CONTRIBUTIONS TO RESOURCE ALLOCATION IN COGNITIVE RADIO NETWORKS

A Dissertation

by

XU WANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Erchin Serpedin
Co-Chair of Committee,	Khalid A. Qaraqe
Committee Members,	Xi Zhang
	Radu Stoleru
Head of Department,	Miroslav M. Begovic

August 2017

Major Subject: Electrical Engineering

ABSTRACT

The continuous increase in the number of wireless devices and the huge demand for higher data rates have promoted the development of new wireless communications technologies with improved spectrum sharing features. Recently, the concept of cognitive radio (CR) has gained increased popularity for the efficient utilization of radio frequency (RF) spectrum. A CR is characterized as a communication system which is capable to learn the spectrum environment through sensing, and to adapt its signaling schemes for a better utilization of the radio frequency resources. Resource allocation, which involves scheduling of spectrum and power resources, represents a crucial problem for the performance of CR networks in terms of system throughput and bandwidth utilization.

In this dissertation, we investigate resource allocation problems in a CR network by exploring a variety of optimization techniques. Specifically, in the first part of the dissertation, our goal is to maximize the total throughput of secondary users (SUs) in an orthogonal frequency division multiple access (OFDMA) CR network. In addition, the power of SUs is controlled to keep the interference introduced to primary users (PUs) under certain limits, which gives rise to a non-convex mixed integer non-linear programming (MINLP) optimization problem. It is illustrated that the original non-convex MINLP formulation admits a special structure and the optimal solution can be achieved efficiently using any standard convex optimization method under a general and practical assumption.

In the second part of the dissertation, considering the imperfect sensing information, we study the joint spectrum sensing and resource allocation problem in a multi-channel-multi-user CR network. The average total throughput of SUs is maximized by jointly optimizing the sensing threshold and power allocation strategies. The problem is also formulated as a non-convex MINLP problem. By utilizing the continuous relaxation and

convex optimization tools, the dimension of the non-convex MINLP problem is significantly reduced, which helps to reformulate the optimization problem without resorting to integer variables. A newly-developed optimization technique, referred to as the monotonic optimization, is then employed to obtain an optimal solution. Furthermore, a practical low-complexity spectrum sensing and resource allocation algorithm is proposed to reduce the computational cost.

DEDICATION

This dissertation is dedicated to my beloved wife Anni, the new member of our family - my lovely son Hugh, my parents, and my grandma.

ACKNOWLEDGMENTS

First, I would like to express my sincere gratitude to my advisor Dr. Erchin Serpedin for his continuous support of my PhD study and research. His advices on both my research and my career have been priceless to me. This work would not have been possible without his help and guidance.

My thanks also go to my committee members Dr. Khalid A. Qaraqe, Dr. Xi Zhang and Dr. Radu Stoleru for their valuable time, insightful comments and suggestions. I also would like to thank Dr. Tie Liu, Dr. Yang Shen, and Dr. Krishna Narayanan. Interacting with them has benefited me a lot for my PhD study. I am also grateful to have the opportunity to collaborate with Dr. Muhammad Ismail and Dr. Sabit Ekin. Thanks for their excellent suggestions and ideas.

I am very grateful to meet my great friends during my PhD program, such as Dr. Ali Ekti, Dr. Bilal Wajid, Dr. Mustafa Alshawaqfeh, Mr. Celal Bilgi, Dr. Alphan Sahin, Miss Xuan Dang and many others. I will always cherish the time we were working together and all the fun we had.

Last but not the least, I would like to thank my wife and my parents. It is their support and encouragement that helped me overcome difficulties and obstacles throughout my studies and my life in general. Thanks for being with me all the time.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Erchin Serpedin (advisor), Khalid A. Qaraqe (co-advisor), and Xi Zhang of the Department of Electrical and Computer Engineering and Professor Radu Stoleru of the Department of Computer Science and Engineering.

All work for the dissertation was completed by the student, in collaboration with Professor Sabit Ekin of the Department of Electrical and Computer Engineering in Oklahoma State University.

Funding Sources

This work was supported by NSF under Grant Number 1547447. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NSF.

NOMENCLATURE

AO	Alternating Optimization
CR	Cognitive Radio
CSI	Channel State Information
FCC	Federal Communications Commission
FFT	Fast Fourier Transform
ISI	Inter-Symbol Interference
KKT	Karush-Kuhn-Tucker
LTE	Long Term Evolution
MDKP	Multi-Dimensional Knapsack Problem
MIMO	Multiple-Input Multiple-Output
MINLP	Mixed Integer Nonlinear Programming
MISO	Multiple-Input Single-Output
NBI	Narrowband Inteference
NP	Non-Deterministic Polynomial-Time
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PBS	Primary Base Station
PSD	Power Spectral Density
PU	Primary User
QoS	Quality of Service
RF	Radio Frequency

SBS	Secondary Base Station
SNR	Signal-to-Noise Ratio
SINR	Signal-to-Interference-and-Noise Ratio
SU	Secondary User

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

1. INTRODUCTION

1.1 Cognitive Radio

With the increasing demands for smart mobile devices and other bandwidth consuming wireless applications, radio frequency (RF) spectrum is becoming more and more crowded. Even though most of the available spectrum has already been licensed, recent reports, performed by many agencies such as Federal Communications Commission (FCC), have indicated that the RF spectrum is being used inefficiently due to the conventional static frequency allocation scheme [1, 2]. As a result, the concept of cognitive radio (CR) was proposed as a promising technology by exploring the opportunistic usage of the frequency bands that are not heavily occupied by the licensed primary users (PUs) [3]. Based on the definition adopted by FCC [4]: "*Cognitive radio: A radio or system that senses its operational electromagnetic environment and can dynamically and autonomously adjust its radio operating parameters to modify system operation, such as maximize throughput, mitigate interference, facilitate interoperability, access secondary markets.*", one can observe that CR represents a potential new approach to spectrum access and aims at providing a more efficient utilization of RF spectrum.

1.1.1 Cognitive Radio Paradigms

The paradigms assumed by CR networks can be generally classified into three broad categories: underlay, overlay and interweave, based on the willingness of secondary users (SUs) to access a channel licensed to the PU [5]. In particular, a PU represents a licensed user with the highest priority on a given frequency band. SUs transmit opportunistically on PUs' licensed bands with the only requirement of not causing harmful interference to PUs.

