+1)}, \tag{3.3.9c}$$

where $\mathbf{P} \in \mathbb{R}_{+}^{K}$ is the vector of the power budgets of the K users in $\mathcal{K}$, and $\mu_1, \mu_2 > 0$ are two (diminishing) stepsizes. Our Structured Spectrum Balancing (SSB) can be summarized using the flow chart in Figure 3.3. As shown in this chart, the tone allocation parameters δ_1 and δ_2 in (3.2.3) and (3.2.4) are determined using a two-dimensional bisection search with convergence accuracy ϵ_2. For the steepest ascent algorithm, we used a convergence accuracy ϵ_1. Using the complexity orders given in

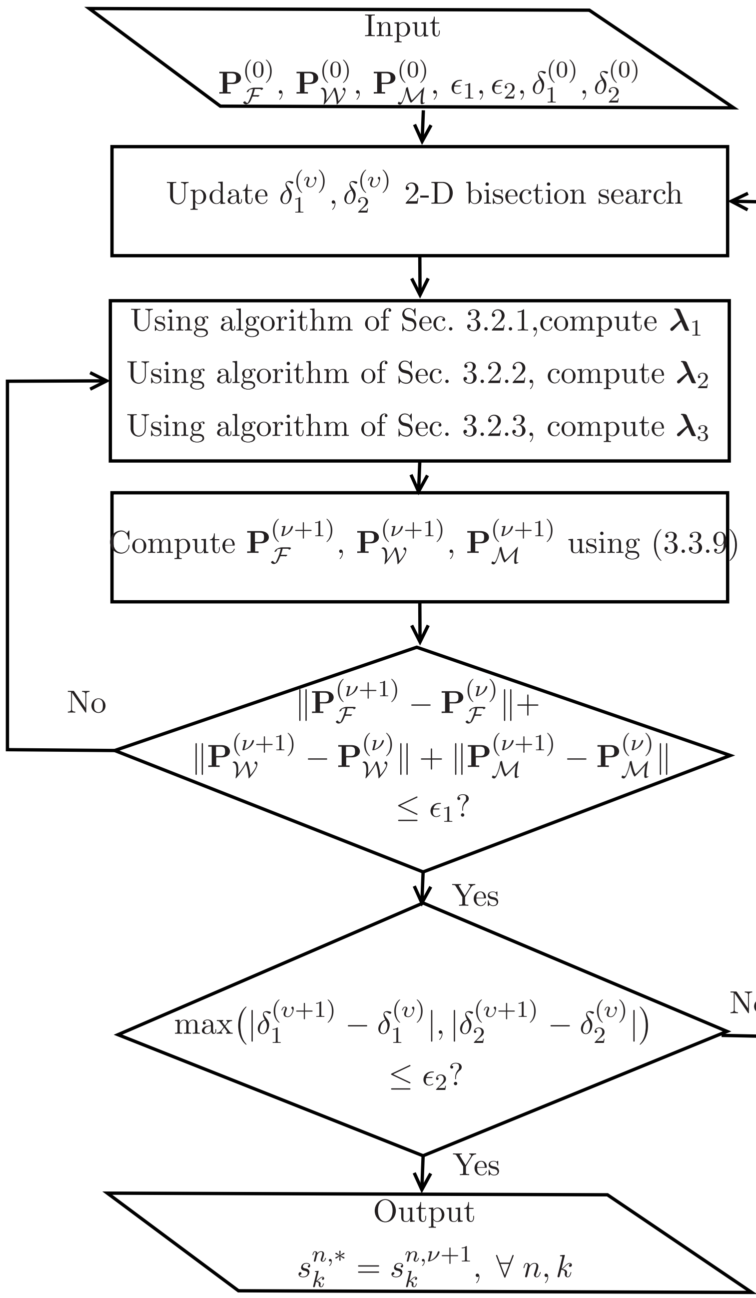


Figure 3.3: A Flow Chart of the SSB algorithm

Section 3.2, and the exponential convergence of the bisection method, one can show that the complexity of the SSB algorithm is $\mathcal{O}\big(N_P(2KN + K^2)\log^2\epsilon_2 \log^2\epsilon_1\big)$, where N_P is the maximum number of iterations of the gradient ascent algorithm.

It is worth mentioning that since the power partitioning problem is not convex, the performance of the algorithm in (3.3.9) depends, in general, on the initial power partitions, $\mathbf{P}_{\mathcal{F}}^{(0)}$, $\mathbf{P}_{\mathcal{W}}^{(0)}$ and $\mathbf{P}_{\mathcal{M}}^{(0)}$. In order to generate 'good' initial partitions, we begin by assuming that α_{jk}^n is zero for all $n \in \mathcal{N}$ and $j \neq k \in \mathcal{K}$. In this case the optimum power allocation is given by the classic water-filling technique. Denoting the power allocated by User k to the n-th tone by $s_k^{n,0}$, we choose the k-th entries of the initial power partitions to be $P_{\mathcal{F},k}^{(0)} = \sum_{n\in\mathcal{F}} s_k^{n,0}$, $P_{\mathcal{W},k}^{(0)} = \sum_{n\in\mathcal{W}} s_k^{n,0}$, and $P_{\mathcal{M},k}^{(0)} = P_k - P_{\mathcal{F},k}^{(0)} - P_{\mathcal{W},k}^{(0)}$. Our extensive numerical experiments have shown that this initialization procedure typically results in sum-rates that are close to the optimal ones achieved by the significantly more complex OSB algorithm.

3.4 Numerical Results for SSB

In this section we compare the sum-rate and the power spectral density (PSD) obtained by OSB, IWFA and SCALE with the sum-rate and PSD obtained by the proposed SSB. Due to the prohibitive computational complexity of OSB, we restrict our attention in this example to a 2-user scenario and a 256-tone DSL system. The crosstalk coefficients and spectral masks of this system were generated using a practical DSL simulator.[3] In particular, we simulated a scenario with one 5 km Central Office (CO) line and one 5 km Remote Terminal (RT) line, where the distance between the CO and the RT was taken to be 2.5 km. The overall power budget of both users was set at 20 dBm, the capacity gap, Γ, at 15, the background noise variance

[3]This simulator was provided by R. Cendrillon of Huawei Tech. Co. Ltd.

at -140 dBm/Hz, and the bit-cap at 15 bits per tone.

For this scenario, IWFA and SCALE achieve relatively low sum-rates of about 5.82 and 5.84 Mbps, respectively, whereas OSB achieves an 'optimal' rate of about 7.62 Mbps. On the other hand, the proposed SSB algorithm achieves a sum-rate of about 7.60 Mbps, which is only slightly less than the sum-rate achieved by OSB. Figure 3.4 shows the powers allocated by the four algorithms, and as can be seen from this figure, the power allocations of SSB resemble, to a large extent, those of OSB. However, these allocations vary quite significantly from the power allocations of both IWFA and SCALE.

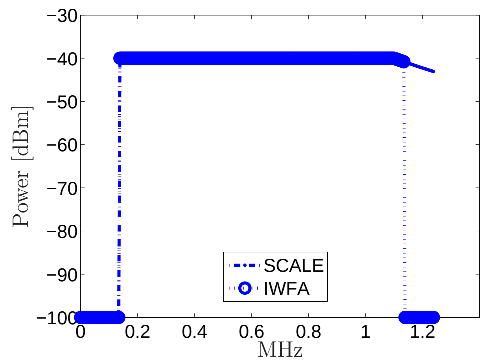
A key advantage of SSB is that it exploits the structure of optimal power allocations to avoid the exhaustive search and the discretization that underlie the OSB algorithm. In order to provide a rough comparison between the computational complexity of OSB and SSB, we measured the Matlab running time of both algorithms for the current 2-user example. For OSB this time was about 530 seconds, whereas for SSB this time was only 22 seconds.[4] This running time difference becomes more dramatic for systems with more users because the proposed SSB relies on polynomial-time algorithms that are significantly more efficient than the exhaustive search of OSB.

3.5 SSB with Pairwise Tone Assignment

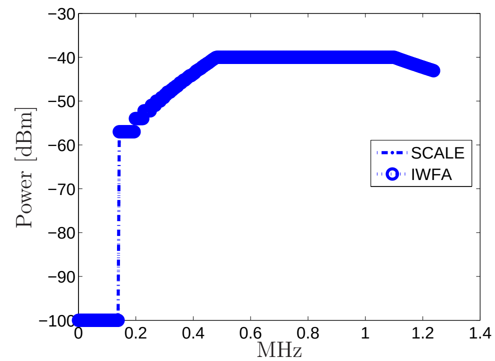
3.5.1 Pairwise Tone Assignment

In previous section, we devised an algorithm based on [28] which states that when the normalized channel gain coefficients α_{jk}^n's satisfy certain conditions users should not share tones. Loosely speaking, when the channel gain coefficients are high - that is, interference between users is strong - it is more beneficial to dedicate each tone to only

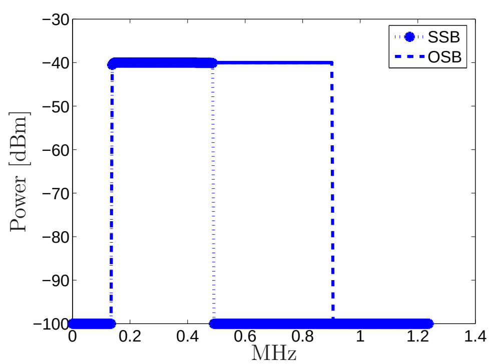
[4]The corresponding running times of IWFA and SCALE are 1 and 15 seconds, respectively.



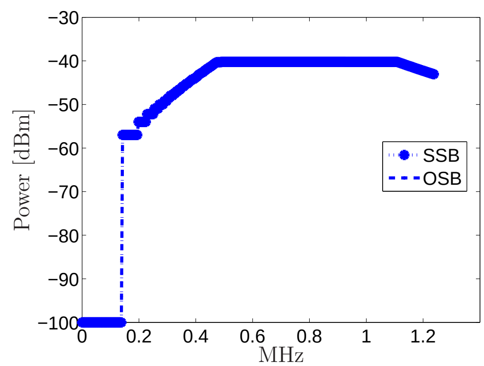
(a) IWFA and SCALE PSD of CO



(b) IWFA and SCALE PSD of RT



(c) OSB and SSB PSD of CO



(d) OSB and SSB PSD of RT

Figure 3.4: A PSD Comparison for IWFA, SCALE, OSB and SSB

one user. We utilize this theorem to assist in power allocation for the service provider in tone section $\mathcal{M}$. As described in the previous SSB, we partition the spectrum into three sections where the middle section is the section where not all the users can satisfy the conditions in either (3.2.3) or (3.2.4). This partitioning provides coarse partitioning on the tones and performs satisfactorily when this middle section is small. However, in the case where the middle section is large, i.e. the number of tones where not all the users can satisfy the conditions in either (3.2.3) or (3.2.4) is large, we should devise a scheme to refine this section. Therefore, a finer partitioning method is needed in order to enhance the performance of SSB. We propose to utilize the (3.2.3) pairwise in the middle section and refine the tone partitioning in this section, that is, further refine the assignment of tones to a certain subset of users. Specifically, for each tone n, we check the compatibility between users using (3.2.3). When (3.2.3) is satisfied for a user pair, say User j and User k, we consider these two users to be non-compatible; otherwise they can share this tone n.

We check (3.2.3) for every pair of users and determine the compatibility of these user pairs. With this knowledge of user compatibility, we run a max-clique algorithm on this set to select the largest set of compatible users (i.e. every pair of users in this set is compatible) and pre-assign tone n to these users. Then we solve the middle part of (3.1.2) only for this subset of users instead for all users. For a pair of users that satisfy (3.2.3), the theorem in [28] states that the interference among two users can be so detrimental that the system is worse off in term of system utilities when this tone is assigned to both users. This leads us to conclude that there is an advantage of tone pre-assignment using (3.2.3), that is, the separation of reciprocally damaging users can facilitate the algorithm to achieve the optimal tone assignment. We proceed to describe how to utilize an approximate max-clique algorithm to realize this tone pre-assignment.

We define

$$c^n_{jk} := \alpha^n_{jk}\alpha^n_{kj} \tag{3.5.10}$$

as the compatibility factor between User j and User k in tone n. A search algorithm will be run on each tone to derive a set of compatible users on tone n which is defined as

$$\mathcal{K}^n := \{k|c^n_{jk} < \frac{1}{4} - \delta_2, \forall j \in \mathcal{K}^n\}. \tag{3.5.11}$$

After obtaining $\mathcal{K}^n$ for all $n \in N$, we solve the tones with unstructured signalling on users $\mathcal{K}^n$ instead of all K users as proposed in Section 3.2.3.

3.5.2 Numerical Results

As another example, we compare the sum-rate of our proposed SSB under pairwise tone partitioning with that of IWFA, gradient search, and SCALE in a 6-user scenario.[5] For this example we use similar parameters to those used in the previous example, and we generate the crosstalk coefficients using the same DSL simulator, but for 2 co-located CO's and 4 Rt's. The lengths of the CO lines were chosen to be 5 and 4 km and those of the RT lines were chosen to be 5, 5, 4, and 4 km, respectively. The distance between the CO's and the RT's was chosen to be 0.2, 0.2, 3 and 3 km, respectively. For this scenario, IWFA, gradient search, and SCALE could achieve sum-rates of only 13.0 Mbps, 13.1 Mbps, and 14.1 Mbps, respectively, whereas SSB could achieve a sum-rate of 16.7 Mbps.

[5]The computational complexity of OSB has made it rather difficult for us to provide a sum-rate comparison for this example.

Chapter 4

Power Allocation for Harmonic Mean Objective

4.1 Maximizing "User Capacity" with Per-tone Harmonic Mean Approximation

Consider the situation in which the service provider wishes to maximize the number of users that the DSL system can support. For the service provider to do that, it may maximize the sum rate of all users, which is given by $\sum_{k=1}^{K} \sum_{n=1}^{N} R_k^n(s_1^n, \ldots, s_K^n)$ (Section 3). However, such an approach may result in power allocations that favor strong users over weaker ones. As an alternative, the service provider may consider a more balanced approach in which the objective is to maximize a weighted sum rate, $\sum_{k=1}^{K} w_k \sum_{n=1}^{N} R_k^n(s_1^n, \ldots, s_K^n)$ where the weights, $\{w_k\}_{k=1}^{K}$ are assigned in such a way that favors weak users over stronger ones. The drawback of this approach is that the way in which the weights should be assigned depends on the channel gains and the power budget in a non-linear fashion. Therefore, in this section, we consider maximization of the number of users that can be accommodated by the

DSL system. The users belong to different categories depending on the QoS that they purchase from the system provider. This design objective is integer-valued and hence generally difficult to handle directly. As an alternative, we consider a design problem in which we maximize the (weighted) harmonic mean of the users' rates. This objective possesses two desirable features: first, from the users' perspective, this objective is known to be fairer than maximizing the sum-rate [10]; second, for single-tone systems, maximizing this objective can be cast as a convex optimization problem for which the global solution can be obtained efficiently. The first feature renders the harmonic mean a natural design objective for maximizing the number of users. This is because for one to be able to compare the number of users that can be supported by two systems, one should guarantee that the users obtain the same service in both systems. The second feature, on the other hand, enables us to design an efficient algorithm that exploits the per-tone convexity to provide power allocations that yield relatively high harmonic mean-rates. The harmonic mean-rate can be written as

$$H(\mathbf{s}_0, \mathbf{s}_1, \cdots, \mathbf{s}_K) = \left(\sum_{k=1}^{K} \left(\sum_{n=1}^{N} R_k^n \right)^{-1} \right)^{-1}. \tag{4.1.1}$$

Now, in order to find the power allocations that maximize this objective, we ought to solve the following optimization problem:

$$\begin{aligned} \min \quad & 1/H(\mathbf{s}_1, \cdots, \mathbf{s}_K), && \text{(4.1.2a)} \\ \text{subject to} \quad & \sum_{n=1}^{N} s_k^n \leq P_k, \forall\, k, && \text{(4.1.2b)} \\ & 0 \leq s_k^n \leq S_{\max,k}^n, \forall\, k, && \text{(4.1.2c)} \end{aligned}$$

where in (4.1.2a), we have used the fact that maximizing $H(\mathbf{s}_1, \cdots, \mathbf{s}_K)$ is equivalent to minimizing $1/H(\mathbf{s}_1, \cdots, \mathbf{s}_K)$. We have also used $\mathbf{s}_k$ to denote the vector

$[s_k^1, \cdots, s_k^N]^T$, P_k to denote the total power budget of User k, and $S^n_{\max,k}$ to denote the maximum signal power that User k can allocate to the n-th tone. In order for (5.0.3a) not to be redundant, we assume that $P_k \leq \sum_{n=1}^N S^n_{\max,k}$. Although it is desirable from the service provider perspective to be able to solve (4.1.2), for $N > 1$ this problem is known to be NP-hard [10], and hence difficult to solve in a computationally-efficient manner. As an alternative, in the next section we derive an upper bound on the objective in (4.1.2a). Unlike the original problem in (4.1.2), this upper bound is convex and hence can be minimized using highly efficient numerical techniques.

4.1.1 Harmonic Mean Reformulation

It is stated in [10] that when there is only one tone in the network (i.e. $N = 1$), (4.1.2a) is a convex optimization problem after transformation. Inspired by this claim, we derive an upper bound for the objective function of (4.1.2a) by using the convexity of the function $f(x) = \frac{1}{x}$ and Jensen's inequality. Specifically,

$$\begin{aligned} H_{\mathrm{U}B}(\mathbf{s}_0, \mathbf{s}_1, \cdots, \mathbf{s}_K) &= \sum_{k=1}^{K} \sum_{n=1}^{N} (R_k^n)^{-1} \\ &\geq \sum_{k=1}^{K} \left(\sum_{n=1}^{N} R_k^n \right)^{-1} . \end{aligned}$$

Note that $H_{\mathrm{U}B}$ not only is an upper bound on the harmonic mean defined as H but also can be interpreted as the inverse of the harmonic mean of the per tone rates of users while H is the harmonic mean of the rates of users. Furthermore, as will be shown later, by using $H_{\mathrm{U}B}$ as the objective function, the optimization problem becomes convex. Another benefit of using $H_{\mathrm{U}B}$ is that the optimization problem can be decomposed across tones via its dual form, which enables us to use efficient convex optimization techniques to solve the problem and achieve the goal of accommodating users with certain prescribed rate ranges. By employing the upper bound of H, we consider the following optimization problem:

$$\begin{aligned} \min \quad & H_{\mathrm{U}B}(\mathbf{s}_1, \cdots, \mathbf{s}_K), \\ \text{s.t.} \quad & \sum_{n=1}^{N} s_k^n \leq P_k, \forall\, k, \\ & 0 \leq s_k^n \leq S_{\max,k}^n, \forall\, k, \forall\, n. \end{aligned} \tag{4.1.3}$$

We proceed to show that (4.1.3) is a convex optimization problem. Consider the following transformation [10]:

$$t_k^n = \left(\log \left(1 + \frac{s_k^n}{\sigma_k^n + \sum_{j \neq k} \alpha_{jk}^n s_j^n} \right) \right)^{-1}; \tag{4.1.4}$$

$$y_k^n = \log s_k^n. \tag{4.1.5}$$

(4.1.3) can be rewritten as

$$\begin{aligned} \min \quad & \sum_{k=1}^{K} \sum_{n=1}^{N} t_k^n, \\ \text{s.t.} \quad & y_k^n \leq \log(S_{\max,k}^n),\ t_k^n \geq 0, \forall\, k,\ \forall\, n, \\ & \sum_{n=1}^{N} 2^{y_k^n} \leq P_k, \forall\, k, \\ & \log \left(\sigma_k^n 2^{(-y_k^n)} + \sum_{j \neq k} \alpha_{jk}^n 2^{(y_j^n - y_k^n)} \right) + \log(2^{\frac{1}{t_k^n}} - 1) \leq 0, \forall\, k,\ \forall\, n. \end{aligned} \tag{4.1.6}$$

Using the observation in [10], it can be directly derived that (4.1.6) is convex, and hence, for small-to-moderate numbers of users and tones, it can be solved using an efficient convex optimization algorithm [15]. Even though (4.1.3) is convex, it is coupled across tones. The computational complexity is still prohibitive for real-time power allocation when the number of tones is high. Fortunately, the problem can be decomposed into several per-tone convex optimization problems in its dual form, hence we can solve it in a more efficient fashion. The dual problem of (4.1.3) can be written as follows

$$
\begin{aligned}
d(\lambda) = & \min_{0 \le s_k^n \le S_{\max,k}^n,\ \forall\, n,\ \forall\, k} \left[\sum_{k=1}^{K} \sum_{n=1}^{N} \left(\log\left(1 + \frac{s_k^n}{\sigma_k^n + \sum_{j \ne k} \alpha_{jk}^n s_j^n}\right)\right)^{-1} + \sum_{k=1}^{K} \lambda_k \sum_{n=1}^{N} s_k^n \right] \\
& - \sum_{k=1}^{K} \lambda_k P_k, \\
= & \sum_{n=1}^{N} \min_{0 \le s_k^n \le S_{\max,k}^n,\ \forall\, n,\ \forall\, k} \left[\sum_{k=1}^{K} \left(\log\left(1 + \frac{s_k^n}{\sigma_k^n + \sum_{j \ne k} \alpha_{jk}^n s_j^n}\right)\right)^{-1} + \sum_{k=1}^{K} \lambda_k s_k^n \right] \\
& - \sum_{k=1}^{K} \lambda_k P_k. \qquad (4.1.7)
\end{aligned}
$$

It is worthwhile to point out that the optimization in (4.1.7) can be solved on a per-tone basis due to the dual decomposition of the original problem. To solve the minimization in (4.1.7), we need to solve the following minimization problem for each n:

$$
\begin{aligned}
\min \quad & \sum_{k=1}^{K} \left(\log\left(1 + \frac{s_k^n}{\sigma_k^n + \sum_{j \ne k} \alpha_{jk}^n s_j^n}\right)\right)^{-1} + \sum_{k=1}^{K} \lambda_k s_k^n \qquad (4.1.8) \\
\text{s.t.} \quad & 0 \le s_k^n \le S_{\max,k}^n,\ \forall\, k.
\end{aligned}
$$

Under the above transformation, the equivalent optimization problem of (4.1.8) after transformation for each n is

$$\begin{aligned} \min \quad & \sum_{k=1}^{K} t_k^n + \sum_{k=1}^{K} \lambda_k 2^{y_k^n} \\ \text{s.t.} \quad & y_k^n \le \log(S_{\max,k}^n),\ t_k^n \ge 0, \forall\, k, \\ & \log\left(\sigma_k^n 2^{(-y_k^n)} + \sum_{j \ne k} \alpha_{jk}^n 2^{(y_j^n - y_k^n)}\right) + \log(2^{\frac{1}{t_k^n}} - 1) \le 0, \forall\, k. \end{aligned} \tag{4.1.9}$$

Compare (4.1.6) and (4.1.9), we conclude that the above optimization problem is also convex and can be solved by existing convex optimization techniques such as interior point method.

4.1.2 QoS: Max Rate and Min Rate Constraints

From a viewpoint of a DSL service provider, the targets of system design are to maximize the system capacity and enhance users' satisfaction level. However, there is usually a trade-off between these two goals. Therefore, to have control over the trade-off as well as to facilitate an agreement on the quality of provided DSL service between users and system providers, we propose to impose a rate range constraint that consists of a maximum rate and a minimum rate in the optimization problem (4.1.3). A maximum rate constraint is to suppress the power used by stronger users in a network on one hand and to encourage users to subscribe to higher end service when needed. A minimum rate constraint can ensure that the users achieve rates that have been contracted, which can be viewed as a mean for a service provider to guarantee quality-of-service. That fact that the maximum and the minimum rate comprise the prescribed rate range that a user and a system provider negotiate is essential in a commercial DSL service contract since users and system providers usually should agree on rates that are bilaterally beneficial. If we denote the low and high endpoints by $R_{\min}$ and $R_{\max}$, respectively, the rate range over which User $k \in \mathcal{K}$ operates can be expressed as the set of rates that satisfy the following constraints:

$$\sum_{n=1}^{N} R_k^n \leq R_{\max,k}$$
$$\sum_{n=1}^{N} R_k^n \geq R_{\min,k}$$

Considering these two constraints, (4.1.3) can be rewritten as

$$\begin{aligned}
\min_{\mathbf{s}} \quad & \sum_{k=1}^{K}\sum_{n=1}^{N}\left(\log\left(1+\frac{s_k^n}{\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n s_j^n}\right)\right)^{-1} \qquad (4.1.10)\\
\text{s.t.} \quad & 0 \leq s_k^n \leq S_{\max,k}^n,\ \forall\, n,\ \forall\, k\\
& \sum_{n=1}^{N} s_k^n \leq P_k,\ \forall\, k\\
& \sum_{n=1}^{N} \log\left(1+\frac{s_k^n}{\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n s_j^n}\right) \leq R_{\max,k}\ \forall\, k,\\
& \sum_{n=1}^{N} \log\left(1+\frac{s_k^n}{\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n s_j^n}\right) \geq R_{\min,k}\ \forall\, k,
\end{aligned}$$

where $R_{\max,k}$ and $R_{\min,k}$ are the maximum rate and minimum rate respectively that the service provider agrees to provide to User k, respectively. By imposing this maximum rate constraint so as to suppress the inter-user interference, the system can accommodate more users as desired while the min rate constraint serves to ensure QoS. While we can cast the maximum rate constraint into a convex function as part of the objective function which will be shown below, the minimum rate constraint cannot be transformed in a similar fashion. However, we can find a lower bound of the left hand side of the minimum rate constraint that is convex under the transformation in (4.1.4) and (4.1.5). Hence, we start by rewriting the left hand side of the minimum rate constraint as the following:

$$\begin{aligned}
&\sum_{n=1}^{N} \log\left(1 + \frac{s_k^n}{\sigma_k^n + \sum_{j\neq k} \alpha_{jk}^n s_j^n}\right) \\
&= \sum_{n=1}^{N} \left(\log\left(\sigma_k^n + \sum_{j=1}^{K} \alpha_{jk}^n s_j^n\right) - \log\left(\sigma_k^n + \sum_{j\neq k} \alpha_{jk}^n s_j^n\right)\right). \qquad (4.1.11)
\end{aligned}$$

We have

$$\begin{aligned}
&\log\left(\sigma_k^n + \sum_{j=1}^{K} \alpha_{jk}^n s_j^n\right) \qquad (4.1.12)\\
&= \log\left(1 + \sum_{j=1}^{K} \alpha_{jk}^n\right) + \log\left(\frac{\sigma_k^n}{1 + \sum_i \alpha_{ik}^n} + \sum_{j=1}^{K} \frac{\alpha_{jk}^n}{1 + \sum_i \alpha_{ik}^n} s_j^n\right) \\
&\geq \log\left(1 + \sum_{j=1}^{K} \alpha_{jk}^n\right) + \frac{1}{1 + \sum_i \alpha_{ik}^n} \log \sigma_k^n + \sum_{j=1}^{K} \frac{\alpha_{jk}^n}{1 + \sum_i \alpha_{ik}^n} \log s_j^n.
\end{aligned}$$

The inequality in (4.1.12) is due to the convexity of $-\log$. Consequently, we have the following inequality that is readily derived from (4.1.12):

$$\begin{aligned}
&\sum_{n=1}^{N} \log\left(1 + \frac{s_k^n}{\sigma_k^n + \sum_{j\neq k} \alpha_{jk}^n s_j^n}\right) \qquad (4.1.13)\\
&\geq \sum_{n=1}^{N} \left(\log\left(1 + \sum_{j=1}^{K} \alpha_{jk}^n\right) + \frac{1}{1 + \sum_i \alpha_{ik}^n} \log \sigma_k^n + \sum_{j=1}^{K} \frac{\alpha_{jk}^n}{1 + \sum_i \alpha_{ik}^n} \log s_j^n \right.\\
&\left. - \log\left(\sigma_k^n + \sum_{j\neq k} \alpha_{jk}^n s_j^n\right)\right). \qquad (4.1.14)
\end{aligned}$$

Note that it can be seen that for s_k^n's that make the right hand side of (4.1.13) greater than $R_{\min,k}$ we have $\sum_{n=1}^{N} \log\left(1 + \frac{s_k^n}{\sigma_k^n + \sum_{j\neq k} \alpha_{jk}^n s_j^n}\right) \geq R_{\min,k}$. Therefore, we consider the following optimization problem:

$$
\begin{aligned}
\min \quad & \sum_{k=1}^{K}\sum_{n=1}^{N}\left(\log\left(1+\frac{s_k^n}{\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n s_j^n}\right)\right)^{-1} && (4.1.15)\\
\text{s.t.} \quad & 0 \le s_k^n \le S_{\max,k}^n,\ \forall\, k,\ \forall\, n,\\
& \sum_{n=1}^{N} s_k^n \le P_k,\ \forall\, k,\\
& \sum_{n=1}^{N}\log\left(1+\frac{s_k^n}{\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n s_j^n}\right) \le R_{\max,k}\ \forall\, k,\\
& \sum_{n=1}^{N}\left(-\log\left(1+\sum_{j=1}^{K}\alpha_{jk}^n\right)-\frac{1}{1+\sum_i \alpha_{ik}^n}\log\sigma_k^n\right.\\
& \quad \left. -\sum_{j=1}^{K}\frac{\alpha_{jk}^n}{1+\sum_i\alpha_{ik}^n}\log s_j^n+\log\left(\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n s_j^n\right)\right) \le -R_{\min,k}\ \forall\, k.
\end{aligned}
$$

From (4.1.15), the dual function can be written as

$$
\begin{aligned}
d(\lambda) := & -\sum_{k=1}^{K}\lambda_k P_k - \sum_{k=1}^{K}\mu_k R_{\max,k}\\
& +\sum_{k=1}^{K}\varsigma_k\left(R_{\min,k}+\sum_{n=1}^{N}\left(-\log\left(1+\sum_{j=1}^{K}\alpha_{jk}^n\right)-\frac{1}{1+\sum_i\alpha_{ik}^n}\log\sigma_k^n\right)\right)\\
& +\sum_{n=1}^{N}\min_{0\le s_k^n\le S_{\max,k}^n,\ \forall\, n,\ \forall\, k}\left[\sum_{k=1}^{K}\left(\log\left(1+\frac{s_k^n}{\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n s_j^n}\right)\right)^{-1}\right.\\
& +\sum_{k=1}^{K}\lambda_k s_k^n+\sum_{k=1}^{K}\mu_k\log\left(1+\frac{s_k^n}{\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n s_j^n}\right)\\
& \left.+\sum_{k=1}^{K}\varsigma_k\left(-\sum_{j=1}^{K}\frac{\alpha_{jk}^n}{1+\sum_i\alpha_{ik}^n}\log s_j^n+\log\left(\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n s_j^n\right)\right)\right]. \qquad (4.1.16)
\end{aligned}
$$

Hence, for each n, the minimization problem becomes the following:

$$\min \sum_{k=1}^{K}\left(\log\left(1+\frac{s_k^n}{\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n s_j^n}\right)\right)^{-1}+\sum_{k=1}^{K}\lambda_k s_k^n+\sum_{k=1}^{K}\mu_k\log\left(1+\frac{s_k^n}{\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n s_j^n}\right)$$
$$+\sum_{k=1}^{K}\varsigma_k\left(-\sum_{j=1}^{K}\frac{\alpha_{jk}^n}{1+\sum_i\alpha_{ik}^n}\log s_j^n+\log\left(\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n s_j^n\right)\right) \tag{4.1.17}$$
$$\text{s.t. } 0\leq s_k^n\leq S_{\max,k}^n,\ \forall\, k.$$

Note that μ_k and ς_k's are non-negative. We apply the transformation (4.1.4) and (4.1.5) and obtain the following

$$\min \sum_{k=1}^{K}\left(t_k^n+\mu_k(t_k^n)^{-1}\right)+\sum_{k=1}^{K}\lambda_k 2^{y_k^n}$$
$$+\sum_{k=1}^{K}\varsigma_k\left(-\sum_{j=1}^{K}\frac{\alpha_{jk}^n}{1+\sum_i\alpha_{ik}^n}y_j^n+\log\left(\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n 2^{y_j^n}\right)\right) \tag{4.1.18}$$
$$\text{s.t. } y_k^n\leq\log(S_{\max,k}^n),\ t_k^n\geq 0,\forall\, k,$$
$$\log\left(\sigma_k^n 2^{(-y_k^n)}+\sum_{j\neq k}\alpha_{jk}^n 2^{(y_j^n-y_k^n)}\right)+\log(2^{\frac{1}{t_k^n}}-1)\leq 0,\forall\, k.$$

It can also be verified that (4.1.18) is a convex optimization problem. We can solve (4.1.18) in a similar fashion to the way we solved (4.1.3).

So far we have derived the dual form of (4.1.3) under which the problem decomposes into a series of per-tone optimization problems (4.1.9). We further transform (4.1.9) to its convex form (4.1.18). We will complete this section by summarizing the algorithm to solve (4.1.3).

4.1.3 Per-tone Harmonic Mean Rate Optimization Algorithm

The procedure of solving (4.1.3) consists of two loops. The outer loop is a Lagrange dual update procedure [15] while the inner loop is to solve (4.1.18) for each tone for each $k \in \mathcal{K}$. This procedure, per-tone Harmonic Mean Rate Optimization (HMRO), can be described in detail as follows:

- Step 0: Initialize primal variables $\{s_k^{n,0}\}$ and dual variables $\{\mu_k^0,\ \lambda_k^0,\ \varsigma_k^0\}$.
- Step 1: Inner loop ν: For each n, solve for

$$
\begin{aligned}
s^{(\nu)} = & \sum_{n=1}^{N} \text{argmin}_{0 \le s_k^n \le S_{\max,k}^n,\, \forall\, n,\, \forall\, k} \sum_{k=1}^{K} \left(\log \left(1 + \frac{s_k^n}{\sigma_k^n + \sum_{j \neq k} \alpha_{jk}^n s_j^n} \right) \right)^{-1} \\
& + \sum_{k=1}^{K} \lambda_k s_k^n + \sum_{k=1}^{K} \mu_k \log \left(1 + \frac{s_k^n}{\sigma_k^n + \sum_{j \neq k} \alpha_{jk}^n s_j^n} \right) \\
& + \sum_{k=1}^{K} \varsigma_k \left(- \sum_{j=1}^{K} \frac{\alpha_{jk}^n}{1 + \sum_i \alpha_{ik}^n} \log s_j^n + \log \left(\sigma_k^n + \sum_{j \neq k} \alpha_{jk}^n s_j^n \right) \right)
\end{aligned}
$$

  by the following steps:

  - Step 1.1: Initialize

$$
\begin{aligned}
t_k^{n,\nu} &= \left(\log \left(1 + \frac{s_k^{n,\nu-1}}{\sigma_k^n + \sum_{j \neq k} \alpha_{jk}^n s_j^{n,\nu-1}} \right) \right)^{-1}; \\
y_k^{n,\nu} &= \log s_k^{n,\nu-1}.
\end{aligned}
$$

– Step 1.2: Solve for $t_k^{n,\nu}$, $y_k^{n,\nu}$ using interior point method [29]:

$$
\begin{aligned}
\min_{t_k^{n,\nu}, y_k^{n,\nu}\ \forall k\in\mathcal{K}} \ & \sum_{k\in\mathcal{K}} \big(t_k^{n,\nu} + \mu_k (t_k^{n,\nu})^{-1}\big) + \sum_{k\in\mathcal{K}} \lambda_k 2^{y_k^{n,\nu}} \\
& + \sum_{k\in\mathcal{K}} \varsigma_k \left(-\sum_{j=1}^{K} \frac{\alpha_{jk}^n}{1+\sum_i \alpha_{ik}^n} y_j^{n,\nu} + \log\left(\sigma_k^n + \sum_{j\neq k} \alpha_{jk}^n 2^{y_j^{n,\nu}}\right)\right) \\
\text{s.t. } & y_k^{n,\nu} \le \log(S_{\max,k}^n),\ t_k^n \ge 0, \forall\, k\in\mathcal{K}, \\
& \log\left(\sigma_k^n 2^{(-y_k^{n,\nu})} + \sum_{j\neq k} \alpha_{jk}^n 2^{(y_j^{n,\nu} - y_k^{n,\nu})}\right) + \log(2^{\frac{1}{t_k^{n,\nu}}} - 1) \le 0, \forall\, k\in\mathcal{K}.
\end{aligned}
$$

– Step 1.3: Let $s_k^{n,\nu} = 2^{y_k^{n,\nu}}$, $\forall k \in \mathcal{K}$.

- Step 2: For all k and n, calculate

$$
R_k^{n,\nu} = \log\Big(1 + \frac{s_k^{n,\nu}}{\sigma_k^n + \sum_{j\neq k} \alpha_{jk}^n s_j^{n,\nu}}\Big).
$$

- Step 3: Apply steepest decent on dual variables

$$
\begin{aligned}
\lambda_k^\nu =& \lambda_k^{\nu-1} + \beta_\lambda^\nu (\sum_{k=1}^{K} s_k^{n,\nu} - P_k), \\
\mu_k^\nu =& \mu_k^{\nu-1} + \beta_\mu^\nu (\sum_{n=1}^{N} R_k^{n,\nu} - R_{\max,k}), \\
\varsigma_k^\nu =& \varsigma_k^{\nu-1} + \beta_\varsigma^\nu (R_{\min,k} - \sum_{n=1}^{N} R_k^{n,\nu}),
\end{aligned}
$$

where β_λ^ν, β_μ^ν and β_ς^ν are the step sizes for the λ, β and ς, respectively. For a discussion on the choices of step size, one can refer to [15] and [28].

- Step 4: Go to Step 1 until the optimality condition is satisfied.

Since (4.1.15) is convex, the proposed algorithm converges to a fixed point of the problem, which hence is a minimizer of (4.1.15) due to its convexity. Furthermore, from (4.1.13) we know that any minimizer of (4.1.15) is feasible for (4.1.10).

4.1.4 Simulation Results

Simulation Case 1:

In this example, we compare the performance of HMRO, SSB (Section 3, [30]), SCALE [12], and IWFA [4] [1] in terms of the achieved harmonic mean and the number of users in prescribed rate range for a six user DSL communication system with two hundred and fifty-six tones. For this system, the cross-talk coefficients, the noise variances and spectrum masks were generated by a DSL simulation program[2]. The system model consists of two CO lines and four RT lines, the topology of which are shown in Fig. 4.1. Other system parameters are set as follows: noise variance is -140 dBm/Hz, power budget is 20 dBm, and bit-cap is 15 bits for all users. We have two groups of users: one subscribes to basic service and the other consists of high-end users. We let User 1 and User 2 be in first group and User 3, User 4, User 5 and User 6 be in the second group. The two groups of users have prescribed rate ranges of $(0.5, 2)$(Mbps) and $(2, 10)$(Mbps) corresponding to the basic service rate range and high-end residential rate range, respectively. As a consequence, we set $R_{\max} = [2, 2, 10, 10, 10, 10]$(Mbps) and $R_{\min} = [0.5, 0.5, 2, 2, 2, 2]$(Mbps). Furthermore, we examine four sets of weights for SSB, SCALE and IWFA that are numbered as Case 1 to Case 4 in Table 4.1 since the objective of these three algorithms are weighted sum-rate. It is worthwhile to point out that we do not consider weights for HMRO since harmonic mean objective ”automatically” provides relatively fair rates to all users and the max and min rate constraints enable the system to provide contracted rate range without adjusting the

[1] Due to the high complexity of OSB and unknown data of ASB, we do not run simulation on OSB and ASB.

[2] The program is provided by Raphael Cendrillon of Huawei Technologies Co. Ltd.

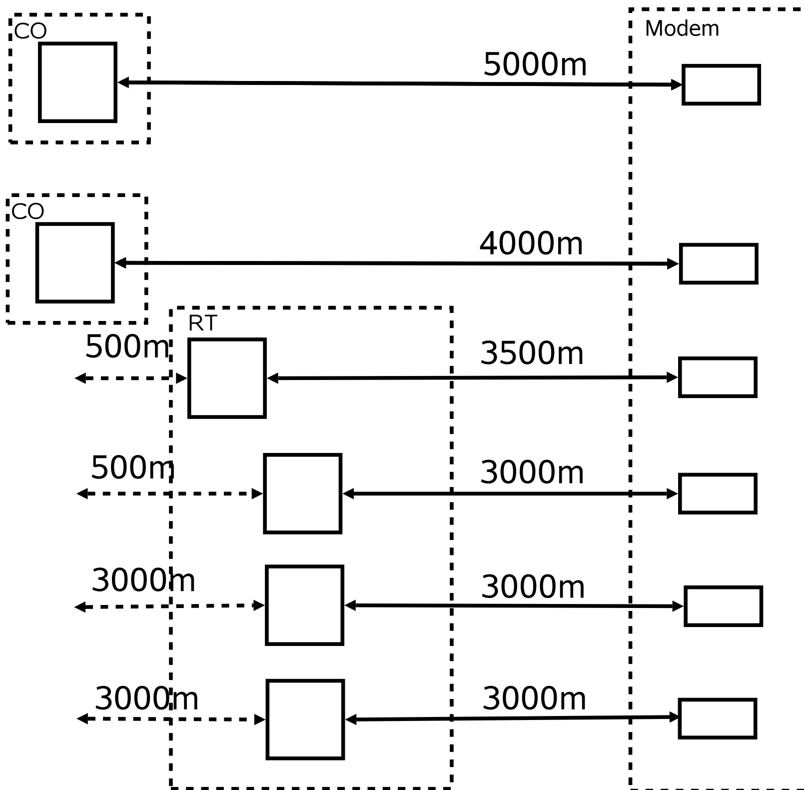


Figure 4.1: A 6 User DSL System Topology

weights on users.

The programs were coded in MATLAB 7.0 and the computational time for HMRO, SSB, SCALE and IWFA to run one of the six users case is 65 sec, 85 seconds, 47 seconds and 10 seconds. Figure 4.3, Figure 4.4 and Figure 4.5 are the individual rates achieved by the users. It is shown in Figure 4.3, Figure 4.4 and Figure 4.5 that among the four methods, HMRO is the only method that can fulfill the desired rate range

Case Number	Weights
Case 1	$[0.2, 0.1, 0.2, 0.2, 0.2, 0.1]$
Case 2	$[0.1, 0.2, 0.2, 0.2, 0.2, 0.1]$
Case 3	$[0.2, 0.1, 0.2, 0.2, 0.1, 0.2]$
Case 4	$[0.1, 0.2, 0.2, 0.2, 0.1, 0.2]$

Table 4.1: Weights for 6 Users Cases

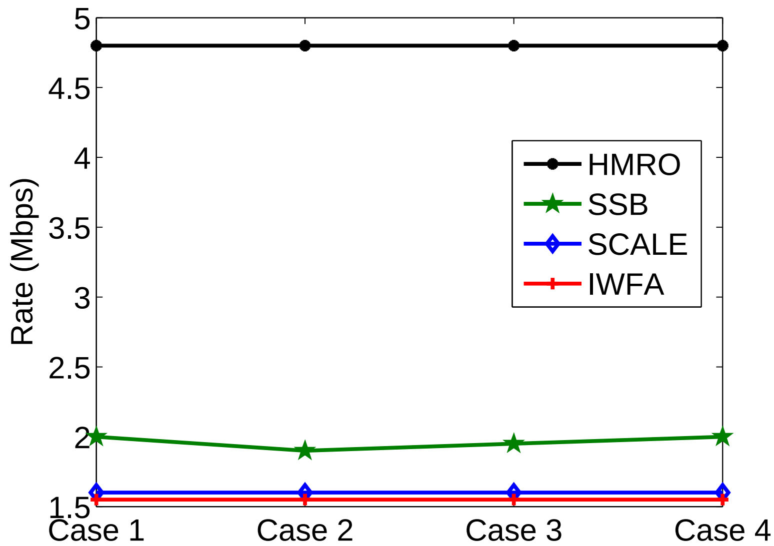


Figure 4.2: Harmonic Mean Rate Achieved by HMRO, SSB, SCALE and IWFA in 6 User Case

for all users in both groups. SSB, SCALE and IWFA can only guarantee service for the group of basic service subscribers. Table 4.2 lists the number of users that can supported by the four algorithms according to Figure 4.3, Figure 4.4 and Figure 4.5.