In the underlay paradigm, the simultaneous coexistence of PUs and SUs on the same frequency band is allowed as long as the interference to PUs does not exceed a given threshold. The interference threshold can be defined in terms of the spectral mask, which stands for the power spectral density (PSD) distribution of the interference across the frequency band, or with respect to an average power constraint [5].

In overlay CR networks, SUs overhear the transmission of PUs and gather the knowledge of the PUs' transmission sequence and coding scheme. Then, SUs use this knowledge to choose a transmission strategy that enhances the performance of its own transmission while causing an acceptable amount of interference to PUs. For instance, SUs can utilize the knowledge of PUs' transmission to eliminate the interference caused by PUs at the secondary receiver.

The interweave paradigm is based on the concept of opportunistic spectrum access. Specifically, SUs sense the licensed spectrum and find the portions of the spectrum that are left temporarily free by PUs. These free frequency bands are referred to as the spectrum holes, and are available for SUs' transmission. That is to say, SUs are allowed to access the channel only if the channel is detected to be vacant. In this dissertation, we will focus on the interweave CR paradigm, which represents also the original motivation for CR networks.

1.1.2 Multi-Channel-Multi-User Setup in Cognitive Radio Networks

In this dissertation, we consider an interweave CR network with multiple channels, and multiple PUs and SUs. In this way, the scenarios of single wideband channel CR network and the CR network with a single PU and/or single SU can be considered as special cases. An example of multi-channel-multi-user CR network is illustrated in Figure 1.1. In this model, the CR network carries out the spectrum sensing on the multi-channel spectrum to detect the spectrum holes. Based on the spectrum sensing information, multiple SUs try to access the spectrum holes. While doing so, SUs also need to monitor the channels they occupy and vacate them whenever PUs become active on these channels. This procedure repeats in time since the spectrum occupancy information may change as time evolves.

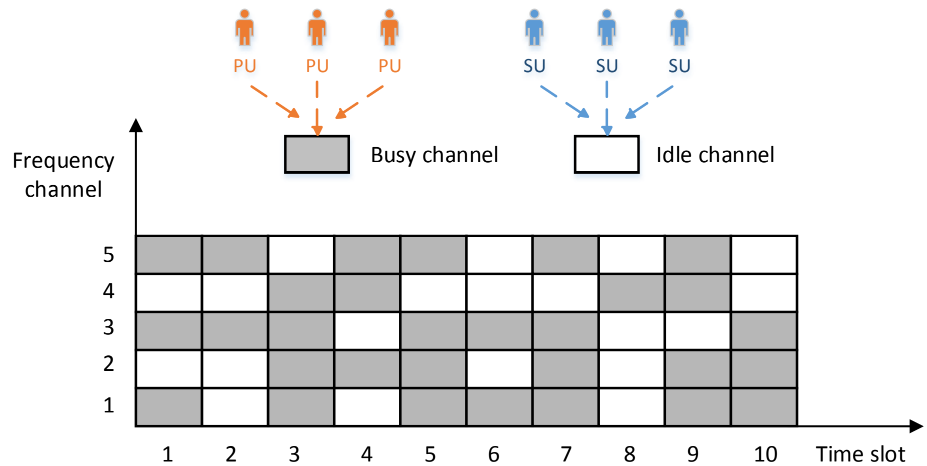


Figure 1.1: Interweave multiple-channel-multiple-user CR model

In CR networks, the orthogonal frequency division multiple access (OFDMA) signaling scheme, which represents an orthogonal frequency division multiplexing (OFDM)-based technology, is typically employed to support multiple access and spectrum allocation for multiple users.

OFDM represents a multicarrier modulation scheme that splits the data stream into several lower rate streams and transmit these streams in parallel on different carriers. Since the duration of each symbol is long, the OFDM signaling scheme suffers less from the multipath echoes. Remaining inter-symbol interference (ISI) can be further removed by implementing a cyclic prefix into the OFDM symbols. Moreover, the channel equalization process is significantly simplified since OFDM can be regarded as a sequence of slowly modulated narrowband signals rather than a rapidly modulated wideband signal. Other advantages of OFDM signaling scheme include the robustness against narrowband interferences (NBIs), very efficient implementation using fast Fourier transform (FFT), and the high spectral efficiency achieved in combination with the multiple-input multiple-output (MIMO) technique [6]. For these reasons, OFDM has been adopted as the physical layer signaling scheme by a number of current and potential future communications standards including IEEE 802.11 WLAN [7], IEEE 802.16 WMAN [8], IEEE 802.22 WRAN [9], and 3GPP Long Term Evolution (LTE) [10].

OFDM perhaps also presents the most suitable signaling scheme to meet all the requirements of a CR network at the physical level. The most important element in CR networks is the ability of measuring and learning the conditions of RF spectrum. OFDM presents intrinsic spectrum sensing abilities due to the inherent FFT operation employed for signal conversion from the time domain to the frequency domain. In this way, all the points in the time-frequency grid for the operating channels can be sensed and measured without any extra hardware components of spectrum sensing. After the CR network senses the spectrum and identifies the spectrum holes, the next step is to allocate the bandwidth to SUs and prevent a congested radio spectrum, i.e., the radio spectrum occupied by licensed PUs to be interfered by SUs. This procedure can be easily carried out by turning off the subcarriers that are used already by PUs, and this is easily implementable due to the very nature of OFDM signaling [6].

1.2 Overview of Resource Allocation

As discussed earlier, the spectrum sensing process resumes to finding the spectrum holes that are not utilized by PUs and to determine SUs which are available to transmit on these idle spectrum channels. In OFDMA CR networks, the problems to intelligently allocate these channels to SUs, and to adjust the transmission power of SUs to guarantee the QoS of PUs and to maximize the total profit of SUs, represent a paramount task for the sake of efficient utilization and scheduling of CR networks. These problems are referred to as the resource allocation problems, and have drawn great interest in the literature of CR networks. In this subsection, the state-of-the-art works on resource allocation problems in OFDMA CR networks are presented. In addition, our contributions to resource allocation are highlighted.