As can see from Table 4.2, HMRO can achieve the targeted rate range for users at all weight cases while SSB, SCALE and IWFA fail to fulfill the rate requirement in any cases. In fact, this can be viewed as an advantage of HMRO that it does

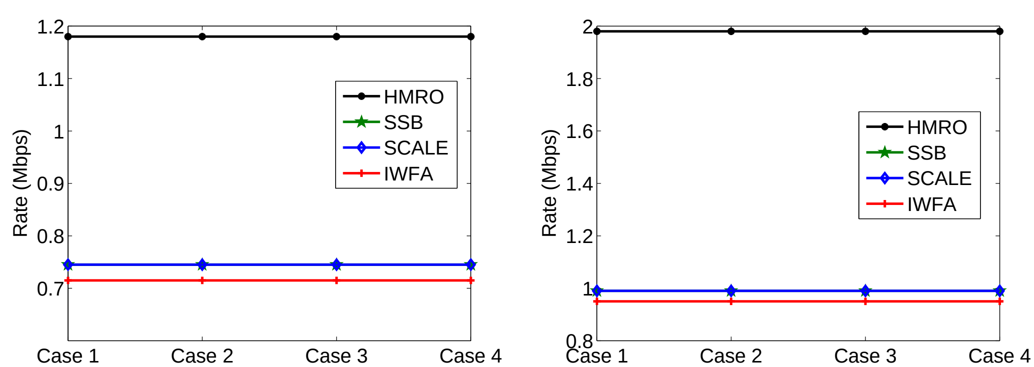


Figure 4.3: User 1 and User 2 Rate in 6 User Case

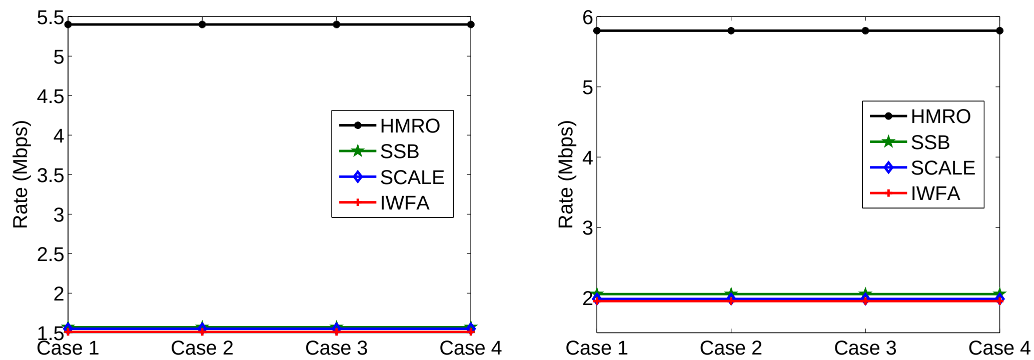


Figure 4.4: User 3 and User 4 Rate in 6 User Case

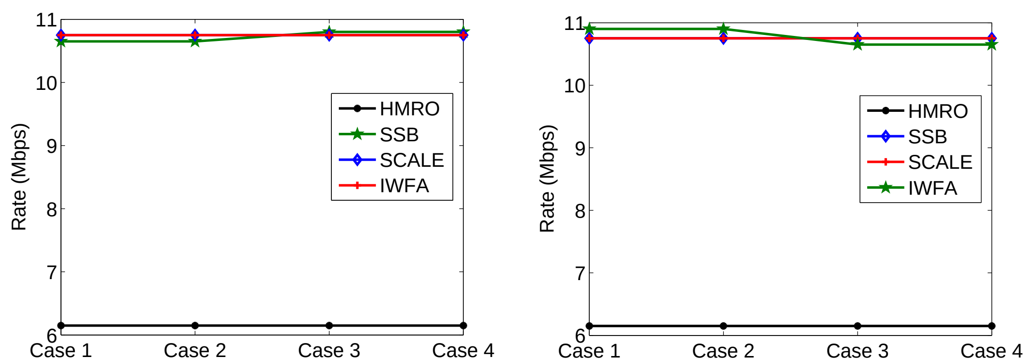


Figure 4.5: User 5 and User 6 Rate in 6 User Case

Number of users	Case 1	Case 2	Case 3	Case 4
HMRO	6	6	6	6
SSB	3	3	3	3
SCALE	2	2	2	2
IWFA	2	2	2	2

Table 4.2: Number of Users in Prescribed Rate Range in 6 User Cases

not require the service provider to blindly adjust weights but change the max and min rate constraints when the service requirement changes. This gives the service provider a better control over the quality of service by a guarantee of achieving users' prescribed rates.

Simulation Case 2:

In this example we consider a DSL communication system with 256 tones. Assuming that there are seven users that we wish to accommodate in the DSL system, the crosstalk coefficients and the noise parameters of these users were generated using a practical DSL simulator 1. The system model consists of 2 Central Office (CO) and 5 Remote Terminal (RT) lines, and all users are assumed to have identical power budgets. The lengths of the CO and RT lines are 5, 4, 3.5, 3.5, 3, 3 and 3 km, respectively, and the distances from the 5 RTs to the COs are set to be 0.3, 0.5, 0.5, 3 and 3 km, respectively. The background noise variance is assumed to be $N_0 = -140$ dBm/Hz and the capacity gap is set to be 15 dB. The users are divided into basic and high-end service groups. The basic service group consists of Users 1 and 2 with $R_{\min,k} = 0.5$ Mbps and $R_{\max,k} = 2$ Mbps, $k = 1, 2$, and the high-end service group consists of Users 3 to 7 with $R_{\min,k} = 2$ Mbps and $R_{\max,k} = 12$ Mbps, $k = 3, \ldots, 7$. Using these parameters, in Figure 4.6 we compare the number of users supported by SCALE and IWFA and the number of users supported by the proposed HMRO algorithm. From this figure it can be seen that for the considered range of power budgets, both SCALE and IWFA support fewer users than the proposed HMRO. For instance, SCALE can only support 4 users (Users 1 and 5 to 7) throughout the entire range of the considered power budgets. However, IWFA exhibits a more interesting behavior. At an input power of 11 dBm, IWFA supports up to 6 users, but as the power increases, IWFA tends to favor stronger users (Users 6 and 7 in the

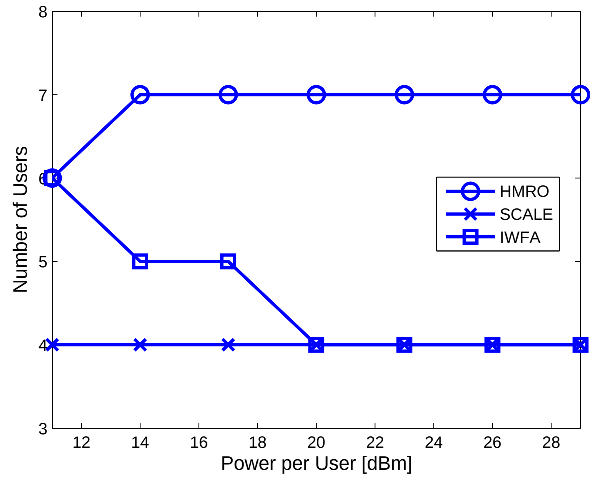


Figure 4.6: Number of Users Supported by HMRO, SCALE,IWFA

current example) over weaker ones. This tendency eventually incurs a decrease in the number of users that the system can support. Finally, we consider the number of users that can be supported by the proposed HMRO. For an input power of 11 dBm, similar to IWFA, this algorithm supports 6 users. However, by increasing the power budget, HMRO manages to accommodate all 7 users in the system. This performance advantage follows from the inherent fairness of the harmonic mean objective and the versatility with which the system designer can control the QoS of different classes of users.

Finally, we compare the complexity of SCALE and IWFA with that of HMRO. Denoting the tolerance by ϵ, the complexity of IWFA and SCALE can be shown to be $\mathcal{O}(KN\log^2\epsilon)$ and $\mathcal{O}(KNL\log^2\epsilon)$, where L is the number of SCALE updates ([31]), whereas the complexity of HMRO is $\mathcal{O}(KNlog^2\epsilon)$. As a rough comparison, the average Matlab running times of the IWFA and SCALE for the scenario considered in this example are about 1 and 15 seconds, respectively, whereas that of HMRO is about 26 seconds. Hence, it can be seen that the computational complexity of HMRO is comparable to that of IWFA and SCALE, but it can support significantly more users than either of the other algorithms.

4.2 Solving Harmonic Mean Objective via Decomposition and Convexification

As stated in the previous section, a traditional objective is to maximize the total achievable data rate of the system (i.e., the sum of users' data rate $= \sum_{k=1}^{K} R_k$, [30], [32] and [1]). However, in this case some users may have dominant achievable data rates while others suffer from low achievable data rates. This is not desirable in practice from a service provider's perspective because it cannot guarantee service subscribers (i.e., the system users) promised data rates, hence the absence of QoS. Alternatively, other utility functions are considered such as weighted sum rate, proportional fairness, minimum rate utility function or the harmonic mean utility function [10]. Among these objective functions, the harmonic mean utility function is more practical. The reasoning is two folds: no requirement for additional weight assignment and fairness to all users [11] [33]. Furthermore, when compared to sum rate objective, harmonic mean rate objective guarantees a larger minimum rate among the users. Figure 4.7 illustrates a 2 user case, where the red dot and blue dot indicate the sum rate and harmonic mean rate maximum, respectively. It is obvious that blue dot has a larger minimum rate between R_1 and R_2 than the red dot.

The harmonic mean rate can be written as $K\left(\sum_{k=1}^{K}\left(\sum_{n=1}^{N} R_k^n\right)^{-1}\right)^{-1}$. Considering each user's available power budget and spectrum mask, the harmonic mean

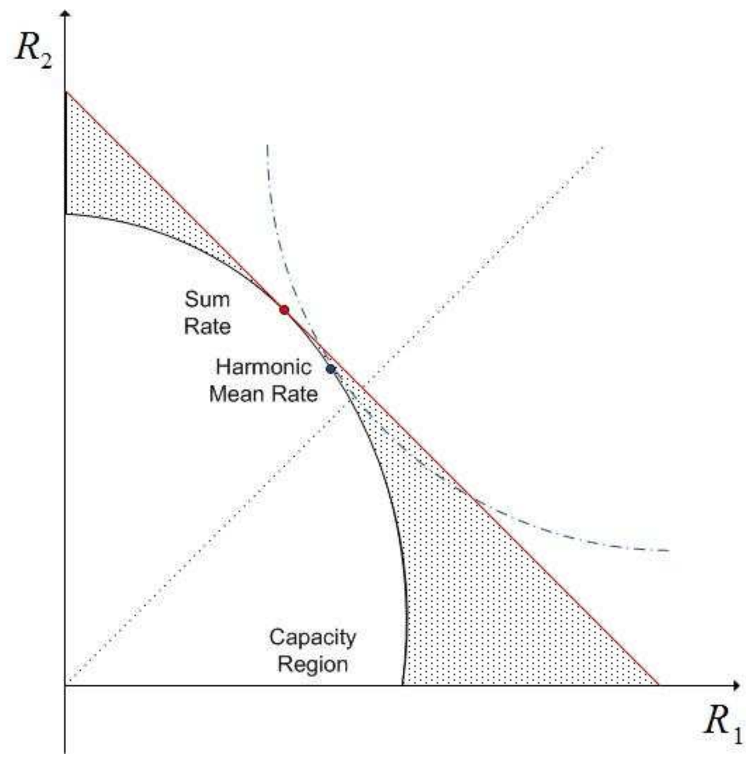


Figure 4.7: Minimum Rate Comparison between Harmonic Mean Rate Objective and Sum Rate Objective in a 2 User Case

maximization problem can be written as:

$$
\begin{aligned}
\min_{\{s_1^n,\ldots,s_K^n\}_{n=1}^N} \quad & \sum_{k=1}^{K}(\sum_{n=1}^{N} R_k^n)^{-1} \\
\text{s.t.} \quad & \sum_{n=1}^{N} s_k^n \leq P_k, \quad \forall k \\
& 0 \leq s_k^n \leq S_{k,\max}^n, \ \forall k \quad \forall n
\end{aligned}
\tag{4.2.19}
$$

where in the objective function we use the fact that maximizing harmonic mean rate is equivalent to minimizing its inverse. We also use P_k to denote the kth user's power constraint as a transmitter while $S_{k,\max}^n$ denotes the maximum signal power that user k can allocate to the n-th tone. Note that we assume that $\sum_{n=1}^{N} S_{k,\max}^n \geq P_k$ in order for the power constraint not to be redundant. When compared (4.1.3), (4.2.19) is fundamentally different in the sense that (4.2.19) is per user harmonic mean and (4.1.3) is per tone per user harmonic mean. Furthermore, (4.2.19) cannot be transformed into a convex problem.

Although solving (4.2.19) is preferable, it is easy to verify that the harmonic mean objective function is non-convex and it has been shown that this problem is NP-hard for the case $N > 1$ [10]. Hence there is not a computationally efficient algorithm to solve for global optimum of (4.2.19). The difficulties stem from two aspects: the objective function is not convex and it is not dual-decomposable over tones due to the inverse of R_k^n, which raises higher computational complexity compared to the sum rate mean objective function. Besides, in real system set-ups, N and K are usually so large that brute force algorithms are not feasible for practical use. In order to circumvent these difficulties, in the next section we utilize Taylor expansion of the objective function in (4.2.19) and derive its lower bound, based on which we develop a novel algorithm that solves the problem efficiently.

4.2.1 Taylor Expansion Based Iterative Algorithm:

Taylor Expansion of the Harmonic Mean Rate

The harmonic mean objective function is not decomposable over tones which leads to the problem of computational intractability. As an alternative, we resort to an iterative algorithm via first order Taylor expansion of R_k at the current power allocation at each iteration which yields the following expansion:

$$(\tilde{R}_k^{(i)})^{-1} = \frac{1}{R_k^{(i)}} - \frac{1}{(R_k^{(i)})^2}(R_k - R_k^{(i)}), \tag{4.2.20}$$

where $(\tilde{R}_k^{(i)})^{-1}$ is the first order Taylor expansion of $(R_k)^{-1}$ at $R_k^{(i)}$ and $R_k^{(i)}$ is the value of R_k at the ith iteration. Denote $\left\{s_k^{n(i)}\right\}$ are the set of resulting power allocation of user k at the i-th iteration. Then $R_k^{(i)}$ is equal to

$$\begin{aligned} R_k^{(i)} &= \sum_{n=1}^{N} R_k^{n(i)} \\ &= \sum_{n=1}^{N} \log(1 + \frac{s_k^{n(i)}}{\sigma_k^n + \sum\limits_{\ell \neq k} \alpha_{\ell k}^n s_\ell^{n(i)}}). \end{aligned} \tag{4.2.21}$$

From (4.2.20), it is easy to notice that minimizing $\sum_{k=1}^{K}(\tilde{R}_k^{(i)})^{-1}$ is equivalent to maximizing $\sum_{k=1}^{K} \frac{R_k}{(R_k^{(i)})^2}$ at the i-th iteration. In other words, at iteration i we solve the following optimization problem based on the above analysis:

$$\begin{aligned} \max_{\{s_k^{n(i+1)}\}} \quad & \sum_{k=1}^{K} (R_k^{(i)})^{-2} \sum_{n=1}^{N} R_k^{n,(i+1)} \\ \text{s.t.} \quad & \sum_{n=1}^{N} s_k^{n(i+1)} \leq P_k, \ \forall k \\ & 0 \leq s_k^{n(i+1)} \leq S_{k,\max}^n, \ \forall k \ \forall n, \end{aligned} \tag{4.2.22}$$

where $R_k^{n(i+1)}$ is in the form of (4.2.21). It is interesting to observe that the objective of (4.2.22) is a weighted sum rate function with weights that are inverse of squared user rates of the previous iteration. Therefore, for each user, the larger the data rate in the previous iteration, the less weight it receives in the current iteration, hence the algorithm yields a lower data rate for this user. As a result, the algorithm favors the users with lower data rates (i.e., higher weights) at each iteration. This mechanism suggests an iterative weight adjustment scheme for the algorithm which implicitly achieves the goal of relatively fair rates among the users. The objective function in (4.2.22) combines weighted sum rate with special data rate weights and has high complexity. In the next part, we introduce a lower bound for the weighted sum rate that allows us to solve the problem in a computationally effective manner and we will explain the overall scheme.

The Iterative Algorithm

The objective in (4.2.22) is non-concave and NP-hard. In attempting to efficient algorithms to solve (4.2.22), we derive a lower bound for this objective function. This lower bound is obtained via concavity of the logarithm function and Jensen's Inequality as follows:

$$\begin{aligned} R_k^n \geq \log(1 + \sum_{j=1}^{K} \alpha_{jk}^n) + \frac{\sigma_k^n}{1 + \sum_{i=1}^{K} \alpha_{ik}^n} \\ + \sum_{j=1}^{K} \frac{\alpha_{jk}^n \log s_j^n}{1 + \sum_{i=1}^{K} \alpha_{ik}^n} - \log(\sigma_k^n + \sum_{j \neq k} \alpha_{jk}^n s_j^n) \triangleq R_{\mathrm{LB},k}^n \end{aligned} \quad (4.2.23)$$

Therefore, at each iteration, instead of solving (4.2.22) we can simply solve the following optimization problem:

$$\begin{aligned}
\min_{\{s_k^{n(i+1)}\}} \; & \sum_{k=1}^{K} (R_k^{(i)})^{-2} \sum_{n=1}^{N} \Bigg[\log(\sigma_k^n + \sum_{j \neq k} \alpha_{jk}^n s_j^{n(i+1)}) \\
& - \sum_{j=1}^{K} \frac{\alpha_{jk}^n \log s_j^{n(i+1)}}{1 + \sum_{i=1}^{K} \alpha_{ik}^n} \Bigg] \\
\text{s.t.} \; & \sum_{n=1}^{N} s_k^{n(i+1)} \leq P_k, \; \forall k; \\
& 0 \leq s_k^{n(i+1)} \leq S_{k,\max}^n, \; \forall k \;\; \forall n.
\end{aligned} \tag{4.2.24}$$

At first glance, it may seem difficult to solve (4.2.24). As a matter of fact, (4.2.24) is convex under the transformation $y_k^n = \log s_k^n$ ([11]) and (4.2.24) can be reformulated as the following problem:

$$\begin{aligned}
\min_{\{y_k^{n(i+1)}\}} \; & \sum_{k=1}^{K} (R_k^{(i)})^{-2} \sum_{n=1}^{N} \Bigg[\log(\sigma_k^n + \sum_{j \neq k} \alpha_{jk}^n \exp(y_j^{n(i+1)})) \\
& - \sum_{j=1}^{K} \frac{\alpha_{jk}^n y_j^{n(i+1)}}{1 + \sum_{i=1}^{K} \alpha_{ik}^n} \Bigg] \\
\text{s.t.} \; & \sum_{n=1}^{N} \exp(y_k^{n(i+1)}) \leq P_k, \; \forall k \\
& y_k^{n(i+1)} \leq \log S_{k,\max}^n, \forall k \; \forall n.
\end{aligned} \tag{4.2.25}$$

Note that the above transformation not only convexifies the objective function, but also implicitly implies the constraints $(0 \leq s_k^n)$ which yield less computation on projection thus a faster algorithm. Furthermore, (4.2.25) is decomposable in the sense that in order to calculate the gradient of the objective function, we need only the partial derivatives $\frac{\partial R_{LB,k}^n}{\partial s_j^m}$. Obviously, this needs less computational time compared to calculating the partial derivatives $\frac{\partial (R_k^{-1})}{\partial s_j^m}$ in (4.2.19). Moreover, (4.2.25) can be decomposed into N independent subproblems in its dual form. To further illustrate this idea, we let λ_k be the Lagrange multiplier associated to the total power constraint and the Lagrange function is

$$\begin{aligned}
&L(s,\lambda)\\
=&\sum_{k=1}^{K}(R_k^{(i)})^{-2}\sum_{n=1}^{N}\Bigg[\log(\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n s_j^{n(i+1)})\\
&-\sum_{j=1}^{K}\frac{\alpha_{jk}^n\log s_j^{n(i+1)}}{1+\sum_{i=1}^{K}\alpha_{ik}^n}\Bigg]+\sum_{k=1}^{K}\lambda_k\left(\sum_{n=1}^{N}\exp(y_k^{n(i+1)})-P_k\right)\\
=&\sum_{k=1}^{K}\left\{\sum_{n=1}^{N}\Bigg[(R_k^{(i)})^{-2}\Bigg(\log(\sigma_k^n+\sum_{j\neq k}\alpha_{jk}^n s_j^{n(i+1)})\right.\\
&\left.-\sum_{j=1}^{K}\frac{\alpha_{jk}^n\log s_j^{n(i+1)}}{1+\sum_{i=1}^{K}\alpha_{ik}^n}+\lambda_k\exp(y_k^{n(i+1)})\Bigg)\Bigg]-\lambda_k P_k\right\}. \qquad (4.2.26)
\end{aligned}$$

From (4.2.26), one can see for fixed λ_k's that $L(s,\lambda)$ is not coupled over tones, and hence each subproblem in the sum term is independent. Furthermore, each subproblem is convex and can be evaluated with efficient optimization techniques. This is the major advantage of our algorithm. Now, the overall proposed algorithm can be stated as the following:

Step 1 Initialize $s_k^{n,(0)}$ by water-filling for each user;

Step 2 At iteration i, solve (4.2.25) via dual optimization algorithms to obtain the $\{y_k^{n,(i)}\}$'s which gives $\{s_k^{n,(i)}\}$'s, with which we calculate the $R_k^{(i)}$'s and update (4.2.25);

Step 3 If the increase in the original harmonic mean objective function is more than the tolerance level ε, go back to Step 2, otherwise stop.

4.2.2 Discussions

In this section, we address two important issues regarding the accuracy and practicality of the proposed algorithm. In the development of the algorithm, (4.2.22) and (4.2.25) are obtained through two approximations. However these two approximations are only tight under certain conditions – roughly speaking, that the achievable rates are large. We further discuss the usefulness of both approximations by explaining cases when system parameters are relatively small. Moreover, we offer insights on solving the subproblems arising from solving (4.2.25) and provide techniques to overcome them in a computationally efficient manner.

Accuracy of the approximations:

Using the Lagrange form for the remainder term in Taylor expansion, R_k^{-1} can be written as

$$R_k^{-1} = (R_k^{(i)})^{-1} - (R_k^{(i)})^{-2}(R_k - R_k^{(i)}) + \xi^{-3}(R_k - R_k^{(i)})^2$$

where ξ is a real number between R_k and $R_k^{(i)}$. It is easy to see that when ξ is large the third term in the above equation is small. Therefore, the approximation of R_k^{-1} using the first two terms of the Taylor expansion in the proposed algorithm can be tight when R_k and $R_k^{(i)}$ are large, hence ξ^{-3} is small. On the other hand, the bound in (4.2.23) is obtained by using Jensen's inequality and it is tight when one of the weights in Jensen's inequality dominates all the other weights. In translation to our case, the inequality is tight when α_{jk}^n's ($j \neq k$) are small compared to α_{kk}^n, which has the value 1 as it is normalized. Therefore, the second bound is tight when the cross-talk coefficients are small. This condition coincides the one for the first approximation to be tight that R_k's are large. Here we use the fact that R_k increases as the cross-talk coefficients decrease. In summary, two inequalities are generally tight under the conditions that cross-talk coefficients are small and R_k's are large. Under these conditions, the resulting solutions can be close to optimum.

Solving the subproblem:

In the proposed algorithm, (4.2.25) is evaluated iteratively as a subproblem at each iteration. To solve this problem, we utilize its dual problem (4.2.26) and optimize the Lagrangian function to update primal/dual variables simultaneously for certain number of steps as it is in Lagrangian method [15]. However, convergence is not guaranteed for this method. Fortunately, for our problem, this method converges in few steps (less than ten) in most cases. Therefore, we do seven update steps to optimize the Lagrangian at each iteration and then the objective function is updated by re-evaluating the $R_k^{(i)}$ according to the first order Taylor expansion of the objective function. It is worth mentioning that the dual problem also decouples the problem across tones. Hence this method allows us to use cheap computational methods to solve this difficult problem.

4.2.3 Numerical Results

In this section, the performance of the proposed algorithm is compared with the performance of the traditional IWFA algorithm. In the simulations, we consider the case where there are three users sharing 64 tones to communicate in a system with interference channels. The crosstalk coefficients are generated uniformly in the interval of [0 , 0.3] and the noise is randomly generated according to the Gaussian distribution with a zero mean and unit variance. In order to make the users dissimilar (to check the user fairness issue), the cross-talk coefficients are generated in the interval [0 , 0.9] for User 2. The input signal power constraint is set to 15 dB. All users are assumed to have the same power budget constraint P_k and spectrum mask $S_{\max}$.

Using the above parameters, Figure 4.8 shows the convergence of the proposed algorithm and makes a comparison in term of harmonic mean achieved by the proposed algorithm and the levels achieved by traditional IWFA and SCALE over iteration. Note that our algorithm and IWFA have the same initialization. However, since there are a lot of oscillations at the first steps of IWFA and SCALE, only the resulting level of harmonic mean for IWFA and SCALE are plotted.

Figure 4.9 presents a comparison between the achievable data rate for the minimum rate user (minimum rate utility function [10]) of the proposed algorithm and that achieved by using IWFA as a function of different power constraints (P_k=14, 17, and 20 dB) for the input signal. As one can see, the performance of the worst user is much better in the proposed algorithm than under the other methods.

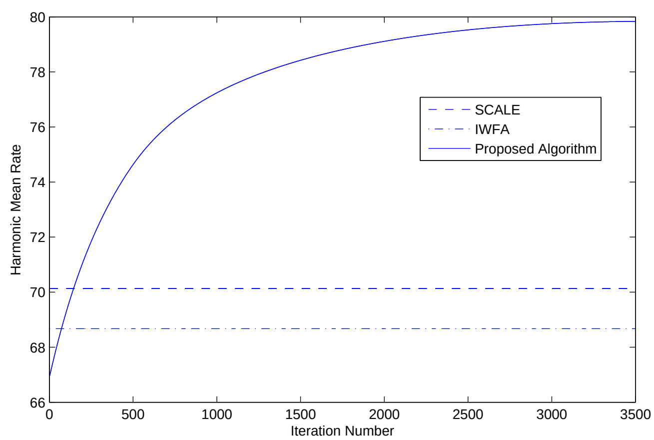


Figure 4.8: Harmonic Mean Rate versus Iteration Number

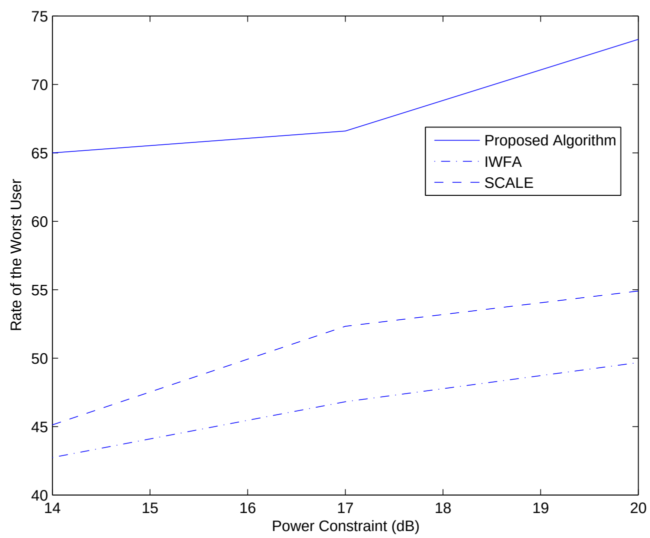


Figure 4.9: Harmonic Mean Rate versus Iteration Number

Chapter 5

Multi-user Communication Systems in the Presence of a Jammer and the Generalized Iterative Water-Filling Algorithm

In this section, we consider a similar communication with the presence of a jammer, denoted as User 0. Suppose that User $k \in \mathcal{K},\ (k \neq 0)$ is interested in maximizing its own sum-rate, so its utility is given by

$$U_k(\mathbf{s}_0, \mathbf{s}_1, \cdots, \mathbf{s}_K) = \sum_{n=1}^{N} R_k^n(s_1^n, \ldots, s_K^n) = \sum_{n=1}^{N} \log\Big(1 + \frac{s_k^n}{\sigma_k^n + \sum_{j \neq k} \alpha_{jk}^n s_j^n + \alpha_{0k}^n s_0^n}\Big), \tag{5.0.1}$$

while the utility of the jammer is

$$U_0(\mathbf{s}_0, \mathbf{s}_1, \cdots, \mathbf{s}_K) = -\sum_{k=1}^{K} U_k = -\sum_{k=1}^{K} \sum_{n=1}^{N} \log\Big(1 + \frac{s_k^n}{\sigma_k^n + \sum_{j \neq k} \alpha_{jk}^n s_j^n + \alpha_{0k}^n s_0^n}\Big), \tag{5.0.2}$$

where we use $\mathbf{s}_k$ to denote the vector $[s_k^1, \cdots, s_k^N]^T$.

Given a limited power budget, and a maximum power constraint on each tone, the goal of User k, is to maximize U_k; that is, User k wishes to choose $s_k^n, n = 1, \ldots, N$ to solve the following optimization problem,

$$\begin{aligned} \max \quad & U_k(\mathbf{s}_0, \mathbf{s}_1, \cdots, \mathbf{s}_K), \\ \text{subject to} \quad & \sum_{n=1}^{N} s_k^n \leq P_k, && (5.0.3a) \\ & 0 \leq s_k^n \leq S_{\max,k}^n, && (5.0.3b) \end{aligned}$$

where, P_k denotes the total power budget of User k, $S_{\max,k}^n$ denotes the maximum signal power that User k can use on the n-th tone, and in order for (5.0.3a) not to be redundant, we assume that $P_k \leq \sum_{n=1}^{N} S_{\max,k}^n$. We will denote the feasible set of User k as $\mathcal{P}_k$; that is,

$$\mathcal{P}_k \triangleq \left\{ \mathbf{s}_k = [s_k^1, \cdots, s_k^N]^T \middle| \sum_{n=1}^{N} s_k^n \leq P_k, 0 \leq s_k^n \leq S_{\max,k}^n \right\}. \tag{5.0.4}$$

Since individual users do not collaborate among themselves nor do they collaborate with the jammer, and since all users and the jammer selfishly maximize their own utilities, this communication scenario can be modelled as a non-cooperative game [16]. In this game individual users and the jammer are non-cooperative players, and the power allocations of any User k, including the jammer, that lie in $\mathcal{P}_k$ (cf., (5.0.4)) represent the set of admissible strategies of this user. A Nash equilibrium of this game [16] is a tuple of power strategies $\{\mathbf{s}_k^*\}_{k=0}^K$, such that for any $k \in \{0\} \cup \mathcal{K}$

$$U_k(\mathbf{s}_0^*, \mathbf{s}_1^*, \cdots, \mathbf{s}_{k-1}^*, \mathbf{s}_k^*, \mathbf{s}_{k+1}^*, \cdots, \mathbf{s}_K^*) \geq U_k(\mathbf{s}_0^*, \mathbf{s}_1^*, \cdots, \mathbf{s}_{k-1}^*, \mathbf{s}_k, \mathbf{s}_{k+1}^*, \cdots, \mathbf{s}_K^*), \quad \forall \mathbf{s}_k \in \mathcal{P}_k. \tag{5.0.5}$$

In other words, a Nash equilibrium of the game is a locally optimal strategy for each player such that no player has an incentive to unilaterally change [16]. In the next section, we will show that, for this game, a Nash equilibrium always exists. Moreover, we will propose a decentralized algorithm for updating the jammer's and the users' power allocations. By analyzing the convergence of this algorithm, we also derive sufficient conditions under which the Nash equilibrium is unique.

5.1 Existence and Uniqueness of a Nash Equilibrium

Since, for every $k = 1, \cdots, K$, $U_k(s_0, s_1, \cdots, s_{k-1}, \bullet, s_{k+1}, s_K)$ is a continuously concave function, and so is $U_0(\bullet, s_1, \cdots, s_K)$, and since each $\mathcal{P}_k$ is a compact convex set, it follows readily from [34, Proposition 2.2.9] that a Nash equilibrium exists. Such an equilibrium can be found using a standard fixed-point algorithm, an instance of which is given in the next section.

5.1.1 A Generalized Iterative Water-Filling Algorithm (GIWFA)—Synchronous Version

A simple distributed algorithm for the users and the jammer to update their power allocation is the following generalized iterative water-filling algorithm (GIWFA). Let $s_k^{n,\nu}$ be the power allocation of User k on the n-th tone at iteration ν, and $\mathbf{s}_k^{\nu}$ be the vector $[s_k^{1,\nu}, \cdots, s_k^{N,\nu}]^T$. For the time being consider synchronous operation, whereby the users update their power allocations sequentially. Assume, without loss of generality, that the users are ordered so that User 1 updates its power allocation first then User 2 and so on, and that the jammer (User 0) updates its power allocation last.

Hence, in each iteration User $k \in \mathcal{K}$ updates its power allocations in order to solve

$$\mathbf{s}_k^{\nu+1} = \Big[\mathbf{s}_k^{\nu+1} + \nabla_{\mathbf{s}_k} U_k(\mathbf{s}_0^{\nu}, \mathbf{s}_1^{\nu+1}, \cdots, \mathbf{s}_{k-1}^{\nu+1}, \mathbf{s}_k, \mathbf{s}_{k+1}^{\nu}, \cdots, \mathbf{s}_K^{\nu})\Big|_{\mathbf{s}_k = \mathbf{s}_k^{\nu+1}}\Big]_{\mathcal{P}_k}, \tag{5.1.6}$$

whereas the jammer solves

$$\mathbf{s}_0^{\nu+1} = \Big[\mathbf{s}_0^{\nu+1} + \nabla_{\mathbf{s}_0} U_0(\mathbf{s}_0, \mathbf{s}_1^{\nu+1}, \cdots, \mathbf{s}_K^{\nu+1})\Big|_{\mathbf{s}_0 = \mathbf{s}_0^{\nu+1}}\Big]_{\mathcal{P}_0}, \tag{5.1.7}$$

where we use $[\cdot]_{\mathcal{P}_k}$ to denote the projection operator onto the polyhedron defined in (5.0.4). That is, for any vector $x \in \mathbb{R}^N$

$$[x]_{\mathcal{P}_k} = \arg\min_{y \in \mathcal{P}_k} \|y - x\|. \tag{5.1.8}$$

Using (5.0.1) and (5.0.2), we can compute the gradients $\nabla_{\mathbf{s}_k} U_k$ explicitly. In particular, the n-th entry of $\nabla_{\mathbf{s}_k} U_k$ for $k \in \{0\} \cup \mathcal{K}$, $[\nabla_{\mathbf{s}_k} U_k]_n$, can be expressed as

$$\begin{aligned}\Big[\nabla_{\mathbf{s}_k} U_k(\mathbf{s}_0^{\nu}, \mathbf{s}_1^{\nu+1}, \cdots, \mathbf{s}_{k-1}^{\nu+1}, \mathbf{s}_k, \mathbf{s}_{k+1}^{\nu}, \cdots, \mathbf{s}_K^{\nu})\Big|_{\mathbf{s}_k = \mathbf{s}_k^{\nu+1}}\Big]_n \\ = \frac{1}{\sigma_k^n + \sum_{j=1}^{k} \alpha_{jk}^n s_j^{n,\nu+1} + \sum_{j=k+1}^{K} \alpha_{jk}^n s_j^{n,\nu} + \alpha_{0k}^n s_0^{n,\nu}}, \qquad \forall k \in \mathcal{K},\end{aligned} \tag{5.1.9}$$

$$\begin{aligned}\Big[\nabla_{\mathbf{s}_0} U_0(\mathbf{s}_0, \mathbf{s}_1^{\nu+1}, \cdots, \mathbf{s}_K^{\nu+1})\Big|_{\mathbf{s}_0 = \mathbf{s}_0^{\nu+1}}\Big]_n \\ = \sum_{k=1}^{K} \frac{\alpha_{0k}^n s_k^{n,\nu+1}}{(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n s_k^{n,\nu+1} + \sigma_k^n + \alpha_{0k}^n s_0^{n,\nu})(\sum_{j=1}^{K} \alpha_{jk}^n s_k^{n,\nu+1} + \sigma_k^n + \alpha_{0k}^n s_0^{n,\nu})},\end{aligned} \tag{5.1.10}$$

where, in (5.1.9) and (5.1.10), we have used that $\alpha_{kk}^n = 1$ for all $k \in \mathcal{K}$.

From (5.1.9) and (5.1.10) we observe that for User $k \in \mathcal{K}$ to update its power allocation, it is sufficient to measure its own noise-plus-interference level on each tone, whereas for the jammer to update its power allocation, it needs, not only to know the power transmitted by each user, but also to know the noise-plus-interference level experienced by each user on every tone.

5.1.2 Convergence Analysis—Synchronous Version

We now present sufficient conditions under which this algorithm converges to the unique Nash equilibrium of the game. Applying [35, Proposition 11.13] it can be seen that a tuple of power strategies $\{\mathbf{s}_k^*\}_{k=0}^K$ achieves equilibrium if and only if

$$\mathbf{s}_k^* = \left[\mathbf{s}_k^* + \theta \nabla_{\mathbf{s}_k} U_k(\mathbf{s}_0^*, \mathbf{s}_1^*, \cdots, \mathbf{s}_{k-1}^*, \mathbf{s}_k, \mathbf{s}_{k+1}^*, \cdots, \mathbf{s}_K^*)\Big|_{\mathbf{s}_k=\mathbf{s}_k^*}\right]_{\mathcal{P}_k}, \qquad k \in \mathcal{K} \quad (5.1.11a)$$

$$\mathbf{s}_0^* = \left[\mathbf{s}_0^* + \theta \nabla_{\mathbf{s}_0} U_0(\mathbf{s}_0, \mathbf{s}_1^*, \cdots, \mathbf{s}_K^*)\Big|_{\mathbf{s}_0=\mathbf{s}_0^*}\right]_{\mathcal{P}_0}, \qquad (5.1.11b)$$

for some $\theta > 0$. Since our generalized iterative water-filling algorithm (5.1.6)–(5.1.7) corresponds to setting $\theta = 1$ in (5.1.11), then if this algorithm converges to a power strategy $\{\mathbf{s}_k^*\}_{k=0}^K$, then it must be a Nash equilibrium of the game (5.0.5). We now present sufficient conditions under which the generalized IWFA converges to a unique Nash equilibrium. In particular, let

$$A = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -\alpha_{12} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\alpha_{1K} & -\alpha_{2K} & \cdots & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & \alpha_{21} & \alpha_{31} & \cdots & \alpha_{K1} \\ 0 & 0 & \alpha_{32} & \cdots & \alpha_{K2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha_{K,K-1} \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \text{and} \quad \beta = \begin{bmatrix} \alpha_{01} \\ \vdots \\ \alpha_{0K} \end{bmatrix}, \qquad (5.1.12)$$

where we define $\alpha_{jk} \triangleq \left\|[\alpha_{jk}^1, \cdots, \alpha_{jk}^N]\right\|_2$ for all $j \in \{0\} \cup \mathcal{K}$, $k \in \mathcal{K}$, $j \neq k$. Furthermore, for every $k \in \mathcal{K}$, let F_k be a $N \times NK$ block-diagonal matrix whose n-th $1 \times K$

diagonal block is f_k^n. That is,

$$F_k \triangleq \begin{bmatrix} f_k^1 & 0 & \cdots & 0 \\ 0 & f_k^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f_k^N \end{bmatrix}, \tag{5.1.13}$$

where the i-th entry of f_k^n, $[f_k^n]_i$, $i = 1, \ldots, K$, be defined as follows.

$$\begin{aligned}[f_k^n]_k = &\frac{(S_{\max,0}^n)^2}{(d_{\min,k}^n)^2(c_{\min,k}^n + S_{\max,0}^n)^2} + \frac{\sum_{j=1,\, j\neq k}^K \alpha_{jk}^n S_{\max,j}^n}{(\sum_{j=1,\, j\neq k}^K \alpha_{jk}^n S_{\max,j}^n + \eta_k^n) c_{\min,k}^n d_{\min,k}^n} \\ &+ \frac{S_{\max,0}^n}{d_{\min,k}^n(c_{\min,k}^n + S_{\max,0}^n)}\left(\frac{1}{d_{\min,k}^n} + \frac{1}{c_{\min,k}^n} + \frac{S_{\max,k}^n}{c_{\min,k}^n(\alpha_{0k}^n c_{\min,k}^n + S_{\max,k}^n)}\right),\end{aligned} \tag{5.1.14}$$

$$\begin{aligned}[f_k^n]_i = &\frac{(S_{\max,k}^n)^2 d_{\min,k}^n + 2S_{\max,k}^n c_{\min,k}^n(\alpha_{0k}^n c_{\min,k}^n + S_{\max,k}^n)}{d_{\min,k}^n(c_{\min,k}^n)^2(\alpha_{0k}^n c_{\min,k}^n + S_{\max,k}^n)^2}\alpha_{ik}^n \\ &+ \frac{2S_{\max,0}^n S_{\max,k}^n}{c_{\min,k}^n(\alpha_{0k}^n c_{\min,k}^n + S_{\max,k}^n) d_{\min,k}^n(c_{\min,k}^n + S_{\max,0}^n)}\alpha_{ik}^n, \qquad i \neq k,\ i \in \mathcal{K},\end{aligned} \tag{5.1.15}$$

where

$$c_{\min,k}^n = \frac{1}{\alpha_{0k}^n}\left(\sum_{j=1,\, j\neq k}^K \alpha_{jk}^n \eta_j^n + \sigma_k^n\right), \tag{5.1.16}$$

$$d_{\min,k}^n = c_{\min,k}^n + \frac{1}{\alpha_{0k}^n}\eta_k^n, \tag{5.1.17}$$

with η_k^n being a lower bound on $s_k^{n,\nu}$. That is, for every iteration ν, $\eta_k^n \leq s_k^{n,\nu}$, $\forall k \in \mathcal{K}, n \in \mathcal{N}$. In Appendix A.2 we show that η_k^n is given by

$$\eta_k^n = \Big[\frac{1}{N}\big(P_k + \sum_{i=1}^{m_k} \sigma_k^{\pi_k(i)}\big) + \big(\frac{1}{N} - 1\big)\sum_{j=0,\, j\neq k}^K \alpha_{jk}^n S_{\max,j}^n - \sigma_k^{\pi_k(n)}\Big]^+, \tag{5.1.18}$$

where m_k is the largest integer for which

$$(m_k - 1)(\sigma_k^{\pi_k(j)} + \sum_{i=0,\, i\neq k}^{K} \alpha_{ik}^{\pi_k(j)} S_{\max,i}^{\pi_k(j)}) \leq P_k + \sum_{i=1}^{m_k - 1} \sigma_k^{(i)},$$

is satisfied for all $j \leq m_k$. For each User $k \in \mathcal{K}$ we use $\sigma_k^{(i)}$ to denote the noise variance that satisfies $\sigma_k^{(i)} \leq \sigma_k^{(i+1)}$, for all $i = 1, \ldots, N-1$. We also use $\pi_k(\cdot)$ to denote the tone permutation that satisfy

$$\sigma_k^{\pi_k(1)} + \sum_{\substack{j=0\\ j\neq k}}^{K} \alpha_{jk}^{\pi_k(1)} S_{\max,j}^{\pi_k(1)} \leq \cdots \leq \sigma_k^{\pi_k(N)} + \sum_{j=0,\, j\neq k}^{K} \alpha_{jk}^{\pi_k(N)} S_{\max,j}^{\pi_k(N)}.$$

Theorem 1 (Convergence of GIWFA) *Suppose there exists a scalar $\tau \in (0,1)$ such that the following conditions are satisfied*

$$\Big(1 + \frac{\left\|\sum_{k=1}^{K} F_k\right\|_2^2}{(1-\tau)^2}\Big)(\|A^{-1}B\|_2^2 + \|A^{-1}\beta\|_2^2) < 1, \tag{5.1.19}$$

$$\max_n \sum_{k=1}^{K} \frac{S_{\max,k}^n (2c_{\min,k}^n + \frac{S_{\max,k}^n}{\alpha_{0k}^n})}{(c_{\min,k}^n)^2 (c_{\min,k}^n + \frac{S_{\max,k}^n}{\alpha_{0k}^n})^2} \leq \tau + 1, \tag{5.1.20}$$

$$\min_n \sum_{k=1}^{K} \Bigg(\frac{(\alpha_{0k}^n)^3 \eta_k^n}{\big(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\big)^2 \big(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + \eta_k^n + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\big)}$$
$$+ \frac{(\alpha_{0k}^n)^3 \eta_k^n}{\big(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\big)\big(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + \eta_k^n + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\big)^2} \Bigg)$$
$$\geq 1 - \tau. \tag{5.1.21}$$

Then the noncooperative game (5.0.5) *has a unique Nash equilibrium, and the iterates generated by the GIWFA algorithm converges to this unique equilibrium linearly.*

Proof 1 *Fix any equilibrium solution and any starting power allocation. We define the error vector at each iteration to be the difference between the current power al-*

location and the power allocation at equilibrium. In Appendix A.1, we show that the conditions (5.1.19)–(5.1.21) *imply the error vectors converge to zero at a geometric rate. Since the choice of equilibrium solution is arbitrary, it follows that the noncooperative game* (5.0.5) *has a unique Nash equilibrium.*

Notice that the conditions (5.1.19)–(5.1.21) only depend on the power budget of each user, its maximum allowable power on each tone and the cross-talk coefficients. In many practical scenarios, these parameters, or a reasonably accurate estimate thereof, may be known *a priori* to the system designer. Hence, these conditions allow the system designer to study the impact of a potential jammer on the users' utilities as well as the sum rate of the whole system. In Section 5.2 we will present numerical results that show that for scenarios in which the conditions of Theorem 1 are met, the choice of both the users and the jammer converge. We also provide instances showing that the violation of these conditions may cause the algorithm to oscillate.