1.2.1 Related Work on Resource Allocation

Due to its significance in managing radio resources and resolving the users' contention for resources, resource allocation problems in OFDMA CR networks have been investigated intensively in literature. In the case of a single SU, the resource allocation problems were studied in [11–13]. Particularly, the authors in [11] studied the resource allocation problem by maximizing the throughput of a SU while limiting the interference introduced to PUs. The scenario of one primary transmitter-receiver and one secondary transmitter-receiver pairs wes considered. A modified water-filling algorithm was proposed to obtain the optimal solution. In addition, two suboptimal loading algorithms that are less computationally complex were also developed. Reference [12] employed a similar problem setup compared to [11] with an additional transmit power constraint. The problem was formulated as a non-convex optimization problem with respect to the power allocated to each OFDM subchannel. A low computational complexity algorithm was proposed to obtain a suboptimal solution. In [13], the resource allocation problem for a single SU was originally expressed as non-linear integer programming problem and then was re-formulated as a 0-1 multi-dimensional knapsack problem (MDKP). An efficient greedy max-min algorithm was then proposed to solve the corresponding problem.

The more practical multiuser set-up in OFDMA CR networks was investigated in [14–16]. In these works, the multiuser RA problem involves two tasks: i) the user assignment task, i.e., to assign the subchannels of OFDM spectrum to SUs; and ii) the power allocation task, i.e., to allocate power to SUs such that a certain objective is achieved. Particularly, the user assignment variables can only take binary values with 1 indicating the subchannel is occupied and 0 otherwise. Reference [14] proposed an efficient algorithm to maximize the total throughput of SUs. Instead of jointly optimizing the user assignment and power allocation tasks, the proposed suboptimal algorithm in [14] first performs the subchannel assignment task and then finds the optimal power allocation for a given SU layout. The authors in [15] assumed a time-sharing property among SUs, which allows multiple SUs to share a common subchannel. The energy efficiency of the CR network, which is defined as the ratio between the total capacity of SUs and the total power consumption, was investigated. A fast interior-point method was developed to jointly optimize the user assignment and power allocation tasks. To obtain a binary user assignment, the time-sharing variables were simply rounded, and the corresponding power levels were derived based on the rounded result. Reference [16] adopted a similar algorithm in [14] to perform the user assignment and power allocation tasks separately. The heterogenous users' rate requirements as well as the spectrum sensing errors were also taken into account by this study.

1.2.2 Contributions to Resource Allocation

The RA algorithms in these works suffer from two main drawbacks: i) no optimal solution is obtained, and all the algorithms proposed in these works are suboptimal in general; and ii) these algorithms sometimes lead to infeasible solutions. For example, in [15], the rounded user assignment variables might not represent a feasible solution for the corresponding power allocation. In this dissertation, a simplified problem formulation is proposed compared to [14–16]. We show that under a general and practical assumption, the proposed formulation can be always solved optimally by standard convex optimization techniques. Specifically, the subchannel layout for SUs can be directly obtained using the signal-to-interference-and-noise ratio (SINR) information of SUs and the corresponding optimal power allocation can be calculated via the proposed subgradient method.

1.3 Overview of Joint Spectrum Sensing and Resource Allocation

Spectrum sensing is of significant importance in CR networks since that's how SUs detect and access the available spectrum. A summary of the main spectrum sensing methods in CR networks is depicted in Table 1.1. The interested reader can refer to [17] and references therein for more details about these sensing methods as well as their advantages, disadvantages and implementation challenges. Apparently, a well-designed spectrum sensing scheme can accurately find the spectrum holes and improve the performance of secondary network. In the field of energy detection based sensing methods, several works studied the spectrum sensing problem by selecting the sensing parameters to maximize the throughput of secondary network under the constraint that PUs are sufficiently protected [18–20]. In particular, the total throughput of secondary network is expressed as the multiplication of the false alarm rate and the achievable throughput, with the false alarm rate expressed as a function of sensing parameters and the achievable throughput as fixed values. However, as discussed earlier, the achievable throughput of secondary network also depends on the resource allocation strategy, i.e., on how to allocate the channels and power to SUs. The references [11, 14, 21] focused on maximizing the throughput of secondary network assuming the perfect spectrum sensing information, i.e., no sensing errors are considered in these works. However, perfect spectrum sensing is extremely difficult to acquire in practical wireless networks. Considering the false alarm and misdetection rates as fixed values, the authors in [16, 22] studied the resource allocation problem in CR networks by maximizing the total throughput of SUs.

In a nutshell, both the selection of sensing parameters and the physical-layer resource allocation strategy have an impact on the performance of the secondary network. The aforementioned references either optimize the spectrum sensing scheme under a pre-defined

SU resource layout or design a resource allocation strategy while fixing the sensing parameters. As an extension and generalization of our resource allocation work introduced in Subsubsection 1.2.2, a joint spectrum sensing and resource allocation optimization problem is proposed for a more practical scheduling of CR network.

Spectrum sensing method	Description
Energy Detection	Also known as radiometry of periodogram, the most common method of spectrum sensing. Detects the primary signal by comparing the output of the energy detector with a sensing threshold.
Waveform-based sensing	Also known as coherent sensing, applicable to systems with known signal patterns. Correlates the received signal with a known pattern of itself and compare the decision metric with a threshold.
Cyclostationarity-based sensing	Detects the presence of PUs' transmission using the cyclic correlation function, which is a cyclostationary feature of the received signal.
Radio identification based sensing	Extracts several features from the received signal and identifies the presence of known transmission technologies via classification methods.
Matched filter	Implements the matched filter to detect a known transmitted signal. Requires perfect knowledge of the PUs' signaling scheme.

Table 1.1: Category of spectrum sensing methods

1.3.1 Related Work on Joint Optimization

In the CR literature, the state-of-the-art joint spectrum sensing and resource allocation problems in CR networks can be roughly categorized into two broad classes: those corresponding to single channels and those associated with multiple channels. In the scenario of a single channel band, the authors in [23] studied the joint optimization of sensing time and power allocation. The goal was to maximize the average throughput of the secondary network and minimize the outage probability of SUs' transmission under the average power constraint of SUs and the average interference threshold introduced to PUs. In [24], the sensing parameters and the transmission power of SUs were optimized to minimize the total power consumption. The cooperative sensing scheme was adopted to jointly consider the sensing data from different locations to prevent the potential sensing issues due to the multipath fading/shadowing and hidden primary user. In terms of the multi-channel setup, which is also the concentration of this dissertation, reference [25] formulated the joint optimization problem for one PU and one SU in OFDM-based CR networks. The problem was further generalized to the case of a non-cooperative power allocation game for two SUs. In [26], Fan *et al.* considered the joint optimization problem for multiple SUs in a multi-channel CR network. The time-sharing property among SUs was assumed in [26], which allows multiple SUs to share a common channel. In both [25] and [26], the alternating optimization (AO) method was employed to maximize the average throughput of SUs by alternatively optimizing the sensing parameters and power variables while setting constant the other variables.