Observe that for any τ, the condition in (5.1.19) implies the standard IWFA convergence conditions. In particular, for any such τ for which (5.1.19) holds, we have

$$\|A^{-1}B\|_2 < 1. \tag{5.1.22}$$

Condition (5.1.21) implies that

$$\min_n \sum_{k=1}^{N} s_k^{n,\nu} \geq \min_n \sum_{k=1}^{N} \eta_k^{n,\nu} > 0.$$

Thus if $s_k^{n,*} \equiv \lim_{\nu\to\infty} s_k^{n,\nu}$, then

$$\min_n \sum_{k=1}^{N} s_k^{n,*} > 0.$$

In words, this says that the Nash equilibrium computed by the GIWFA has the property that every tone n is used by at least one user k. Another insight offered by Theorem 1 is that if the jammer's maximum signal power $S^n_{\max,0}$ on tone n is sufficiently large so that $\eta^n_k = 0$ for all k, then (5.1.21) cannot be satisfied and the convergence of the GIWFA is in jeopardy.

5.1.3 Extension to Asynchronous GIWFA

In Sections 5.1.1 and 5.1.2 we considered the case in which the users and the jammer update their power allocations sequentially in a predetermined order according to a common clock. However, in many practical scenarios a common clock may not be available for the users and the jammer to operate in such a synchronous fashion. Moreover, even if such a clock is available, due to practical implementation issues, either the users or the jammer may not have access to the most recent multi-user interference. In this case an asynchronous version of the GIWFA algorithm may be more desirable and more robust to implement than a synchronous one.

In a totally asynchronous scheme, the users and the jammer update their power allocations at arbitrary time instants using possibly outdated multi-user interference [18]. Under certain mild conditions a fundamental result in [18, Proposition 2.1] ensures that the asynchronous scheme converges to a unique Nash equilibrium of the game (5.0.5) if: 1) each user and the jammer update their power allocations at least once within any sufficiently large, but finite, time interval, and 2) the iterates contract with respect to some norm. This contraction condition is precisely the same as the set of conditions given in Theorem 1; see also Appendix A.1.[1] In other words, the conditions given in Theorem 1 ensure convergence of both the synchronous and the

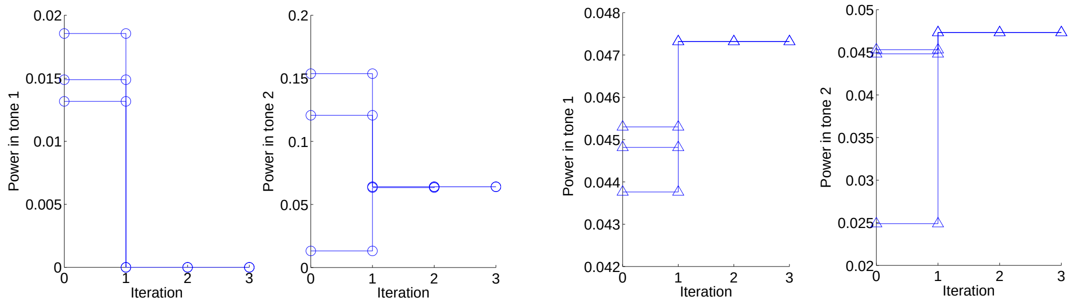
[1] For the asynchronous scheme the iteration indices ν and $\nu + 1$ in Appendix A.1 ought to be interpreted as the time instants within which each user and the jammer will have updated their power allocations at least once.

asynchronous versions of the GIWFA algorithm.

5.2 Numerical Results

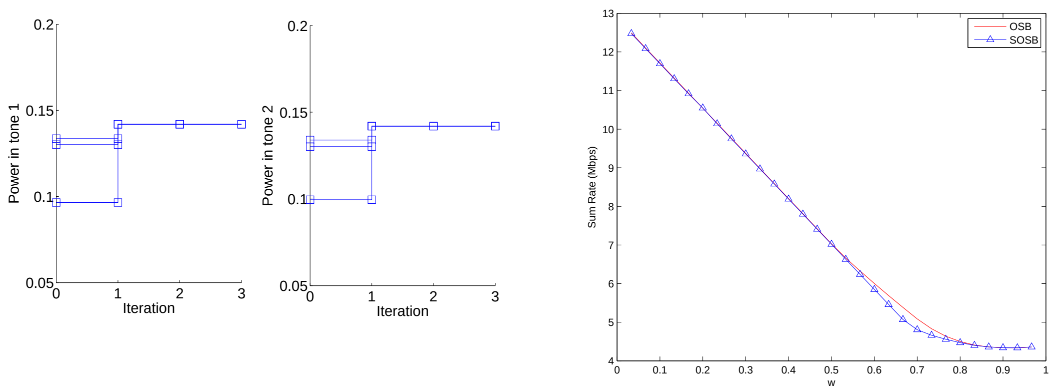
In this section we provide a numerical example that illustrates the sufficiency of the conditions given in Theorem 1 for the convergence of the decentralized GIWFA algorithm. We also provide an example that shows that when the conditions in Theorem 1 are violated the users and the jammer may fail to converge and the behavior of the GIWFA becomes dependent on the initial point. For the numerical examples in this section, the number of users, $K = 4$, and the number of tones $N = 10$, and the maximum allowable power per tone is set to be constant across tones for each user as well as for the jammer; i.e., we set $S^n_{\max,k} = S_{\max,k}$, $n = 1, \ldots, 10$ for $k = 0, \ldots, 4$.

Example 1 *In this example, the system parameters (i.e,* $\alpha^n_{jk}, \sigma^n_k, P_k, S_{\max}, k, \forall\ j \neq k,\ k = 0, \ldots, 4$*) are selected at random so as to satisfy the conditions in Theorem 1. The users and the jammer update their power allocations using the GIWFA algorithm described in Section 5.1.1. For this scenario, in Figures 5.1(a) and 5.1(b) we plot the power allocations of Users 1 and 2 versus the iteration number for all the tones. For the same scenario, in Figure 5.1(c) we plot the power allocations of the jammer versus the iteration number. In each of the plots, three randomly chosen allocations were used to initialize the fixed-point algorithm. Since the system parameters were chosen to meet the conditions of Theorem 1, the algorithm converges to a unique Nash equilibrium, irrespective of the initial power allocations. In order to quantify the jammer's impact on the overall system performance, the sum rate of all the users over the ten tones is plotted versus the iteration number in Figure 5.1(d), where Users 1 and 2 are Marked by 'o' and '△', respectively, whereas the power allocations*



(a) Power allocations of User 1

(b) Power allocations of User 2



(c) Power allocations of the jammer

(d) Sum rate with and without a jammer

Figure 5.1: The Power Allocations of Users and Jammer

of the jammer is Marked by '□'. It is shown that the GIWFA iterates converge to a unique Nash equilibrium irrespective of the initial power allocation.

Example 2 *In this example, we retain the channel gains of the users as per Example 1. (Since, in Example 1 the gains were selected to meet the conditions in Theorem 1, these gains also meet the IWFA convergence condition* (5.1.22)*. However, the channel gains of the jammer are chosen such that the conditions in Theorem 1 are violated. In this example, we consider two random instances of this scenario. For*

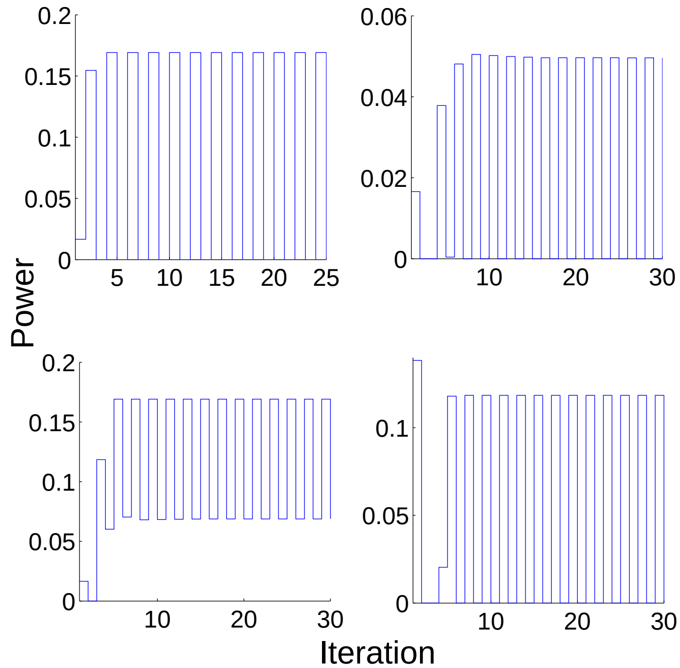


Figure 5.2: The Power Allocations of User 1 DO NOT Converge on Four Different Tones.

the first instance, we show the power allocations of one of the users on some of the tones. As can be seen from Figure 5.2, on these tones the user's allocations do not converge, and, in fact, they keep fluctuating. In the second instance of this example, we initialize the GIWFA algorithm using three different randomly chosen power allocations. In Figure 5.3 we plot the sum rate versus the iteration number in this case. It can be seen that the sum rate fluctuates and no equilibrium is reached.

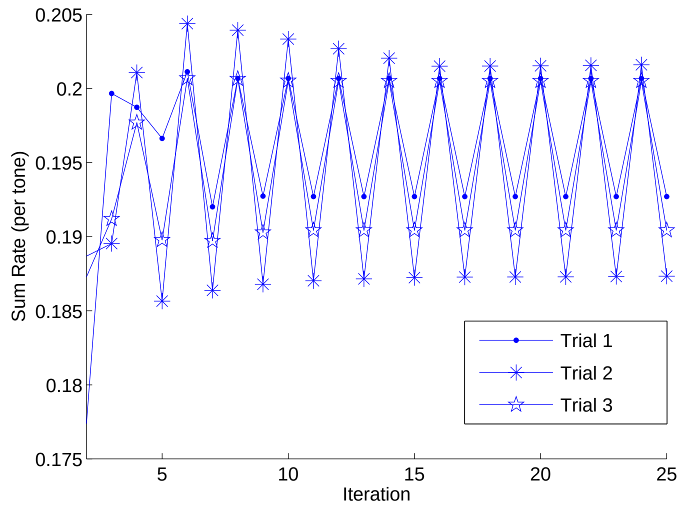


Figure 5.3: By Changing the Initial Power Allocation, the Iterates may Oscillate.

Chapter 6

Conclusions and Open Questions

6.1 Conclusions

In this work, we studied a multiuser communication scenario in which K user pairs share N orthogonal tones. We have proposed three dynamic power allocation algorithms targeting weighted sum rate, user capacity and harmonic mean rate system utilities, respectively.

SSB, which has the goal to maximize the weighted sum rate of the system, divides the spectrum into three sets of tones $\mathcal{F}$, $\mathcal{NF}_{\mathrm{IWFA}}$ and $\mathcal{NF}_{\mathrm{M}}$, within each of which it applies FDMA, IWFA and Lagrange dual algorithm, respectively. We showed by simulation that SSB achieves a performance close to optimal with relative low complexity.

In the second part of the dissertation, we defined user capacity which is the maximum number of users supported by the system and we designed an algorithm to approach this capacity. Inspired by the observation that a harmonic mean objective can provide fairness for users' data rates, we indirectly achieved the goal of approaching user capacity by maximizing the harmonic mean rate of the system. However, this optimization problem is non-convex when the number of tones is greater than one. To overcome this non-convexity, we took advantage of the convexity feature of per-tone harmonic mean and proposed a computationally-efficient algorithm. Furthermore, we incorporated maximum and minimum rate constraints into the proposed algorithm to achieve the rate ranges that users demand, thereby providing a convenient tool for a service provider to guarantee quality-of-service. Furthermore, we studied the harmonic mean rate objective function and proposed an iterative algorithm based on the lower bound for the local Taylor expansion of the harmonic mean function. A transformation of the variables is also provided for convexifing the subproblem. By considering the dual problem, the subproblem is decomposed across the tones. This decomposition makes the optimization algorithm much faster. The numerical results also show improvements in the sense of user fairness and achievable data rate compared to some well-known algorithms.

In the last part of the dissertation, we considered the system scenario where there is a malicious jammer wishing to disrupt the communication of the users and modeled the behaviors of the users and the jammer as a noncooperative game. We derived sufficient conditions under which the iterates of both synchronous and totally asynchronous decentralized GIWFA algorithms converge to a unique Nash equilibrium. Our theoretical analysis and numerical simulations show that the presence of a strong jammer can not only reduce the total network throughout, but also cause an otherwise convergent IWFA to oscillate.

6.2 Suggestions on Future Research

The study in this dissertation opens up the following research directions that merit further investigation:

- Performance analysis of the proposed algorithms

- Optimality conditions for FDMA signalling in harmonic mean objective

- Price setting for the resources and utilization of prices to facilitate the resource management

6.2.1 Performance Analysis of the Proposed Algorithms

Convergence and Complexity Analysis on the Proposed Algorithms:

Further studies on the convergence and complexity of the proposed SSB and harmonic mean optimization algorithm. The goal of SSB is to solve weighted sum rate optimization problems which is are non-convex. The algorithm converges to a fixed point of the optimization problem. However, due to the non-convex nature of the problem, this fixed point may be a local maximum or saddle point. It is of system designers interests to quantify what is the performance loss in terms of the weighted sum-rate of the system since the ultimate goal is to maximize the system utility.

Recall that in designing the harmonic mean optimization algorithm, we convexified the objective function and the minimum rate constraint so as to facilitate the computation of the original problem. We utilized the upper bound of the harmonic mean rate as the new objective function in the formulation and solved the reformulated optimization problem by the harmonic mean optimization algorithm. This raises the question of optimality of this reformulation –i.e., the harmonic mean achieved by our algorithm compared to the true optimum.

Optimality Conditions for Assigning Tones into Sets:

The proposed SSB algorithm categorizes available tones into three sets and partitions available power budgets to be used by corresponding algorithms for these sets. The criteria utilized to partition the tones are derived from the theorem proposed in [13]. However, there is no proof of the optimality of using these criteria to partition the tones into different sets. Furthermore, the power budget should be added to the criteria so as to improve the performance of the algorithm.

6.2.2 Sufficient Conditions for Harmonic Mean Function

[1] shows the optimality conditions for FDMA structured signalling for spectrum management problems with sum rate objective functions. In Chapter 4 showed that the harmonic mean objective function has the advantage of giving users fairer rates and ensuring users' quality of service. Therefore, it would be very intriguing to study similar conditions under the harmonic mean objective function. The conditions can facilitate solving of the problem using the algorithm developed here.

6.2.3 Price Setting for the Resources and Utilization of Prices to Facilitate the Resource Management

In a free market mechanism, prices usually serve as a leverage to balance demands and supplies as well as to indicate the relationship between them. In a multi-user communication system as considered here, heterogeneous system users compete for a common spectrum so as to satisfy their needs for communication, which creates the demands for system resources. A system administrator or service provider aims to devise a scheme to regulate users' power allocation by utilizing prices. The system administrator sets prices on the communication resources in such a way that system utility (e.g. revenue) is maximized. In the above mentioned scenario, the non-cooperative users' behaviors can be modeled as a non-cooperative game and a system administrator's optimization problem can be cast upon the Nash equilibrium of this game. The existence of Nash equilibrium and an algorithm that finds a Nash equilibrium are very interesting topics for future work.

6.3 Suggestions to Future Researchers

Throughout the course of my PhD program, I have acquired a significant amount of knowledge of optimization methods and other valuable skills such as research methodology and presentation skills. The most important thing I have learned, which may be useful to future researchers in this field, is that when dealing with optimization problems in engineering it is usually not easy to solve them directly from a mathematical point of view. However, with inspiration from the engineering interpretation on the problem, approximations and bounds may help to reduce the complexity of the problem while providing a good quality solution.

Bibliography

[1] Shunsuke Hayashi and Zhi-Quan Luo, "Spectrum management for interference-limited multiuser communication systems," *IEEE Trans. Inform. Theory*, June 2006, Accepted.

[2] Raphael Cendrillon, Wei Yu, Marc Moonen, Jan Verlinden, and Tom Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Trans. Signal Processing*, vol. 54, no. 5, pp. 922–933, May 2006.

[3] Rapheal Cendrillon, Jianwei Huang, Mung Chiang, and Marc Moonen, "Autonomous spectrum balancing for digital subscriber lines," *IEEE Trans. Signal Processing*, vol. 55, no. 8, pp. 4241–4257, Aug. 2007.

[4] Wei Yu, George Ginis, and John M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE J. Select. Areas Commun.*, vol. 20, no. 5, pp. 1105–1115, June 2002.

[5] Nobuo Yamashita and Zhi-Quan Luo, "A nonlinear complementarity approach to multiuser power control for digital subscriber lines," *Optimization Methods Software*, vol. 19, no. 5, pp. 633–652, Oct. 2004.

[6] Chaoxiong Wang, *Power control for multiuser communication systems and computation of generalized Nash equilibria*, Ph.D. thesis, Rensselaer Polytech. Inst., New York, Nov. 2007.

[7] Muriel Médard, "Capacity of correlated jamming channels," in *Proc. 35th Allerton Conf. Commun., Control Comput.*, Monticello, Il, Sept. 1997, pp. 1043–1052, Also available at:`http://www.mit.edu/~medard/pubs.html`.

[8] Ian F. Akyildiz, Won-Yeol Lee, Mehmet C. Vuran, and Shantidev Mohanty, "NeXt Generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Comp. Ntwk.*, vol. 50, no. 13, pp. 2127–2159, Sept. 2006.

[9] Simon Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Select. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[10] Zhi-Quan Luo and Shuzhong Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Select. Topics Signal Processing*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[11] Zhi-Quan Luo and Jong-Shi Pang, "Analysis of iterative waterfilling algorithm for multiuser power control in digital subscriber lines," *EURASIP Appl. Signal Processing*, vol. 2006, no. 1, pp. 1–10, Apr. 2006, Article ID 24012.

[12] John Papandriopoulos and Jamie S. Evans, "Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks," in *Proc. IEEE Int. Conf. Commun.*, Istanbul, June 2006, vol. 9, pp. 3270–3275.

[13] Shunsuke Hayashi and Zhi-Quan Luo, "Dynamic spectrum management: When is FDMA sum-rate optimal?," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Honolulu, Hawaii, Apr. 2007, pp. 15–20.

[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.

[15] Dimitri P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Nashua, NH, 2 edition, 2003.

[16] Herbert Gintis, *Game theory evolving: a problem-centered introduction to modeling strategic behavior*, Princeton Press, New Jersy, 2000.

[17] Gesualdo Scutari, Daniel P. Palomar, and Sergio Barbarossa, "Asynchronous iterative waterfilling for Gaussian frequency-selective interference channels," *IEEE Trans. Inform. Theory*, Jan. 2008, Accepted.