1.3.2 Contributions to Joint Optimization

In the multi-channel setup, both [25] and [26] utilized the AO method to iteratively optimize the sensing parameters and allocate the transmission power using a random initial guess. However, the AO method can be easily trapped in a local minima/maxima near the starting point [27]. Thus, the convergence of the solutions in [25] and [26] heavily depend on the initial point selected and the stability of the proposed AO algorithms is questionable. Moreover, for a multi-channel-multi-user setup, the time-sharing property is not practical in some cases such as the OFDM scheme, in which an IFFT/FFT pair coupled with a cyclic prefix are used for modulation and de-modulation. Since the FFT-size is fixed in advance, no further subdivision of each frequency subchannel is possible [28].

In this dissertation, we address the joint spectrum sensing and resource allocation problem in a downlink multi-channel-multi-user CR network. Our goal is to maximize the average throughput of the secondary network subject to a peak power constraint of SUs and an average interference constraint at the PUs. Furthermore, we assume that each channel in the frequency domain can be occupied by at most one SU, which leads to a non-convex mixed integer nonlinear programming (MINLP) optimization problem. We first show that under a general and practical assumption, the non-convex MINLP problem can be reduced to a formulation which does not resort to any integer variable. It is further illustrated that the reduced formulation represents the canonical form of a monotonic optimization and can be solved globally using the polyblock outer approximation algorithm. As an alternative method for the joint optimization problem, a computationally efficient suboptimal algorithm, which consists of a list of single-variable optimization problems, is proposed based on a modified version of the original non-convex MINLP formulation.

1.4 Outline

The rest of this dissertation is structured as follows:

- As one of the two mainly-used optimization methods in this dissertation, some preliminaries of convex optimization are presented in Section 2. Since convex optimization has been investigated intensively for decades and most of the techniques are well known to researchers in a variety of areas, only the specific techniques employed in this dissertation are briefly described.

- As another optimization method used throughout the entire dissertation, a conceptual and graphical introduction on monotonic optimization is provided in Section 3. Since the concept of monotonic optimization was only reported recently, a detailed illustration is presented.

- In Section 4, the optimal resource allocation algorithm for downlink OFDMA CR networks is presented.

- In Section 5, by considering the imperfect spectrum sensing information, the joint spectrum sensing and resource allocation problem is studied in a multi-channel-multi-user CR network.

- Conclusions and some future research directions are drawn in Section 6.

2. PRELIMINARIES ON CONVEX OPTIMIZATION

In recent years, convex optimization has become a popular computational tool in the general area of engineering, thanks to its ability to solve large-scaled practical engineering problems efficiently and reliably. Nowadays, new applications of convex optimization are continuously being reported in nearly every area of engineering and science including communications, signal processing, estimation and detection theory, networking, control, information theory, computer science, operations research, machine learning, bioinformatics, economics and statistics. For detailed examples of convex optimization applications, the reader is referred to [29–33] and references therein.

In this section, some preliminaries of convex optimization techniques that are utilized in this dissertation are briefly presented. In particular, some basic concepts and theorems of convex optimization are introduced in Subsection 2.1. The relevant convex optimization techniques in the dual domain are given in Subsection 2.2. In Subsection 2.3, we briefly discuss how convex optimization can be applied to solve convex MINLP optimization problems.

2.1 Basic Definitions and Theorems

In this subsection, most of the definitions and theorems about convex optimization are presented based on the materials reported in [29, 34] without resorting to proofs.

Definition 1. *A set $\mathcal{S} \subset \mathbb{R}^n$ is a convex set if it contains the entire line segment joining any pair of its points, i.e.,*

$$\mathbf{x}, \mathbf{y} \in \mathcal{S} \Rightarrow \lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in \mathcal{S},\ \forall 0 \leq \lambda \leq 1.$$

Definition 2. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex on a convex set $\mathcal{S}$ if*

$$f(\theta \mathbf{x} + (1-\theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y}),\ \forall \mathbf{x}, \mathbf{y} \in \mathcal{S}, 0 \leq \theta \leq 1.$$

A function f is said to be concave if $-f$ is a convex function.

Theorem 1. *A more general way to determine the convexity of a function is by means of the Hessian matrix. Specifically, a continuous, twice differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is convex on a convex set $\mathcal{S}$ if and only if its Hessian matrix of second partial derivatives is positive semidefinite. Similarly, a function f is said to be concave if its Hessian matrix is negative semidefinite.*

Definition 3. *The general form of a convex optimization problem is defined as*

$$\min_{\mathbf{x} \in \mathcal{S}} f(\mathbf{x}),$$

for a convex function $f : \mathbb{R}^n \to \mathbb{R}$ and a convex set $\mathcal{S} \in \mathbb{R}^n$. Alternatively, the standard form of convex optimization problems can be expressed as

$$\begin{aligned} &\min\ f(\mathbf{x}) \\ &\text{s.t. } g_i(\mathbf{x}) \leq 0,\ i = 1, \cdots, m \\ &\qquad h_j(\mathbf{x}) = 0,\ j = 1, \cdots, p \end{aligned} \tag{2.1}$$

where "s.t." stands for "subject to", $f, g_i : \mathbb{R}^n \to \mathbb{R}$ are convex functions and $h_j : \mathbb{R}^n \to \mathbb{R}$ are affine functions.

Definition 4. *If $f : \mathbb{R}^n \to \mathbb{R}$, then the perspective function of f is the function $g : \mathbb{R}^{n+1} \to \mathbb{R}$ defined as*

$$g(\mathbf{x}, t) = t f(\mathbf{x}/t),$$

for $t > 0$. The perspective operator preserves convexity. That is to say, if f is a convex function, so is its perspective function g. Similarly, if f is concave, then so is g.

Theorem 2. *If $f_1, f_2, \cdots, f_n : \mathbb{R}^n \to \mathbb{R}$ are convex functions and $\omega_1, \omega_2, \cdots, \omega_n \geq 0$, then $f : \mathbb{R}^n \to \mathbb{R}$, defined as*

$$f = \omega_1 f_1 + \omega_2 f_2 + \cdots + \omega_n f_n$$

is also convex. Similarly, a nonnegative weighted sum of concave functions is concave.

2.2 Convex Optimization in Dual Domain

The topics of duality and Karush-Kuhn-Tucker (KKT) conditions are our focus of this subsection. In this subsection, most of the results are borrowed from the convex optimization tutorial paper [35]. First, we consider a standard optimization formulation in (2.1). However, for now, we do not assume the convexity of (2.1). The Lagrange function $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ is defined as

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^{p} \mu_j h_j(\mathbf{x}),$$

where vectors $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_m]^T$ and $\boldsymbol{\mu} = [\mu_1, \cdots, \mu_p]^T$ are referred to as the Lagrange multipliers or dual variables. The Lagrange dual problem is defined as

$$\begin{aligned} &\max\, d(\boldsymbol{\lambda}, \boldsymbol{\mu}) \\ &\text{s.t. } \boldsymbol{\lambda} \geq \mathbf{0}, \end{aligned}$$

with $d(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$. In contrast, the original problem (2.1) is referred to as the primal problem compared to the dual problem. It can be stated that regardless of the convexity for the primal problem, the Lagrange dual problem is a convex optimization problem. In addition, it presents an optimal value d^* less than the optimal value p^* of the primal problem, i.e., $d^* \leq p^*$, a condition which is termed as the weak duality. Thanks to this unique property, one can always approximate a lower bound of a non-convex optimization problem by analyzing and addressing the dual problem.

Moreover, if we assume the convexity of the primal problem (2.1), a more powerful theorem can be employed to achieve a necessary and sufficient condition for optimality. Specifically, under the assumption of a certain regularity condition and the differentiability of functions f, g_i and h_j, the following KKT conditions hold:

Theorem 3. *If* $\tilde{\mathrm{x}}$, $\tilde{\boldsymbol{\lambda}}$, *and* $\tilde{\boldsymbol{\mu}}$ *are any points satisfying the KKT conditions*

$$\begin{aligned} g_i(\tilde{\mathbf{x}}) &\leq 0,\ i = 1, \cdots, m \\ h_i(\tilde{\mathbf{x}}) &= 0,\ j = 1, \cdots, p \\ \tilde{\lambda}_i &\geq 0,\ i = 1, \cdots, m \\ \tilde{\lambda}_i g_i(\tilde{\mathbf{x}}) &= 0,\ i = 1, \cdots, m \\ \nabla f(\tilde{\mathbf{x}}) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla g_i(\tilde{\mathbf{x}}) + \sum_{j=1}^{p} \tilde{\mu}_j \nabla h_j(\tilde{\mathbf{x}}) &= \mathbf{0}, \end{aligned} \tag{2.2}$$

where ∇ *stands for the gradient, then* $\tilde{\mathrm{x}}, \tilde{\boldsymbol{\lambda}}$, *and* $\tilde{\boldsymbol{\mu}}$ *are primal and dual optimal solution, respectively. In addition, the strong duality is achieved, i.e.,* $d^* = p^*$.

The regularity condition can be met in various ways. In general, most convex optimization problems satisfy the Slater's condition, which means that there exists a point $\mathbf{x}$ such that $g_i(\mathbf{x}) < 0$ and $h_j(\mathbf{x}) = 0$. Additionally, if g_i and h_j are all affine functions, then the regularity condition is automatically met.

KKT conditions play a significant role in convex optimization. In some special cases, an analytical solution can be directly derived from (2.2). Moreover, some numerical convex optimization algorithms, such as the interior-point method with the barrier funciton, can be interpreted equivalently as tackling the KKT conditions.

2.3 Relaxation for Convex MINLP Optimization

As discussed earlier, the proposed resource allocation and joint optimization problems are originally both formulated as MINLP problems. Therefore, a brief introduction to MINLP problems and their optimization algorithms are illustrated in this subsection.

Mathematically, MINLP expresses the optimization problem using both integer and fractional variables. Furthermore, the objective function and/or its feasible region are described by nonlinear functions [36]. In some instances, when the integrality variable is continuously relaxed, the remaining objective function and feasible set are convex and such kind of MINLP is referred to as the convex MINLP problem. Even though the convex MINLP is NP-hard, there are exact algorithms which terminate in a finite number of steps with a guaranteed optimal solution (if there is one). Some examples of optimal convex MINLP algorithms include the branch-and-bound method [37], the generalized Benders decomposition [38], the outer approximation method [39], the branch-and-cut method [40], and the extended cutting-plan method [41]. However, most of these methods suffer from a relatively high computational complexity and may not be suitable for implementation in CR wireless networks. The main reason is that the resource allocation in CR networks utilizes the instantaneous channel state information (CSI) and spectrum information and must be carried out frequently to cope with the changes in these data.

In the literature of wireless communications, a typical way for tackling convex MINLP problems is to first perform the continuous relaxation by omitting the integrality variable [15]. Then, the remaining problem admits the form of a nonlinear convex optimization and can be addressed efficiently using any standard convex optimization algorithm such as the interior-point method. In this way, the solution obtained for integer variables is mostly likely fractional and it does not represent a feasible solution. To deal with this issue, some rounding policy is proposed to transform the fractional solutions into integer values. Another method is to first perform integer optimization based on the constraints related to integer variables [14, 16]. Thus, an integer feasible but suboptimal solution is obtained. Then, the optimization problem is calculated again only with respect to the remaining fractional variables.

However, the aforementioned two procedures suffer from some drawbacks, mainly due to the potential infeasible solution. For the former case, a rounding policy has to be designed to guarantee the feasibility of the constraints associated with the integer variables. Sometimes this kind of rounding policy might be very complicated or even impossible to carry out. With regard to the later case, even though the feasibility of the integer variables is ensured, the resulting solution for the fractional variables may not satisfy all the related constraints.

In this dissertation, we still adopt the continuous relaxation method to tackle the convex MINLP problem. By exploring the special structure of the formulated problem, we alternatively show that the relaxed problem actually achieves an integer solution for the integer variable. In this way, we do not need to go through the two procedures mentioned above and avoid the potential infeasibility issue.