[18] D. P. Bertsekas and T. N. Tsitsiklis, *Parallel and distributed computation: Numerical methods*, Athena Scientific, Nashua, NH, 2 edition, 1989.

[19] Gesualdo Scutari, Daniel P. Palomar, and Sergio Barbarossa, "Optimal linear precoding strategies for wideband noncooperative systems based on Game Theory—Part I: Nash equilibria," *IEEE Trans. Signal Processing*, vol. 56, no. 3, pp. 1230–1249, Mar. 2008.

[20] Raul Etkin, Abhay Parekh, and David Tse, "Spectrum sharing for unlicensed bands," *IEEE J. Select. Areas Commun.*, vol. 25, no. 3, pp. 517–528, Apr. 2007.

[21] Jong-Shi Pang, Gesualdo Scutari, Francisco Facchinei, and Chaoxiong Wang, "Distributed power allocation with rate constraints in Gaussian frequency-selective interference channels," Jan. 2007, Submitted.

[22] A. Kashyap, T. Başar, and R. Srikant, "Correlated jamming on MIMO Gaussian fading channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 2119–2123, Sept. 2004.

[23] Shahrokh Farahmand, Georgios Giannakis, and Xin Wang, "Max-min strategies for power-limited games in the presence of correlated jamming," in *41st Ann. Conf. Info. Sci. Systms.*, Baltimore, Mar. 2007, pp. 300–305.

[24] Shabnam Shafiee and Sennur Ulukus, "Capacity of multiple access channels with correlated jamming," in *IEEE Milit. Commun. Conf.*, Atlantic City, New Jersey, Oct. 2005, vol. 1, pp. 218–224.

[25] Ramy H. Gohary, Yao Huang (Morin), and Zhi-Quan Luo, "A generalized iterative water-filling algorithm for distributed power control in the presence of a jammer," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Taipei, Taiwan, Apr. 2009, pp. 2373–2376.

[26] Zhi-Quan Luo and Shuzhong Zhang, "Duality gap estimation and polynomial time approximation for optimal spectrum management," Submited.

[27] Wei Yu and Raymond Liu, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, Feb. 2006.

[28] Shunsuke Hayashi and Zhi-Quan Luo, "Dynamic spectrum management: When is fdma sum-rate optimal?," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Honolulu, Hawaii, Apr. 2007, pp. 15–20.

[29] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[30] Yao Huang (Morin), Ramy H. Gohary, and Zhi-Quan Luo, "Structured spectrum balancing in DSL multiuser communications," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Taipei, Taiwan, Apr. 2009, pp. 1004–1008.

[31] Paschalis Tsiaflakis, Moritz Diehl, and Marc Moonen, "Distributed spectrum management algorithms for multi-user dsl netowrks," Accepted by IEEE Transactions on Signal Processing.

[32] Wei Yu and Raymond Liu, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[33] Yao Huang (Morin), Ramy H. Gohary, and Zhi-Quan Luo, "Approaching user capacity in a DSL system via harmonic mean-rate optimization," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Taipei, Taiwan, Apr. 2009, pp. 1000–1004.

[34] F. Facchinei and J. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problem*, Springer-Verlag, New York, 2003.

[35] I. V. Konnov, *Equilibrium models and variational inequalities*, vol. 210 of *Math. Sci. Eng.*, Elsevier, Boston, 2007.

[36] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1994.

Appendix A

Appendix

A.1 Proof of Theorem 1

Recall that we use $s_k^{n,\nu}$ and $s_k^{n,*}$ to denote the power allocated by User $k \in \{0\} \cup \mathcal{K}$ to the n-th tone at the ν-th iteration and at equilibrium, respectively. For the updates of User $k \in \mathcal{K}$, it was shown in [11] that each iteration of the IWFA algorithm in (5.1.6) is equivalent to solving the following fixed-point equation.

$$\begin{bmatrix} s_k^{1,\nu} \\ \vdots \\ s_k^{N,\nu} \end{bmatrix} = \begin{bmatrix} s_k^{1,\nu} - \sigma_k^1 - \sum_{j=1}^{k} \alpha_{jk}^1 s_j^{1,\nu} - \sum_{j=k+1}^{K} \alpha_{jk}^1 s_j^{1,\nu-1} - \alpha_{0k}^1 s_0^{1,\nu-1} \\ \vdots \\ s_k^{N,\nu} - \sigma_k^N - \sum_{j=1}^{k} \alpha_{jk}^N s_j^{N,\nu} - \sum_{j=k+1}^{K} \alpha_{jk}^N s_j^{N,\nu-1} - \alpha_{0k}^N s_0^{N,\nu-1} \end{bmatrix}_{\hat{\mathcal{P}}_k}$$

$$= \begin{bmatrix} -\sigma_k^1 - \sum_{j=1}^{k-1} \alpha_{jk}^1 s_j^{1,\nu} - \sum_{j=k+1}^{K} \alpha_{jk}^1 s_j^{1,\nu-1} - \alpha_{0k}^1 s_0^{1,\nu-1} \\ \vdots \\ -\sigma_k^N - \sum_{j=1}^{k-1} \alpha_{jk}^N s_j^{N,\nu} - \sum_{j=k+1}^{K} \alpha_{jk}^N s_j^{N,\nu-1} - \alpha_{0k}^N s_0^{N,\nu-1} \end{bmatrix}_{\hat{\mathcal{P}}_k}, \qquad \text{(A.1)}$$

where in (A.1) we have used that $\alpha_{kk}^n = 1$ for all $n \in \mathcal{N}$, and $[\cdot]_{\hat{\mathcal{P}}_k}$ to denote the projection onto the polyhedron

$$\hat{\mathcal{P}}_k = \Big\{(s_k^1, \cdots, s_k^N) | 0 \leq s_k^n \leq S_{\max,k}^n, n = 1, \ldots, N, \sum_{n=1}^{N} s_k^n = P_k\Big\}. \tag{A.2}$$

Note that, in contrast with the polyhedron in (5.0.4), in the polyhedron in (A.2), the power constraint is satisfied with equality.

Now, in a similar fashion, the jammer can update its power in order to solve

$$\begin{bmatrix} s_0^{1,\nu} \\ \vdots \\ s_0^{N,\nu} \end{bmatrix} = \begin{bmatrix} s_0^{1,\nu} + \sum_{k=1}^{K} \frac{\alpha_{0k}^1 s_k^{1,\nu}}{\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^1 s_j^{1,\nu} + \sigma_k^1\right)\left(\sum_{j=1}^{K} \alpha_{jk}^1 s_j^{1,\nu} + \alpha_{0k}^1 s_0^{1,\nu} + \sigma_k^1\right)} \\ \vdots \\ s_0^{N,\nu} + \sum_{k=1}^{K} \frac{\alpha_{0k}^N s_k^{N,\nu}}{\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^N s_j^{N,\nu} + \sigma_k^N\right)\left(\sum_{j=1}^{K} \alpha_{jk}^N s_j^{N,\nu} + \alpha_{0k}^N s_0^{N,\nu} + \sigma_k^N\right)} \end{bmatrix}_{\hat{\mathcal{P}}_0}, \tag{A.3}$$

where the set $\hat{\mathcal{P}}_0$ is defined in a fashion similar to (A.2).

Let $s_k^{n,*}$ be the power allocation at equilibrium of User $k \in \mathcal{K}$, at tone $n \in \mathcal{N}$. Furthermore, let

$$t_k^{n,\nu} = s_k^{n,\nu} - s_k^{n,*}, \quad \forall k \in \mathcal{K}, \qquad \text{and} \qquad r^{n,\nu} = s_0^{n,\nu} - s_0^{n,*}. \tag{A.4}$$

At equilibrium we have

$$\begin{bmatrix} s_k^{1,*} \\ \vdots \\ s_k^{N,*} \end{bmatrix} = \begin{bmatrix} -\sigma_k^1 - \sum_{j=1}^{k-1} \alpha_{jk}^1 s_j^{1,*} - \sum_{j=k+1}^{K} \alpha_{jk}^1 s_j^{1,*} - \alpha_{0k}^1 s_0^{1,*} \\ \vdots \\ -\sigma_k^N - \sum_{j=1}^{k-1} \alpha_{jk}^N s_j^{N,*} - \sum_{j=k+1}^{K} \alpha_{jk}^N s_j^{N,*} - \alpha_{0k}^N s_0^{N,*} \end{bmatrix}_{\hat{\mathcal{P}}_k}, \tag{A.5}$$

and

$$\begin{bmatrix} s_0^{1,*} \\ \vdots \\ s_0^{N,*} \end{bmatrix} = \begin{bmatrix} s_0^{1,*} + \sum_{k=1}^{K} \frac{\alpha_{0k}^1 s_k^{1,*}}{\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^1 s_j^{1,*} + \sigma_k^1\right)\left(\sum_{j=1}^{K} \alpha_{jk}^1 s_j^{1,*} + \alpha_{0k}^1 s_0^{1,*} + \sigma_k^1\right)} \\ \vdots \\ s_0^{N,*} + \sum_{k=1}^{K} \frac{\alpha_{0k}^N s_k^{N,*}}{\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^N s_j^{N,*} + \sigma_k^N\right)\left(\sum_{j=1}^{K} \alpha_{jk}^N s_j^{N,*} + \alpha_{0k}^N s_0^{N,*} + \sigma_k^N\right)} \end{bmatrix}_{\hat{\mathcal{P}}_0} . \tag{A.6}$$

We now subtract (A.5) from (A.1), and (A.6) from (A.3). Using the non-expansiveness property of the projection operator [11], one can write

$$\begin{aligned} \left\| \begin{bmatrix} t_k^{1,\nu} \\ \vdots \\ t_k^{N,\nu} \end{bmatrix} \right\| &\leq \left\| \begin{bmatrix} -\alpha_{0k}^1 r^{1,\nu-1} - \sum_{j=1}^{k-1} \alpha_{jk}^1 t_j^{1,\nu} - \sum_{j=k+1}^{K} \alpha_{jk}^1 t_j^{1,\nu-1} \\ \vdots \\ -\alpha_{0k}^N r^{N,\nu-1} - \sum_{j=1}^{k-1} \alpha_{jk}^N t_j^{N,\nu} - \sum_{j=k+1}^{K} \alpha_{jk}^N t_j^{N,\nu-1} \end{bmatrix} \right\| \\ &\leq \left\| \begin{bmatrix} \alpha_{0k}^1 r^{1,\nu-1} \\ \vdots \\ \alpha_{0k}^N r^{N,\nu-1} \end{bmatrix} \right\| + \left\| \sum_{j=1}^{k-1} \begin{bmatrix} \alpha_{jk}^1 t_j^{1,\nu} \\ \vdots \\ \alpha_{jk}^N t_j^{N,\nu} \end{bmatrix} \right\| + \left\| \sum_{j=k+1}^{K} \begin{bmatrix} \alpha_{jk}^1 t_j^{1,\nu-1} \\ \vdots \\ \alpha_{jk}^N t_j^{N,\nu-1} \end{bmatrix} \right\| \\ &\leq \alpha_{0,k} \|r^{\nu-1}\| + \sum_{j=1}^{k-1} \alpha_{jk} \|t_j^{\nu}\| + \sum_{j=k+1}^{K} \alpha_{jk} \|t_j^{\nu-1}\|, \end{aligned} \tag{A.7}$$

where in (A.7) we have used t_k^{ν} and r^{ν} to denote the vectors $[t_k^{1,\nu}, \cdots, t_k^{N,\nu}]^T$ and $[r^{1,\nu}, \cdots, r^{N,\nu}]^T$, respectively, and α_{jk} to denote $\left\|[\alpha_{jk}^1, \cdots, \alpha_{jk}^N]\right\|_2$.

Using a technique similar to the one in [11] we can express the inequalities in (A.7) for all users simultaneously in the following matrix form.

$$A \begin{bmatrix} \|t_1^{\nu}\| \\ \vdots \\ \|t_K^{\nu}\| \end{bmatrix} \leq \begin{bmatrix} B & \beta \end{bmatrix} \begin{bmatrix} \|t_1^{\nu-1}\| \\ \vdots \\ \|t_K^{\nu-1}\| \\ \|r^{\nu-1}\| \end{bmatrix}, \tag{A.8}$$

where

$$A = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -\alpha_{12} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\alpha_{1K} & -\alpha_{2K} & \cdots & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & \alpha_{21} & \alpha_{31} & \cdots & \alpha_{K1} \\ 0 & 0 & \alpha_{32} & \cdots & \alpha_{K2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha_{K,K-1} \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \text{and} \quad \beta = \begin{bmatrix} \alpha_{01} \\ \vdots \\ \alpha_{0K} \end{bmatrix}, \tag{A.9}$$

where the inequality in (A.8) is to be interpreted element-wise. Notice that A is a non-singular Z matrix with (entry-wise) non-negative inverse. The matrix B and the vector β are also non-negative. Hence using [36, Property 2.5.3.18], we have that (A.8) imply that

$$\begin{bmatrix} \|t_1^\nu\| \\ \vdots \\ \|t_K^\nu\| \end{bmatrix} \leq \begin{bmatrix} A^{-1}B & A^{-1}\beta \end{bmatrix} \begin{bmatrix} \|t_1^{\nu-1}\| \\ \vdots \\ \|t_K^{\nu-1}\| \\ \|r^{\nu-1}\| \end{bmatrix}. \tag{A.10}$$

If we use t^ν to denote the vector $\begin{bmatrix} \|t_1^\nu\| & \cdots & \|t_K^\nu\| \end{bmatrix}^T$, then (A.10) implies that

$$\|t^\nu\| \leq \begin{bmatrix} \|A^{-1}B\|_2 & \|A^{-1}\beta\| \end{bmatrix} \begin{bmatrix} \|t^{\nu-1}\| \\ \|r^{\nu-1}\| \end{bmatrix}. \tag{A.11}$$

We now turn our attention to the jammer's updates; cf. (A.3). In order to simplify our exposition, we will use the following notation.

$$\begin{aligned} c_k^{n,*} &= \tfrac{1}{\alpha_{0k}^n} \sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n s_j^{n,*} + \tfrac{\sigma_k^n}{\alpha_{0k}^n}, \\ c_k^{n,\nu} &= \tfrac{1}{\alpha_{0k}^n} \sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n s_j^{n,\nu} + \tfrac{\sigma_k^n}{\alpha_{0k}^n}, \end{aligned} \tag{A.12}$$

$$\begin{aligned} d_k^{n,*} &= c_k^{n,*} + \tfrac{s_k^{n,*}}{\alpha_{0k}^n}, \\ d_k^{n,\nu} &= c_k^{n,\nu} + \tfrac{s_k^{n,\nu}}{\alpha_{0k}^n}. \end{aligned} \tag{A.13}$$

Using a technique similar to the one used for the users' updates and employing the non-expansiveness property of the projection operator, we use (A.3) to write

$$\left\| \begin{bmatrix} r^{1,\nu} \\ \vdots \\ r^{N,\nu} \end{bmatrix} \right\| \leq \left\| \begin{bmatrix} r^{1,\nu} + \sum_{k=1}^{K} \frac{s_k^{1,\nu}}{(c_k^{1,\nu}+s_0^{1,\nu})(d_k^{1,\nu}+s_0^{1,\nu})} - \frac{s_k^{1,*}}{(c_k^{1,*}+s_0^{1,*})(d_k^{1,*}+s_0^{1,*})} \\ \vdots \\ r^{N,\nu} + \sum_{k=1}^{K} \frac{s_k^{N,\nu}}{(c_k^{N,\nu}+s_0^{N,\nu})(d_k^{N,\nu}+s_0^{N,\nu})} - \frac{s_k^{N,*}}{(c_k^{N,*}+s_0^{N,*})(d_k^{N,*}+s_0^{N,*})} \end{bmatrix} \right\|. \tag{A.14}$$

Using partial fraction expansion, the n-th entry of the vector on the right hand side of (A.14) can be written as

$$\begin{aligned} r^{n,\nu} \left(1 - \sum_{k=1}^{K} \left(\frac{\alpha_{0k}^n}{(c_k^{n,\nu} + s_0^{n,*})(c_k^{n,\nu} + s_0^{n,*} + r^{n,\nu})} - \frac{\alpha_{0k}^n}{(d_k^{n,\nu} + s_0^{n,*})(d_k^{n,\nu} + s_0^{n,*} + r^{n,\nu})} \right) \right) \\ - \sum_{k=1}^{K} \left(\frac{s_k^{n,*}}{(c_k^{n,*} + s_0^{n,*})(d_k^{n,*} + s_0^{n,*})} - \frac{s_k^{n,\nu}}{(c_k^{n,\nu} + s_0^{n,*})(d_k^{n,\nu} + s_0^{n,*})} \right). \end{aligned} \tag{A.15}$$

Let Υ^ν be an $N \times N$ diagonal matrix with the n-th diagonal entry given by

$$\left(1 - \sum_{k=1}^{K} \left(\frac{\alpha_{0k}^n}{(c_k^{n,\nu} + s_0^{n,*})(c_k^{n,\nu} + s_0^{n,*} + r^{n,\nu})} - \frac{\alpha_{0k}^n}{(d_k^{n,\nu} + s_0^{n,*})(d_k^{n,\nu} + s_0^{n,*} + r^{n,\nu})} \right) \right). \tag{A.16}$$

Furthermore, let γ_k^ν be an N-dimensional vector whose n-th entry is given by

$$\gamma_k^{n,\nu} = \left| \frac{s_k^{n,*}}{(c_k^{n,*} + s_0^{n,*})(d_k^{n,*} + s_0^{n,*})} - \frac{s_k^{n,\nu}}{(c_k^{n,\nu} + s_0^{n,*})(d_k^{n,\nu} + s_0^{n,*})} \right|. \tag{A.17}$$

Now, (A.14) can be bounded as follows

$$\begin{aligned} \|r^\nu\| &\leq \|\Upsilon^\nu r^\nu\| + \left\| \sum_{k=1}^K \gamma_k^\nu \right\| \\ &\leq \|\Upsilon^\nu\|_2 \|r^\nu\| + \left\| \sum_{k=1}^K \gamma_k^\nu \right\|. \end{aligned}$$

Assuming that $\|\Upsilon^\nu\|_2 < 1$, then we have

$$\|r^\nu\| \leq \big(1 - \|\Upsilon^\nu\|_2\big)^{-1} \left\| \sum_{k=1}^K \gamma_k^\nu \right\|. \tag{A.18}$$

In order to analyze the matrix Υ^ν and the vectors $\{\gamma_k^\nu\}$, we will need a lower bound on $s_k^{n,\nu}$. In Appendix A.2 we provide a lower bound η_k^n such that $0 \leq \eta_k^n \leq s_k^{n,\nu}$, for all iterations ν, $k \in \mathcal{K}$, $n \in \mathcal{N}$. Using this value of η_k^n, we can readily derive a lower bound on $c_k^{n,\nu}$. In particular, if we let $c_{\min,k}^n$ denote this bound, then it follows from (A.12) that

$$c_k^{n,\nu} = \tfrac{1}{\alpha_{0k}^n} \sum_{j=1,\, j\neq k}^K \alpha_{jk}^n s_j^{n,\nu} + \tfrac{\sigma_k^n}{\alpha_{0k}^n} \geq \frac{1}{\alpha_{0k}^n} \left(\sum_{j=1,\, j\neq k}^K \alpha_{jk}^n \eta_j^n + \sigma_k^n \right) \triangleq c_{\min,k}^n. \tag{A.19}$$

Similarly, a lower bound on $d_k^{n,\nu}$ can be derived from (A.13)

$$d_{\min,k}^n \triangleq c_{\min,k}^n + \frac{1}{\alpha_{0k}^n} \eta_k^n. \tag{A.20}$$

Now that we have a lower bound on $s_k^{n,\nu}$, $c_k^{n,\nu}$ and $d_k^{n,\nu}$, we can proceed to analyze

$\gamma_k^{n,\nu}$ in (A.17). From (A.17), our goal is to bound $\{\gamma_k^{n,\nu}\}$ as a linear combination of $|t_j^{n,\nu}|_{j=1}^K$. This requires some detailed computation which we present below. By definition, we have

$$\begin{aligned}
\gamma_k^{n,\nu} &= \frac{|s_k^{n,*}(c_k^{n,\nu}+s_0^{n,*})(d_k^{n,\nu}+s_0^{n,*}) - s_k^{n,\nu}(c_k^{n,*}+s_0^{n,*})(d_k^{n,*}+s_0^{n,*})|}{(c_k^{n,*}+s_0^{n,*})(d_k^{n,*}+s_0^{n,*})(c_k^{n,\nu}+s_0^{n,*})(d_k^{n,\nu}+s_0^{n,*})} \\
&\le \frac{(s_0^{n,*})^2|s_k^{n,*}-s_k^{n,\nu}| + s_0^{n,*}\left|s_k^{n,*}(c_k^{n,\nu}+d_k^{n,\nu}) - s_k^{n,\nu}(c_k^{n,*}+d_k^{n,*})\right|}{(c_k^{n,*}+s_0^{n,*})(d_k^{n,*}+s_0^{n,*})(c_k^{n,\nu}+s_0^{n,*})(d_k^{n,\nu}+s_0^{n,*})} \\
&\qquad + \frac{\left|s_k^{n,*}c_k^{n,\nu}d_k^{n,\nu} - s_k^{n,\nu}c_k^{n,*}d_k^{n,*}\right|}{(c_k^{n,*}+s_0^{n,*})(d_k^{n,*}+s_0^{n,*})(c_k^{n,\nu}+s_0^{n,*})(d_k^{n,\nu}+s_0^{n,*})} \\
&\le \frac{(s_0^{n,*})^2|s_k^{n,*}-s_k^{n,\nu}|}{d_k^{n,\nu}d_k^{n,*}(c_k^{n,*}+s_0^{n,*})(c_k^{n,\nu}+s_0^{n,*})} + \frac{s_0^{n,*}\left|s_k^{n,*}(c_k^{n,\nu}+d_k^{n,\nu}) - s_k^{n,\nu}(c_k^{n,*}+d_k^{n,*})\right|}{c_k^{n,*}d_k^{n,*}d_k^{n,\nu}(c_k^{n,\nu}+s_0^{n,*})} \\
&\qquad + \frac{\left|s_k^{n,*}c_k^{n,\nu}d_k^{n,\nu} - s_k^{n,\nu}c_k^{n,*}d_k^{n,*}\right|}{(c_k^{n,*}+s_0^{n,*})(d_k^{n,*}+s_0^{n,*})(c_k^{n,\nu}+s_0^{n,*})(d_k^{n,\nu}+s_0^{n,*})} \qquad \text{(A.21)} \\
&\le \frac{(S_{\max,0}^n)^2|s_k^{n,*}-s_k^{n,\nu}|}{d_k^{n,\nu}d_k^{n,*}(c_k^{n,*}+S_{\max,0}^n)(c_k^{n,\nu}+S_{\max,0}^n)} + \frac{S_{\max,0}^n\left|s_k^{n,*}(c_k^{n,\nu}+d_k^{n,\nu}) - s_k^{n,\nu}(c_k^{n,*}+d_k^{n,*})\right|}{c_k^{n,*}d_k^{n,*}d_k^{n,\nu}(c_k^{n,\nu}+S_{\max,0}^n)} \\
&\qquad + \frac{\left|s_k^{n,*}c_k^{n,\nu}d_k^{n,\nu} - s_k^{n,\nu}c_k^{n,*}d_k^{n,*}\right|}{(c_k^{n,*}+s_0^{n,*})(d_k^{n,*}+s_0^{n,*})(c_k^{n,\nu}+s_0^{n,*})(d_k^{n,\nu}+s_0^{n,*})} \qquad \text{(A.22)} \\
&\le \frac{(S_{\max,0}^n)^2|t_k^{n,\nu}|}{(d_{\min,k}^n)^2(c_{\min,k}^n+S_{\max,0}^n)^2} + \frac{S_{\max,0}^n\left|s_k^{n,*}(c_k^{n,\nu}+d_k^{n,\nu}) - s_k^{n,\nu}(c_k^{n,*}+d_k^{n,*})\right|}{c_k^{n,*}d_k^{n,*}d_k^{n,\nu}(c_k^{n,\nu}+S_{\max,0}^n)} \\
&\qquad + \frac{\left|s_k^{n,*}c_k^{n,\nu}d_k^{n,\nu} - s_k^{n,\nu}c_k^{n,*}d_k^{n,*}\right|}{c_k^{n,*}d_k^{n,*}c_k^{n,\nu}d_k^{n,\nu}}, \qquad \text{(A.23)}
\end{aligned}$$

where in (A.22) we have used the fact that both the first and the second term of (A.21) are monotone increasing in $s_0^{n,*}$.