3. PRELIMINARIES ON MONOTONIC OPTIMIZATION

Optimization techniques have been widely used in a variety of disciplines in engineering and the solutions of most optimization problems rely heavily on the convexity of the problem formulation. However, many practical engineering problems are non-convex in their original formulation and even cannot be transformed into an equivalent convex form by any means. An interesting observation is that most non-convex engineering problems, especially in the areas of communications, networking and information theory, show the monotonic or hidden monotonic properties. One simple example lies in the field of information theory, where the Shannon capacity monotonically increases with regard to the bandwidth and the signal-to-noise ratio (SNR) [42]. The analysis of these monotonic properties has led to a number of concepts and theorems, which may enable the researchers to reduce the difficulty level in obtaining the global optimal solution of the problem, and this is the fundamental principle of monotonic optimization.

Most mathematical concepts and theories for monotonic optimization have been established since 2000 by a series of papers written by Tuy *et al.* [43–46]. Only very recently the concept of monotonic optimization was introduced to the ares of communications and networking. Qian *et al.* made the first attempt in [47], in which the weighted throughput maximization through power control is achieved globally using a monotonic-based algorithm named as MAPLE. The hidden monotonic property was discovered from the original formulation of the power control problem. A number of follow-up works [19, 48–51] then were reported by exploring the monotonic property or the hidden monotonic property inside the proposed non-convex problem formulation. The applications of monotonic property encompass a variety of topics including power scheduling, multiple-input single-output (MISO) and MIMO channel optimization, and CR management.

In this section, some preliminaries about monotonic optimization are introduced. Most of the material in this section is summarized based on two tutorial-style papers [43] and [52], where the first paper is presented mainly from a mathematical point of view while the second paper focuses on applications of monotonic optimization in communications and networking.

3.1 Basic Definitions and Theorems

Some basic concepts and theorems for monotonic optimization are illustrated in this subsection without proof. The reader is referred to [43, 52] for more details of the proofs.

Definition 5 (Increasing function)**.** *A function* $f : \mathbb{R}_+^n \to \mathbb{R}$ *is increasing if* $f(\mathbf{x}) \le f(\mathbf{y})$ *when*[1] $\mathbf{x} \le \mathbf{y}$. *A function* f *is decreasing if* $-f$ *is increasing.*

Definition 6 (Normal set)**.** *A set* $\mathcal{G} \subset \mathbb{R}_+^n$ *is normal if* $\mathbf{x} \in \mathcal{G} \Rightarrow [\mathbf{0}, \mathbf{x}] \subset \mathcal{G}$.

Definition 7 (Conormal set)**.** *A set* $\mathcal{H} \subset \mathbb{R}_+^n$ *is conormal in* $[\mathbf{0}, \mathbf{b}]$ *if* $\mathbf{x} \in \mathcal{H} \Rightarrow [\mathbf{x}, \mathbf{b}] \subset \mathcal{H}$.

Definition 8 (Upper boundary)**.** *A point* $\bar{\mathbf{x}}$ *in a closed normal set* $\mathcal{G}$ *is called an upper boundary point of* $\mathcal{G}$ *if there does not exist any other point* $\mathbf{x}$ *in* $\mathcal{G}$ *such that* $\mathbf{x} > \bar{\mathbf{x}}$. *The set of all upper boundary points of* $\mathcal{G}$ *is called the upper boundary of* $\mathcal{G}$ *and is denoted as* $\partial^+\mathcal{G}$.

For a better understanding of these concepts, an example illustrating the normal set, the conormal set, and the upper boundary of the normal set is shown in Figure 3.1. In the rectangular region $[\mathbf{0}, \mathbf{b}]$, set $\mathcal{G}$, denoted as the striped area, is clearly a normal set, set $\mathcal{H}$, represented as the shadowed area, is a conormal set, and the upper boundary of $\mathcal{G}$ is marked as the blue line in the figure.

Next, we present the definition of monotonic optimization in its standard form:

[1]For two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we say $\mathbf{x} \le \mathbf{y}$ if $x_i \le y_i,\ i = 1, \cdots, n$.

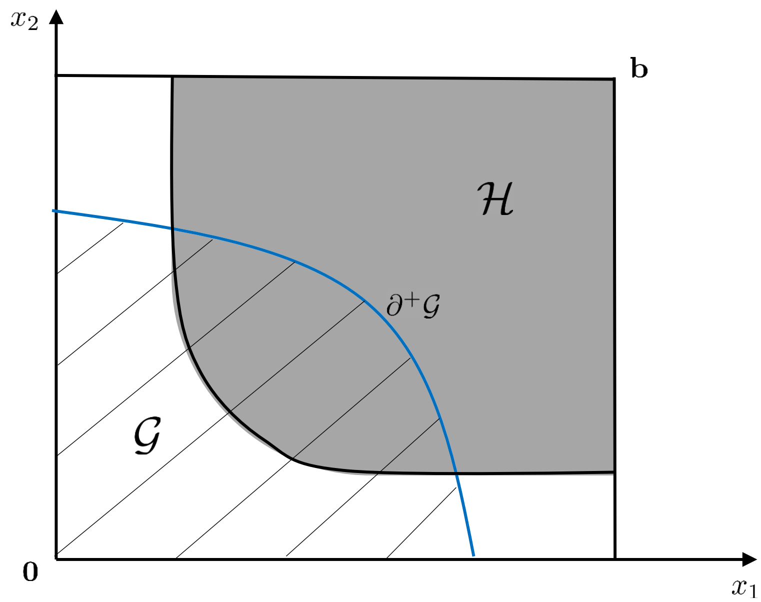


Figure 3.1: Example of normal set, conormal set and upper boundary.

Definition 9 (Standard form of monotonic optimization)**.** *Monotonic optimization is referred to as a problem with the following structure:*

$$\begin{aligned} &\max \ f(\mathbf{x}) \\ &\text{s.t. } \mathbf{x} \in \mathcal{G} \cap \mathcal{H}, \end{aligned} \tag{3.1}$$

where $f(\mathbf{x}) : \mathbb{R}_+^n \to \mathbb{R}$ *is an increasing function,* $\mathcal{G} \subset [\mathbf{0}, \mathbf{b}] \subset \mathbb{R}_+^n$ *is a normal set with nonempty interior, and* $\mathcal{H}$ *is a closed conormal set in* $[\mathbf{0}, \mathbf{b}]$.

Since monotonic optimization is mainly employed in non-convex optimization problems in which a monotonic or a hidden monotonic property can be explored, we first give readers an intuitive idea about how monotonic property can be utilized to alleviate the difficulty to obtain a global solution for non-convex optimization problems.