Next we bound the third and second term in (A.23) separately. Let $a_k^{n,\nu}$ denote the third term of (A.23). Then, using the definition of $d_k^{n,\nu}$ in (A.13), we obtain

$$
\begin{aligned}
a_k^{n,\nu} &= \frac{\left|s_k^{n,*}c_k^{n,\nu}d_k^{n,\nu} - s_k^{n,\nu}c_k^{n,*}d_k^{n,*}\right|}{c_k^{n,*}d_k^{n,*}c_k^{n,\nu}d_k^{n,\nu}} \\
&= \frac{\left|s_k^{n,*}c_k^{n,\nu}(c_k^{n,\nu} + \frac{s_k^{n,\nu}}{\alpha_{0k}^n}) - s_k^{n,\nu}c_k^{n,*}(c_k^{n,*} + \frac{s_k^{n,*}}{\alpha_{0k}^n})\right|}{c_k^{n,*}d_k^{n,*}c_k^{n,\nu}d_k^{n,\nu}} \\
&= \frac{\left|\frac{s_k^{n,*}s_k^{n,\nu}}{\alpha_{0k}^n}(c_k^{n,\nu} - c_k^{n,*}) + s_k^{n,*}(c_k^{n,\nu})^2 - s_k^{n,\nu}(c_k^{n,*})^2\right|}{c_k^{n,*}d_k^{n,*}c_k^{n,\nu}d_k^{n,\nu}} \\
&\le \frac{\frac{s_k^{n,*}s_k^{n,\nu}}{\alpha_{0k}^n}|c_k^{n,\nu} - c_k^{n,*}| + |s_k^{n,*}(c_k^{n,\nu})^2 - (s_k^{n,*} + t_k^{n,\nu})(c_k^{n,*})^2|}{c_k^{n,*}d_k^{n,*}c_k^{n,\nu}d_k^{n,\nu}} \\
&\le \frac{\frac{s_k^{n,*}s_k^{n,\nu}}{\alpha_{0k}^n}|c_k^{n,\nu} - c_k^{n,*}| + s_k^{n,*}|(c_k^{n,\nu})^2 - (c_k^{n,*})^2| + |t_k^{n,\nu}|(c_k^{n,*})^2}{c_k^{n,*}d_k^{n,*}c_k^{n,\nu}d_k^{n,\nu}} \\
&= \frac{s_k^{n,*}s_k^{n,\nu}|c_k^{n,\nu} - c_k^{n,*}|}{\alpha_{0k}^n c_k^{n,*}c_k^{n,\nu}(c_k^{n,*} + \frac{s_k^{n,*}}{\alpha_{0k}^n})(c_k^{n,\nu} + \frac{s_k^{n,\nu}}{\alpha_{0k}^n})} + \frac{s_k^{n,*}(c_k^{n,\nu} + c_k^{n,*})|c_k^{n,\nu} - c_k^{n,*}|}{c_k^{n,*}c_k^{n,\nu}(c_k^{n,*} + \frac{s_k^{n,*}}{\alpha_{0k}^n})d_k^{n,\nu}} + \frac{|t_k^{n,\nu}|c_k^{n,*}}{d_k^{n,*}c_k^{n,\nu}d_k^{n,\nu}}
\end{aligned}
\tag{A.24}
$$

$$
\le \frac{(S_{\max,k}^n)^2|c_k^{n,\nu} - c_k^{n,*}|}{\alpha_{0k}^n(c_{\min,k}^n)^2(c_{\min,k}^n + \frac{S_{\max,k}^n}{\alpha_{0k}^n})^2} + \frac{S_{\max,k}^n(c_k^{n,\nu} + c_k^{n,*})|c_k^{n,\nu} - c_k^{n,*}|}{c_k^{n,*}c_k^{n,\nu}(c_k^{n,*} + \frac{S_{\max,k}^n}{\alpha_{0k}^n})d_k^{n,\nu}} + \frac{|t_k^{n,\nu}|c_k^{n,*}}{d_k^{n,*}c_k^{n,\nu}d_k^{n,\nu}}
\tag{A.25}
$$

$$
\begin{aligned}
&= \frac{(S_{\max,k}^n)^2|c_k^{n,\nu} - c_k^{n,*}|}{\alpha_{0k}^n(c_{\min,k}^n)^2(c_{\min,k}^n + \frac{S_{\max,k}^n}{\alpha_{0k}^n})^2} + \frac{S_{\max,k}^n|c_k^{n,\nu} - c_k^{n,*}|}{c_k^{n,\nu}(c_k^{n,*} + \frac{S_{\max,k}^n}{\alpha_{0k}^n})d_k^{n,\nu}} + \frac{S_{\max,k}^n|c_k^{n,\nu} - c_k^{n,*}|}{c_k^{n,*}(c_k^{n,*} + \frac{S_{\max,k}^n}{\alpha_{0k}^n})d_k^{n,\nu}} \\
&\qquad + \frac{|t_k^{n,\nu}|c_k^{n,*}}{(c_k^{n,*} + \frac{s_k^{n,*}}{\alpha_{0k}^n})c_k^{n,\nu}d_k^{n,\nu}} \\
&\le \frac{(S_{\max,k}^n)^2|c_k^{n,\nu} - c_k^{n,*}|}{\alpha_{0k}^n(c_{\min,k}^n)^2(c_{\min,k}^n + \frac{S_{\max,k}^n}{\alpha_{0k}^n})^2} + \frac{2S_{\max,k}^n|c_k^{n,\nu} - c_k^{n,*}|}{c_{\min,k}^n d_{\min,k}^n(c_{\min,k}^n + \frac{S_{\max,k}^n}{\alpha_{0k}^n})} + \frac{|t_k^{n,\nu}|c_k^{n,*}}{(c_k^{n,*} + \frac{s_k^{n,*}}{\alpha_{0k}^n})c_{\min,k}^n d_{\min,k}^n}
\end{aligned}
\tag{A.26}
$$

$$
\begin{aligned}
&\le \left(\frac{(S_{\max,k}^n)^2}{(\alpha_{0k}^n)^2(c_{\min,k}^n)^2(c_{\min,k}^n + \frac{S_{\max,k}^n}{\alpha_{0k}^n})^2} + \frac{2S_{\max,k}^n}{\alpha_{0k}^n c_{\min,k}^n d_{\min,k}^n(c_{\min,k}^n + \frac{S_{\max,k}^n}{\alpha_{0k}^n})}\right)\sum_{j=1,\, j\neq k}^{K}\alpha_{jk}^n|t_j^{n,\nu}| \\
&\qquad + \frac{|t_k^{n,\nu}|\sum_{j=1,\, j\neq k}^{K}\alpha_{jk}^n S_{\max,j}^n}{(\sum_{j=1,\, j\neq k}^{K}\alpha_{jk}^n S_{\max,j}^n + s_k^{n,*})c_{\min,k}^n d_{\min,k}^n},
\end{aligned}
\tag{A.27}
$$

$$
\begin{aligned}
&\le \frac{(S_{\max,k}^n)^2 d_{\min,k}^n + 2S_{\max,k}^n c_{\min,k}^n(\alpha_{0k}^n c_{\min,k}^n + S_{\max,k}^n)}{d_{\min,k}^n(c_{\min,k}^n)^2(\alpha_{0k}^n c_{\min,k}^n + S_{\max,k}^n)^2}\sum_{j=1,\, j\neq k}^{K}\alpha_{jk}^n|t_j^{n,\nu}| \\
&\qquad + \frac{\sum_{j=1,\, j\neq k}^{K}\alpha_{jk}^n S_{\max,j}^n}{(\sum_{j=1,\, j\neq k}^{K}\alpha_{jk}^n S_{\max,j}^n + \eta_k^n)c_{\min,k}^n d_{\min,k}^n}|t_k^{n,\nu}|,
\end{aligned}
\tag{A.28}
$$

where in (A.25), we have used the fact that the first term in (A.24) is monotonically increasing in both $s_k^{n,*}$ and $s_k^{n,\nu}$, and that the second term in (A.24) is monotonically increasing in $s_k^{n,*}$. Similarly, in (A.27) we have used the fact that in (A.26), the last term is monotonically increasing in $c_k^{n,*}$.

We now consider the second term in (A.23). Denoting this term by $b_k^{n,\nu}$, we have,

$$
\begin{aligned}
b_k^{n,\nu} &= \frac{S_{\max,0}^n \left| s_k^{n,*}(c_k^{n,\nu} + d_k^{n,\nu}) - s_k^{n,\nu}(c_k^{n,*} + d_k^{n,*}) \right|}{c_k^{n,*} d_k^{n,*} d_k^{n,\nu} (c_k^{n,\nu} + S_{\max,0}^n)} \\
&= \frac{S_{\max,0}^n \left| s_k^{n,*}(c_k^{n,\nu} + d_k^{n,\nu}) - (s_k^{n,*} + t_k^{n,\nu})(c_k^{n,*} + d_k^{n,*}) \right|}{c_k^{n,*} d_k^{n,*} d_k^{n,\nu} (c_k^{n,\nu} + S_{\max,0}^n)} \\
&= \frac{S_{\max,0}^n \left| s_k^{n,*}(c_k^{n,\nu} + d_k^{n,\nu} - c_k^{n,*} - d_k^{n,*}) - t_k^{n,\nu}(c_k^{n,*} + d_k^{n,*}) \right|}{c_k^{n,*} d_k^{n,*} d_k^{n,\nu} (c_k^{n,\nu} + S_{\max,0}^n)} \\
&\le \frac{S_{\max,0}^n s_k^{n,*} |c_k^{n,\nu} + d_k^{n,\nu} - c_k^{n,*} - d_k^{n,*}|}{c_k^{n,*} d_k^{n,*} d_k^{n,\nu} (c_k^{n,\nu} + S_{\max,0}^n)} + \frac{S_{\max,0}^n |t_k^{n,\nu}| (c_k^{n,*} + d_k^{n,*})}{c_k^{n,*} d_k^{n,*} d_k^{n,\nu} (c_k^{n,\nu} + S_{\max,0}^n)} \\
&= \frac{S_{\max,0}^n s_k^{n,*} |c_k^{n,\nu} + d_k^{n,\nu} - c_k^{n,*} - d_k^{n,*}|}{c_k^{n,*} (c_k^{n,*} + \frac{s_k^{n,*}}{\alpha_{0k}^n}) d_k^{n,\nu} (c_k^{n,\nu} + S_{\max,0}^n)} + \frac{S_{\max,0}^n |t_k^{n,\nu}|}{d_k^{n,*} d_k^{n,\nu} (c_k^{n,\nu} + S_{\max,0}^n)} + \frac{S_{\max,0}^n |t_k^{n,\nu}|}{c_k^{n,*} d_k^{n,\nu} (c_k^{n,\nu} + S_{\max,0}^n)} \\
&\le \frac{S_{\max,0}^n s_k^{n,*} \left(2 \sum_{j=1,\, j\neq k}^K \alpha_{jk}^n |t_j^{n,\nu}| + |t_k^{n,\nu}|\right)}{\alpha_{0k}^n c_k^{n,*} (c_k^{n,*} + \frac{s_k^{n,*}}{\alpha_{0k}^n}) d_k^{n,\nu} (c_k^{n,\nu} + S_{\max,0}^n)} + \frac{S_{\max,0}^n |t_k^{n,\nu}|}{d_k^{n,*} d_k^{n,\nu} (c_k^{n,\nu} + S_{\max,0}^n)} + \frac{S_{\max,0}^n |t_k^{n,\nu}|}{c_k^{n,*} d_k^{n,\nu} (c_k^{n,\nu} + S_{\max,0}^n)}
\end{aligned}
\tag{A.29}
$$

$$
\begin{aligned}
&\le \frac{S_{\max,0}^n S_{\max,k}^n \left(2 \sum_{j=1,\, j\neq k}^K \alpha_{jk}^n |t_j^{n,\nu}| + |t_k^{n,\nu}|\right)}{c_k^{n,*} (\alpha_{0k}^n c_k^{n,*} + S_{\max,k}^n) d_k^{n,\nu} (c_k^{n,\nu} + S_{\max,0}^n)} + \frac{S_{\max,0}^n |t_k^{n,\nu}|}{d_k^{n,*} d_k^{n,\nu} (c_k^{n,\nu} + S_{\max,0}^n)} \\
&\quad + \frac{S_{\max,0}^n |t_k^{n,\nu}|}{c_k^{n,*} d_k^{n,\nu} (c_k^{n,\nu} + S_{\max,0}^n)}
\end{aligned}
\tag{A.30}
$$

$$
\begin{aligned}
&\le \frac{S_{\max,0}^n S_{\max,k}^n \left(2 \sum_{j=1,\, j\neq k}^K \alpha_{jk}^n |t_j^{n,\nu}| + |t_k^{n,\nu}|\right)}{c_{\min,k}^n (\alpha_{0k}^n c_{\min,k}^n + S_{\max,k}^n) d_{\min,k}^n (c_{\min,k}^n + S_{\max,0}^n)} \\
&\quad + \frac{S_{\max,0}^n}{d_{\min,k}^n (c_{\min,k}^n + S_{\max,0}^n)} \left(\frac{1}{d_{\min,k}^n} + \frac{1}{c_{\min,k}^n}\right) |t_k^{n,\nu}| \\
&= \frac{2 S_{\max,0}^n S_{\max,k}^n}{c_{\min,k}^n (\alpha_{0k}^n c_{\min,k}^n + S_{\max,k}^n) d_{\min,k}^n (c_{\min,k}^n + S_{\max,0}^n)} \sum_{j=1,\, j\neq k}^K \alpha_{jk}^n |t_j^{n,\nu}| \\
&\quad + \frac{S_{\max,0}^n}{d_{\min,k}^n (c_{\min,k}^n + S_{\max,0}^n)} \left(\frac{1}{d_{\min,k}^n} + \frac{1}{c_{\min,k}^n} + \frac{S_{\max,k}^n}{c_{\min,k}^n (\alpha_{0k}^n c_{\min,k}^n + S_{\max,k}^n)}\right) |t_k^{n,\nu}|,
\end{aligned}
\tag{A.31}
$$

where in (A.30), we have used that the first term in (A.29) is monotonically increasing in $s_k^{n,*}$.

Using the bounds on $a_k^{n,\nu}$ and $b_k^{n,\nu}$ in (A.28) and (A.31), respectively, the scalar $\gamma_k^{n,\nu}$

in (A.23) can be now bounded by a linear combination of $\{|t_j^{n,\nu}|\}_{j=1}^K$. In particular, let f_k^n be a $1 \times K$ row vector whose entries are defined as,

$$[f_k^n]_k = \frac{(S_{\max,0}^n)^2}{(d_{\min,k}^n)^2(c_{\min,k}^n + S_{\max,0}^n)^2} + \frac{\sum_{j=1,\, j\neq k}^K \alpha_{jk}^n S_{\max,j}^n}{(\sum_{j=1,\, j\neq k}^K \alpha_{jk}^n S_{\max,j}^n + \eta_k^n) c_{\min,k}^n d_{\min,k}^n} + \frac{S_{\max,0}^n}{d_{\min,k}^n(c_{\min,k}^n + S_{\max,0}^n)} \left(\frac{1}{d_{\min,k}^n} + \frac{1}{c_{\min,k}^n} + \frac{S_{\max,k}^n}{c_{\min,k}^n(\alpha_{0k}^n c_{\min,k}^n + S_{\max,k}^n)} \right), \tag{A.32}$$

$$[f_k^n]_i = \frac{(S_{\max,k}^n)^2 d_{\min,k}^n + 2S_{\max,k}^n c_{\min,k}^n(\alpha_{0k}^n c_{\min,k}^n + S_{\max,k}^n)}{d_{\min,k}^n (c_{\min,k}^n)^2 (\alpha_{0k}^n c_{\min,k}^n + S_{\max,k}^n)^2} \alpha_{ik}^n + \frac{2S_{\max,0}^n S_{\max,k}^n}{c_{\min,k}^n(\alpha_{0k}^n c_{\min,k}^n + S_{\max,k}^n) d_{\min,k}^n (c_{\min,k}^n + S_{\max,0}^n)} \alpha_{ik}^n, \qquad i \neq k, \tag{A.33}$$

and let

$$\underline{t}^{n,\nu} = [|t_1^{n,\nu}|, \cdots, |t_K^{n,\nu}|]^T. \tag{A.34}$$

Using (A.32) and (A.33), $\gamma_k^{n,\nu}$ can be now bounded by

$$\gamma_k^{n,\nu} \le f_k^n \underline{t}^{n,\nu}. \tag{A.35}$$

Hence, the vector γ_k^ν can be element-wise bounded by the product of an $N \times NK$ block-diagonal matrix, F_k and a $KN \times 1$ vector whose entries are $|t_j^{n,\nu}|, n = 1, \ldots, N, k = 1, \ldots, K$. In particular, we define

$$F_k \triangleq \begin{bmatrix} f_k^1 & 0 & \cdots & 0 \\ 0 & f_k^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f_k^N \end{bmatrix}, \tag{A.36}$$

and write

$$\gamma_k^\nu \le F_k t^\nu, \tag{A.37}$$

where t^ν is defined as

$$t^\nu \triangleq \begin{bmatrix} \underline{t}^{1,\nu} \\ \vdots \\ \underline{t}^{N,\nu} \end{bmatrix}. \tag{A.38}$$

Substituting from (A.37) into (A.18), we obtain

$$\|r^\nu\| \le (1 - \|\Upsilon^\nu\|_2)^{-1} \Big\| \sum_{k=1}^{K} F_k \Big\|_2 \|t^\nu\|. \tag{A.39}$$

Now using (A.11), we have

$$\|r^\nu\| \le (1 - \|\Upsilon^\nu\|_2)^{-1} \Big\| \sum_{k=1}^{K} F_k \Big\|_2 \begin{bmatrix} \|A^{-1}B\|_2 & \|A^{-1}\beta\| \end{bmatrix} \begin{bmatrix} \|t^{\nu-1}\| \\ \|r^{\nu-1}\| \end{bmatrix}. \tag{A.40}$$

Writing (A.39) along with (A.11) in a vector form yields

$$\begin{bmatrix} \|t^\nu\| \\ \|r^\nu\| \end{bmatrix} \le \begin{bmatrix} 1 & 0 \\ 0 & (1 - \|\Upsilon^\nu\|_2)^{-1} \Big\| \sum_{k=1}^{K} F_k \Big\|_2 \end{bmatrix} \begin{bmatrix} \|A^{-1}B\|_2 & \|A^{-1}\beta\| \\ \|A^{-1}B\|_2 & \|A^{-1}\beta\| \end{bmatrix} \begin{bmatrix} \|t^{\nu-1}\| \\ \|r^{\nu-1}\| \end{bmatrix}, \tag{A.41}$$

where the inequality is to be interpreted element-wise. A sufficient condition for convergence is to have

$$\left\| \begin{bmatrix} 1 & 0 \\ 0 & (1 - \|\Upsilon^\nu\|_2)^{-1} \Big\| \sum_{k=1}^{K} F_k \Big\|_2 \end{bmatrix} \begin{bmatrix} \|A^{-1}B\|_2 & \|A^{-1}\beta\| \\ \|A^{-1}B\|_2 & \|A^{-1}\beta\| \end{bmatrix} \right\|_2 < 1. \tag{A.42}$$