In general, finding the global solution for a non-convex optimization problem involves searching every point in the feasible domain. However, if the objective function $f(\mathbf{x})$:

$\mathbb{R}^n \to \mathbb{R}$ in a maximization problem is monotonically increasing, once a feasible point $\mathbf{x}$ is known, all the points $\bar{\mathbf{x}}$ satisfying $\bar{\mathbf{x}} < \mathbf{x}$ can be simply discarded since all these points lead to a smaller objective value compared to $\mathbf{x}$. Moreover, if a function $g(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ representing a constraint such that $g(\mathbf{x}) \leq 0$ is a monotonically increasing function, once an infeasible point $\mathbf{x}$ is found such that $g(\mathbf{x}) > 0$, then all the points satisfying $\bar{\mathbf{x}} > \mathbf{x}$ are infeasible at all due to the monotonic property of the constraint. In this way, the monotonic property inside the non-convex optimization problem successfully reduces the size of feasible region, and thus simplifies the searching procedure for a global optimal solution.

3.2 Polyblock Outer Approximation Algorithm

The commonly-used algorithm for solving monotonic optimization problems, referred to as the polyblock outer approximation algorithm, is illustrated in this subsection. We first introduce two important concepts, namely the concept of polyblock and the concept of proper vertices.

Definition 10 (Polyblock)**.** *A set $\mathcal{P} \subset \mathbb{R}^n_+$ is called a polyblock if it is a union of a finite number of rectangular boxes, i.e., $\mathcal{P} = [\mathbf{0}, \mathbf{b}_1] \cup [\mathbf{0}, \mathbf{b}_2] \cup \cdots \cup [\mathbf{0}, \mathbf{b}_m]$. Let $\mathcal{T} = [\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_m]$, then $\mathcal{T}$ is defined as the vertex set of the polyblock. Obviously, a polyblock is a normal set.*

Definition 11 (Improper vertex)**.** *Let $\mathcal{T}$ be the vertex set of a polyblock, a vertex $\mathbf{v} \in \mathcal{T}$ is proper if there does not exist $\mathbf{v}' \neq \mathbf{v}$ and $\mathbf{v}' \in \mathcal{T}$ such that $\mathbf{v}' > \mathbf{v}$. A vertex is improper if it is not proper. Removing improper vertices does not affect the shape of a polyblock.*

An example of polyblocks and improper vertices is shown in Figure 3.2. It can be seen that the poloblock is generated by taking the union of 3 rectangular boxes with vertex $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$, respectively. However, based on the definition of improper vertices, it can be readily observed that vertex $\mathbf{v}_3$ is actually an improper vertex and removing $\mathbf{v}_3$ does not change the shape of polyblock.

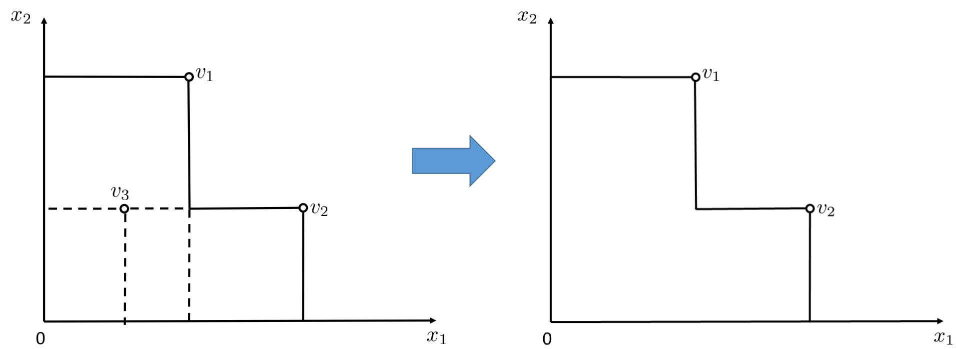
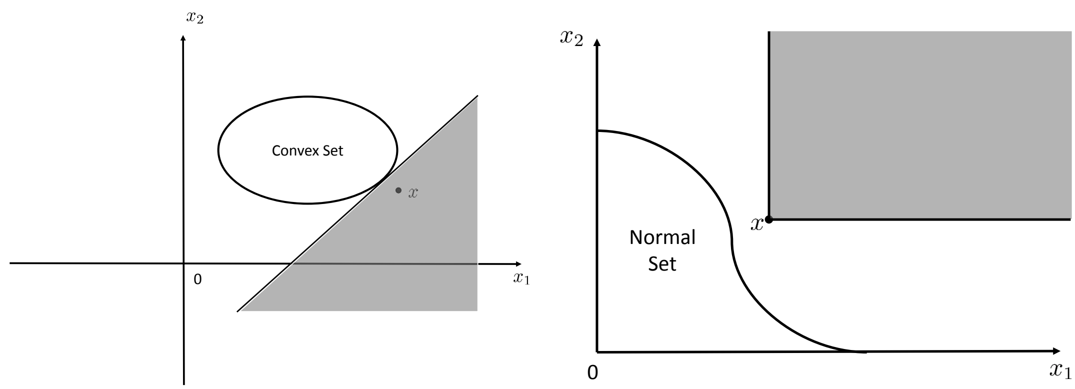


Figure 3.2: Illustration of polyblocks and improper vertices

For convex optimization, the supporting hyperplane theorem states that for any point $\mathbf{x}$ outside the feasible convex set, there is a hyperplane passing through a point on the boundary of the convex set that separates $\mathbf{x}$ from the feasible set. Consequently, the feasible set is enclosed in a halfspace and separated from the other halfspace. The supporting hyperplane theorem plays a fundamental role for the polyhedral outer approximation method [53], since it allows to address the convex optimization based on iteratively approximating the convex feasible set by a nested sequence of polyhedra. In each iteration, the new polyhedron is formed by taking the intersection with the new halfspace generated by the supporting hyperplane theorem. Additionally, for a concave maximization problem, the concave function always achieves its maximum over a bounded polyhedron on one of its vertices.

Similarly, the monotonic optimization problem can be also solved in a way analogous to the polyhedral outer approximation method. Particularly, any point $\mathbf{x}$ outside a normal set can be separated from this normal set by a cone defined as $\{\bar{\mathbf{x}}|\bar{\mathbf{x}} \geq \mathbf{x}\}$. As a result,



(a) (b)

Figure 3.3: (a) Separation pattern for convex optimization; (b) Separation pattern for monotonic optimization.

a normal set can be approximated by a nested sequence of polyblocks. For a clear illustration of the relationship and difference between convex set approximation and normal set approximation, the separation patterns for these two approximations are depicted in Figure 3.3. Moreover, the comparison between a polyhedron and a polyblock is illustrated in Figure 3.4.