In Appendix A.3, we show that the condition in (A.42) is equivalent to the condition that

$$\Big(1 + (1 - \|\Upsilon^\nu\|_2)^{-2}\Big\|\sum_{k=1}^{K} F_k\Big\|_2^2\Big)\big(\|A^{-1}B\|_2 + \|A^{-1}\beta\|\big) < 1. \tag{A.43}$$

Now, $\|\Upsilon^\nu\|_2$ is the only iteration-dependent entry in (A.43). Observe that the left hand side of (A.43) is a monotone increasing function of $\|\Upsilon^\nu\|_2$. Hence, for (A.43) to hold, it is sufficient to have

$$\|\Upsilon^\nu\|_2 \leq \tau, \tag{A.44}$$

where τ is an iteration-independent constant, that satisfies

$$\Big(1 + (1 - \tau)^{-2}\Big\|\sum_{k=1}^{K} F_k\Big\|_2^2\Big)\big(\|A^{-1}B\|_2 + \|A^{-1}\beta\|\big) < 1 \tag{A.45}$$

We now consider the diagonal matrix Υ^ν; cf. (A.16). The spectral norm of this matrix is given by the maximum absolute value of its diagonal entries. Hence, in order to satisfy (A.44), we must have

$$\max_n\left(1 - \sum_{k=1}^{K}\left(\frac{\alpha_{0k}^n}{(c_k^{n,\nu} + s_0^{n,*})(c_k^{n,\nu} + s_0^{n,*} + r^{n,\nu})} - \frac{\alpha_{0k}^n}{(d_k^{n,\nu} + s_0^{n,*})(d_k^{n,\nu} + s_0^{n,*} + r^{n,\nu})}\right)\right) \leq \tau, \tag{A.46}$$

$$\min_n\left(1 - \sum_{k=1}^{K}\left(\frac{\alpha_{0k}^n}{(c_k^{n,\nu} + s_0^{n,*})(c_k^{n,\nu} + s_0^{n,*} + r^{n,\nu})} - \frac{\alpha_{0k}^n}{(d_k^{n,\nu} + s_0^{n,*})(d_k^{n,\nu} + s_0^{n,*} + r^{n,\nu})}\right)\right) \geq -\tau. \tag{A.47}$$

We begin by considering the condition in (A.47). This condition can be written as

$$\max_n \sum_{k=1}^{K}\left(\frac{\alpha_{0k}^n}{(c_k^{n,\nu} + s_0^{n,*})(c_k^{n,\nu} + s_0^{n,*} + r^{n,\nu})} - \frac{\alpha_{0k}^n}{(d_k^{n,\nu} + s_0^{n,*})(d_k^{n,\nu} + s_0^{n,*} + r^{n,\nu})}\right) \leq \tau + 1. \tag{A.48}$$

Let χ_1 denote the term on the left hand side of (A.48). We first note that each term in the summand is a monotonically decreasing function of $r^{n,\nu}$. Since $s_0^{n,*} + r^{n,\nu} =$

$s_0^{n,\nu} \geq 0$, χ_1 can be bounded as follows.

$$\begin{aligned}
\chi_1 &\leq \max_n \sum_{k=1}^{K} \left(\frac{\alpha_{0k}^n}{(c_k^{n,\nu} + s_0^{n,*})c_k^{n,\nu}} - \frac{\alpha_{0k}^n}{(d_k^{n,\nu} + s_0^{n,*})d_k^{n,\nu}} \right) \\
&= \max_n \sum_{k=1}^{K} \alpha_{0k}^n \frac{(d_k^{n,\nu})^2 - (c_k^{n,\nu})^2 + s_0^{n,*}(d_k^{n,\nu} - c_k^{n,\nu})}{(c_k^{n,\nu} + s_0^{n,*})c_k^{n,\nu}(d_k^{n,\nu} + s_0^{n,*})d_k^{n,\nu}} \\
&= \max_n \sum_{k=1}^{K} \frac{s_k^{n,\nu}(2c_k^{n,\nu} + s_0^{n,*} + \frac{s_k^{n,\nu}}{\alpha_{0k}^n})}{(c_k^{n,\nu} + s_0^{n,*})c_k^{n,\nu}(c_k^{n,\nu} + s_0^{n,*} + \frac{s_k^{n,\nu}}{\alpha_{0k}^n})(c_k^{n,\nu} + \frac{s_k^{n,\nu}}{\alpha_{0k}^n})}.
\end{aligned} \tag{A.49}$$

One can check that each term in the summand in (A.49) is a monotonically decreasing function of $s_k^{n,\nu}$. Hence, we have

$$\chi_1 \leq \max_n \sum_{k=1}^{K} \frac{S_{\max,k}^n(2c_k^{n,\nu} + s_0^{n,*} + \frac{S_{\max,k}^n}{\alpha_{0k}^n})}{(c_k^{n,\nu} + s_0^{n,*})c_k^{n,\nu}(c_k^{n,\nu} + s_0^{n,*} + \frac{S_{\max,k}^n}{\alpha_{0k}^n})(c_k^{n,\nu} + \frac{S_{\max,k}^n}{\alpha_{0k}^n})}. \tag{A.50}$$

Similarly, each term in the summand in (A.50) is a monotonically decreasing function of $c_k^{n,\nu}$. Hence,

$$\chi_1 \leq \max_n \sum_{k=1}^{K} \frac{S_{\max,k}^n(2c_{\min,k}^n + s_0^{n,*} + \frac{S_{\max,k}^n}{\alpha_{0k}^n})}{(c_{\min,k}^n + s_0^{n,*})c_{\min,k}^n(c_{\min,k}^n + s_0^{n,*} + \frac{S_{\max,k}^n}{\alpha_{0k}^n})(c_{\min,k}^n + \frac{S_{\max,k}^n}{\alpha_{0k}^n})}. \tag{A.51}$$

Finally, one can check that each term in the summand in (A.51) is a monotonically decreasing function of $s_0^{n,*}$. Therefore, we can write

$$\chi_1 \leq \max_n \sum_{k=1}^{K} \frac{S_{\max,k}^n(2c_{\min,k}^n + \frac{S_{\max,k}^n}{\alpha_{0k}^n})}{(c_{\min,k}^n)^2(c_{\min,k}^n + \frac{S_{\max,k}^n}{\alpha_{0k}^n})^2}.$$

Therefore, a sufficient condition for (A.47) to be satisfied is

$$\max_n \sum_{k=1}^{K} \frac{S_{\max,k}^n(2c_{\min,k}^n + \frac{S_{\max,k}^n}{\alpha_{0k}^n})}{(c_{\min,k}^n)^2(c_{\min,k}^n + \frac{S_{\max,k}^n}{\alpha_{0k}^n})^2} \leq \tau + 1. \tag{A.52}$$

We now proceed to provide a sufficient condition for (A.46) to be satisfied at all iterations. This condition can be written as

$$\chi_2 = \min_n \sum_{k=1}^{K} \left(\frac{\alpha_{0k}^n}{(c_k^{n,\nu} + s_0^{n,*})(c_k^{n,\nu} + s_0^{n,*} + r^{n,\nu})} - \frac{\alpha_{0k}^n}{(d_k^{n,\nu} + s_0^{n,*})(d_k^{n,\nu} + s_0^{n,*} + r^{n,\nu})} \right) \geq 1 - \tau. \tag{A.53}$$

Noting that each term in the summand is monotonically decreasing in $r^{n,\nu}$, we have

$$\begin{aligned}
\chi_2 &\geq \min_n \sum_{k=1}^{K} \left(\frac{\alpha_{0k}^n}{(c_k^{n,\nu} + s_0^{n,*})(c_k^{n,\nu} + S_{\max,0}^n)} - \frac{\alpha_{0k}^n}{(d_k^{n,\nu} + s_0^{n,*})(d_k^{n,\nu} + S_{\max,0}^n)} \right) \\
&= \min_n \sum_{k=1}^{K} \alpha_{0k}^n \frac{(c_k^{n,\nu} + s_0^{n,*})\left(\frac{s_k^{n,\nu}}{\alpha_{0k}^n}\right) + \frac{s_k^{n,\nu}}{\alpha_{0k}^n}(c_k^{n,\nu} + S_{\max,0}^n + \frac{s_k^{n,\nu}}{\alpha_{0k}^n})}{(c_k^{n,\nu} + s_0^{n,*})(c_k^{n,\nu} + S_{\max,0}^n)(d_k^{n,\nu} + s_0^{n,*})(d_k^{n,\nu} + S_{\max,0}^n)} \\
&= \min_n \sum_{k=1}^{K} \frac{s_k^{n,\nu}}{(c_k^{n,\nu} + S_{\max,0}^n)(d_k^{n,\nu} + S_{\max,0}^n)} \left(\frac{1}{c_k^{n,\nu} + s_0^{n,*}} + \frac{1}{d_k^{n,\nu} + s_0^{n,*}} \right) && \text{(A.54)} \\
&\geq \min_n \sum_{k=1}^{K} \frac{s_k^{n,\nu}}{(c_k^{n,\nu} + S_{\max,0}^n)(d_k^{n,\nu} + S_{\max,0}^n)} \left(\frac{1}{c_k^{n,\nu} + S_{\max,0}^n} + \frac{1}{d_k^{n,\nu} + S_{\max,0}^n} \right), && \text{(A.55)} \\
&\geq \min_n \sum_{k=1}^{K} \Bigg(\frac{(\alpha_{0k}^n)^3 s_k^{n,\nu}}{\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\right)^2} \\
&\qquad \frac{1}{\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + s_k^{n,\nu} + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\right)} \\
&\quad + \frac{(\alpha_{0k}^n)^3 s_k^{n,\nu}}{\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\right)\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + s_k^{n,\nu} + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\right)^2} \Bigg), && \text{(A.56)}
\end{aligned}$$

where (A.55) follows from observing that each term in the summand in (A.54) is monotonically decreasing in $s_0^{n,*}$. Since (A.56) is a monotone increasing function of

$s_k^{n,\nu}$, we can use the lower bound $\eta_k^n \leq s_k^{n,\nu}$ (cf. (A.74)) to write

$$\chi_2 \geq \min_n \sum_{k=1}^{K} \Bigg(\frac{(\alpha_{0k}^n)^3 \eta_k^n}{\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\right)^2} \frac{1}{\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + \eta_k^n + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\right)} + \frac{(\alpha_{0k}^n)^3 \eta_k^n}{\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\right)\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + \eta_k^n + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\right)^2} \Bigg). \tag{A.57}$$

Now, a sufficient condition for (A.53) to be satisfied is to have

$$\min_n \sum_{k=1}^{K} \Bigg(\frac{(\alpha_{0k}^n)^3 \eta_k^n}{\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\right)^2 \left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + \eta_k^n + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\right)} + \frac{(\alpha_{0k}^n)^3 \eta_k^n}{\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\right)\left(\sum_{j=1,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n + \eta_k^n + \alpha_{0k}^n S_{\max,0}^n + \sigma_k^n\right)^2} \Bigg) \geq 1 - \tau. \tag{A.58}$$

In summary, if conditions (A.45), (A.52), and (A.58) are simultaneously satisfied, the GIWFA iterations are guaranteed to converge to a unique Nash equilibrium point for the non-cooperative game (5.0.5). This completes the proof of Theorem 1.

A.2 A lower bound on $s_k^{n,\nu}$

Denote the interference level observed by User $k \in \mathcal{K}$ on the n-th tone at the ν-th iteration by $I_k^{n,\nu}$, where

$$I_k^{n,\nu} = \sum_{j=1}^{k-1} \alpha_{jk}^n s_j^{n,\nu} + \sum_{j=k+1}^{K} \alpha_{jk}^n s_j^{n,\nu-1} + \alpha_{0k}^n s_0^{n,\nu-1}. \tag{A.59}$$

Since

$$s_k^{n,\nu} \leq S_{\max,k}^n, \qquad \forall n \in \mathcal{N} \tag{A.60}$$

an upper bound on $I_k^{n,\nu}$ can be expressed as

$$I_k^{n,\nu} \leq I_{\max,k}^n = \sum_{j=0,\, j\neq k}^{K} \alpha_{jk}^n S_{\max,j}^n. \tag{A.61}$$

For every $k \in \mathcal{K}$, let the permutation $\pi_k(\cdot)$ be defined such that

$$\sigma_k^{\pi_k(1)} + I_{\max,k}^{\pi_k(1)} \leq \sigma_k^{\pi_k(2)} + I_{\max,k}^{\pi_k(2)} \leq \cdots \leq \sigma_k^{\pi_k(N)} + I_{\max,k}^{\pi_k(N)}. \tag{A.62}$$

Before we proceed with our analysis, we provide a brief discussion regarding the IWFA algorithm. If we denote the water-level by μ_k^ν. Now, at each iteration, one can categorize the N tones into three classes; tones on which User k allocates power $S_{\max,k}^n$, tones on which User k performs standard water-filling, and tones on which User k puts no power. While the power allocated by User k on the first class of tones is not affected by the increase in water-level, if that exceeds a certain level, the power allocated on the remaining tones can only increase if μ_k^ν increases. Furthermore, we note that the constraints in (A.60) serve to increase to the water-level. In other words, if the constraints in (A.60) were not enforced, the water-level would decrease in order to bring the power level in the respective tones up to the water-level. Since in this section we are considering a lower bound on $s_k^{n,\nu}$, a worst-case scenario would be to assume that none of the constraints in (A.60) is active. In this case we have

$$s_k^{\pi_k(n),\nu} = [\mu_k^\nu - (I_k^{\pi_k(n),\nu} + \sigma_k^{\pi_k(n)})]^+, \qquad \forall n \in \mathcal{N}, \tag{A.63}$$

where $[\cdot]^+$ denotes the projection onto the non-negative real line.

Assuming, for simplicity of exposition, that at the ν-th iteration the noise plus interference assumes distinct values on each tone, it is possible to identify N water-level intervals. In particular, the water-level within a certain interval would only cover a certain subset of tones. Let the number of tones covered by water at the ν-th iteration be m_k^ν and let these tones be denoted by $\hat{\pi}_k(1), \ldots, \hat{\pi}_k(m_k^\nu)$, where, unlike (A.62), $\hat{\pi}_k(\cdot)$ is an iteration-dependent permutation of tones such that

$$\sigma_k^{\hat{\pi}_k(1)} + I_k^{\hat{\pi}_k(1),\nu} \leq \sigma_k^{\hat{\pi}_k(2)} + I_k^{\hat{\pi}_k(2),\nu} \leq \cdots \leq \sigma_k^{\hat{\pi}_k(N)} + I_k^{\hat{\pi}_k(N),\nu}. \tag{A.64}$$

Our goal is to find a lower bound on m_k^ν, and to identify the tones that User $k \in \mathcal{K}$ is guaranteed to activate at every iteration of the GIWFA. For the tones $\hat{\pi}_k(1), \ldots, \hat{\pi}_k(m_k^\nu)$, the term inside the square brackets (A.63) is non-negative, and this term is strictly negative for all remaining tones. Using this notation, we can express the water level explicitly as

$$\mu_k^\nu = \frac{1}{m_k^\nu}\Big(P_k + \sum_{i=1}^{m_k^\nu}(I_k^{\hat{\pi}_k(i),\nu} + \sigma_k^{\hat{\pi}_k(i)})\Big). \tag{A.65}$$

Substituting from (A.65) into (A.63), and noting that the choice of m_k^ν is such the term inside the square brackets (A.63) is non-negative for all j for which

$$\pi_k(j) \in \{\hat{\pi}_k(1), \ldots, \hat{\pi}_k(m_k^\nu)\}. \tag{A.66}$$

$$s_k^{\pi_k(j),\nu} = \frac{1}{m_k^\nu}\Big(P_k + \sum_{i=1}^{m_k^\nu}(I_k^{\hat{\pi}_k(i),\nu} + \sigma_k^{\hat{\pi}_k(i)})\Big) - (I_k^{\pi_k(j),\nu} + \sigma_k^{\pi_k(j)}),$$

$$\forall j \text{ for which (A.66) holds,} \tag{A.67}$$

Observe that if for the $\pi_k(j)$-th tone (A.66) does not hold, then the definition of m_k^ν implies that $s_k^{\pi_k(j),\nu} = 0$, and this tone is not used by User k at the νth iteration, and hence is not in the set of interest.

Let $m_k \in \{1, \ldots, N\}$ be the desired lower bound on m_k^ν. Furthermore, let $\sigma_k^{(i)}$ denote the noise variance of User $k \in \mathcal{K}$ that satisfies $\sigma_k^{(i)} \leq \sigma_k^{(i+1)}$ for $i = 1, \ldots, N-1$. We will show that if m_k is defined to be the largest integer for which

$$(m_k - 1)(\sigma_k^{\pi_k(j)} + I_{\max,k}^{\pi_k(j)}) \leq P_k + \sum_{i=1}^{m_k-1} \sigma_k^{(i)}, \tag{A.68}$$

is satisfied for all $j \leq m_k$, then $m_k \leq m_k^\nu$, $\forall \nu$. Since m_k satisfies (A.68), then m_k also satisfies

$$(m_k - 1)(\sigma_k^{\pi_k(j)} + I_k^{\pi_k(j),\nu}) \leq P_k + \sum_{\substack{i=1 \\ \hat{\pi}_k(i) \neq \pi_k(j)}}^{m_k} (\sigma_k^{\hat{\pi}_k(i)} + I_k^{\hat{\pi}_k(i),\nu}), \tag{A.69}$$

where $\hat{\pi}_k(\cdot)$ is the permutation of tones defined in (A.64). This is because the right hand side of (A.69) is at least as great as the right hand side of (A.68) and the left hand side is less than or equal to the left hand side of (A.68).

Now, (A.69) is equivalent to writing

$$\frac{1}{m_k}\Big(P_k + \sum_{i=1}^{m_k}(I_k^{\hat{\pi}_k(i),\nu} + \sigma_k^{\hat{\pi}_k(i)})\Big) - (I_k^{\pi_k(j),\nu} + \sigma_k^{\pi_k(j)}) \geq 0. \tag{A.70}$$

We now compare (A.70) with (A.67). Since by definition, m_k^ν is the largest integer for which the right hand side of (A.67) is greater than or equal to zero, we conclude that m_k is less than or equal to m_k^ν. However, from (A.68), we note that the definition of m_k does not depend on the iterations. Hence, from (A.68), we know that the tones $\pi_k(1), \ldots, \pi_k(m_k)$ are going to be activated by User k in each iteration.

Using the fact that m_k is a lower bound on the number of tones that are going to be activated, we can write a lower bound on the water level, μ_k^ν at the ν-th iteration. In particular, using (A.65) and (A.66), it is easy to see that

$$\mu_k^\nu \geq \frac{1}{N}\big(P_k + \sum_{i=1}^{m_k}(\sigma_k^{\pi_k(i)} + I_k^{\pi_k(i),\nu})\big). \tag{A.71}$$

Now, substituting from (A.71) into (A.63), we have

$$s_k^{\pi_k(n),\nu} \geq \Big[\frac{1}{N}\big(P_k + \sum_{i=1}^{m_k}\sigma_k^{\pi_k(i)}\big) - \Big(1 - \frac{1}{N}\Big)I_k^{\pi_k(n),\nu} - \sigma_k^{\pi_k(n)}\Big]^+, \qquad \forall n \in \mathcal{N}, \tag{A.72}$$

$$\geq \Big[\frac{1}{N}\big(P_k + \sum_{i=1}^{m_k}\sigma_k^{\pi_k(i)}\big) - \Big(1 - \frac{1}{N}\Big)I_{\max,k}^{\pi_k(n)} - \sigma_k^{\pi_k(n)}\Big]^+, \qquad \forall n \in \mathcal{N}, \tag{A.73}$$

$$= \Big[\frac{1}{N}\big(P_k + \sum_{i=1}^{m_k}\sigma_k^{\pi_k(i)}\big) - \Big(1 - \frac{1}{N}\Big)\sum_{j=0,\, j\neq k}^{K}\alpha_{jk}^{\pi_k(n)}S_{\max,j}^{\pi_k(n)} - \sigma_k^{\pi_k(n)}\Big]^+, \qquad \forall n \in \mathcal{N}. \tag{A.74}$$

Finally, we define η_k^n as

$$\eta_k^n \triangleq \Big[\frac{1}{N}\big(P_k + \sum_{i=1}^{m_k}\sigma_k^{\pi_k(i)}\big) + \big(\frac{1}{N} - 1\big)\sum_{j=0,\, j\neq k}^{K}\alpha_{jk}^{n}S_{\max,j}^{n} - \sigma_k^{n}\Big]^+, \tag{A.75}$$

where m_k is the largest integer for which (A.68) is satisfied, and the tone permutations $\pi_k(\cdot)$ are defined in (A.62) for all $k \in \mathcal{K}$. Hence, from (A.74) we have that η_k^n is an iteration-independent lower bound on $s_k^{n,\nu}$.

A.3 Proving the equivalence of (A.42) and (A.43)

In order to show that the condition in (A.42) is equivalent to that in (A.43), we notice that the 2×2 matrix on the right hand side of (A.42) is rank 1. Let us denote this matrix by Z; i.e.,

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & (1 - \|\Upsilon^\nu\|_2)^{-1} \| \sum_{k=1}^{K} F_k \|_2 \end{bmatrix} \begin{bmatrix} \|A^{-1}B\|_2 & \|A^{-1}\beta\| \\ \|A^{-1}B\|_2 & \|A^{-1}\beta\| \end{bmatrix}. \tag{A.76}$$

The condition in (A.42) is equivalent to $\|ZZ^T\|_2 < 1$. However, because Z is rank 1, then ZZ^T is also rank 1, and we have

$$\|ZZ^T\|_2 = \mathrm{Tr}(ZZ^T) = \Big(1 + (1 - \|\Upsilon^\nu\|_2)^{-2} \| \sum_{k=1}^{K} F_k \|_2^2 \Big) (\|A^{-1}B\|_2^2 + \|A^{-1}\beta\|^2) < 1, \tag{A.77}$$

which is the condition given in (A.43).