The details of the polyblock outer approximation algorithm are omitted in this subsection. However, in Subsection 4.3, a step-by-step illustration will be given for an implementation of the polyblock outer approximation algorithm in CR networks.

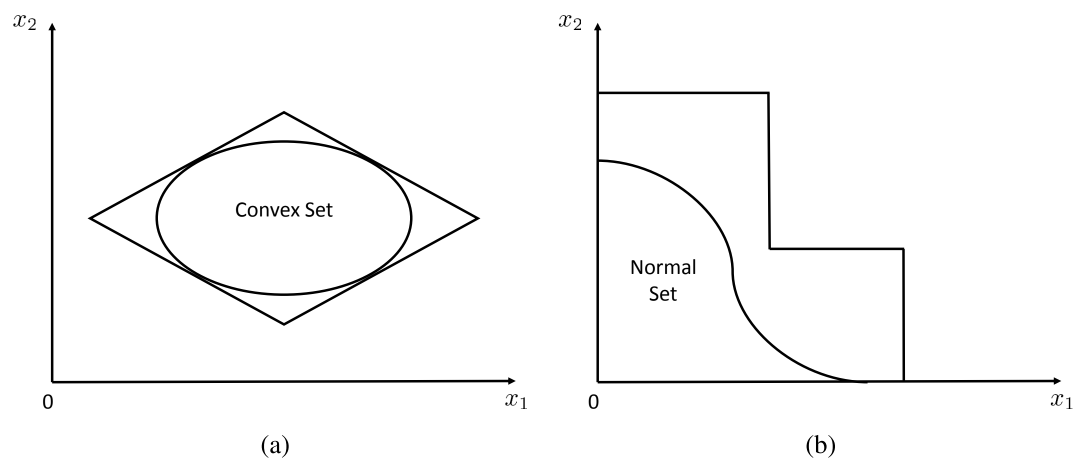


Figure 3.4: (a) A polyhedron enclosing a convex set; (b) A polyblock enclosing a normal set.

4. OPTIMAL RESOURCE ALLOCATION FOR DOWNLINK OFDMA COGNITIVE RADIO NETWORKS

4.1 Introduction

In this section, we study the downlink resource allocation problem in OFDMA CR networks assuming the perfect spectrum sensing knowledge [54]. Our goal is to maximize the aggregated throughput of SUs. In addition, the power of SUs is controlled to keep the interference introduced to PUs under certain limits, which gives rise to a non-convex MINLP optimization problem. It is illustrated that the non-convex MINLP formulation admits a special structure and the optimal solution can be always achieved using standard convex optimization techniques under a general and practical assumption. In particular, the subgradient method is adopted to address the problem in the dual domain. The effectiveness of the proposed algorithms is verified by simulations.

The rest of this section is structured as follows. The system model and problem formulation are presented in Subsection 4.2. By exploring the special structure of our problem, an optimal subchannel and power allocation algorithm is proposed in Subsection 4.3. Numerical results and discussions are given in Subsection 4.4. Finally, conclusions are drawn in Subsection 4.5.

4.2 System Model

Consider a downlink OFDMA CR interweave network with L licensed PUs coexisting with K SUs, as shown in Figure 4.1. In the RF spectrum domain, the CR model employed in [11, 14, 15] is considered. Particularly, as depicted in Figure 4.2, it is assumed that the lth PU's operating band ranges from f_l to $f_l + B_l$ and the remaining available spectrum has N total subchannels with equal bandwidth B. Since SUs can only access to subchannels that are detected to be vacant, the interference introduced from SUs to PUs is mainly due to out-of-band emissions, which are caused by power leakage in the sidelobes of OFDM signals [55]. Specifically, the interference per unit power introduced to the lth PU due to the allocation of the SU on the nth subchannel is given by [11]

$$I_{l,n}^{SP} = \int_{d_{l,n}-B_l/2}^{d_{l,n}+B_l/2} g_{l,n}^{SP} \phi(f) df, \tag{4.1}$$

where g_{ln}^{SP} represents the power gain from the secondary base station (SBS) to the lth PU on the nth subchannel, and $d_{l,n}$ denotes the spectrum distance between the centers of the nth subchannel and the lth PU band. Notation $\phi(f)$ represents the PSD of the OFDM signal and it is given by $\phi(f) = T\left(\frac{\sin \pi f T}{\pi f T}\right)^2$, where T stands for the symbol duration.

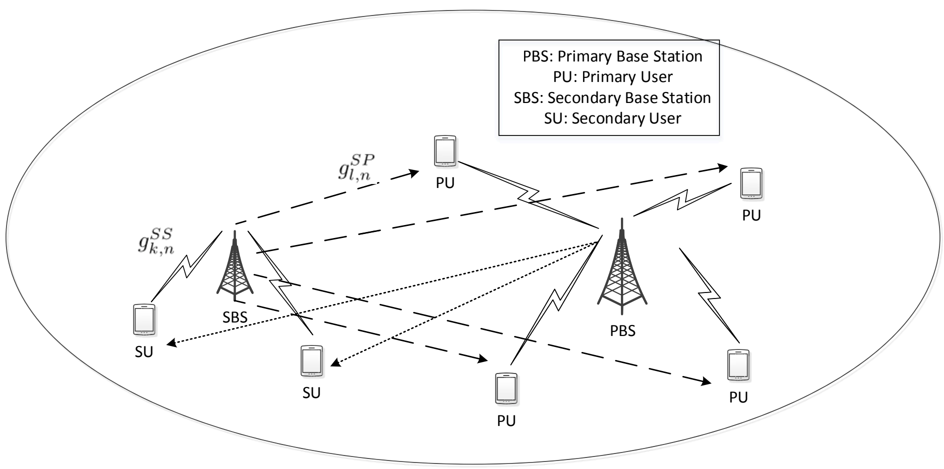


Figure 4.1: A downlink cognitive radio network.

The received throughput of the kth SU on the nth subchannel is expressed via Shannon throughput formula as follows [14]:

$$r_{k,n} = \log\left(1 + \frac{g_{k,n}^{SS} p_{k,n}}{\sigma^2 + I_k^{PS}}\right), \tag{4.2}$$

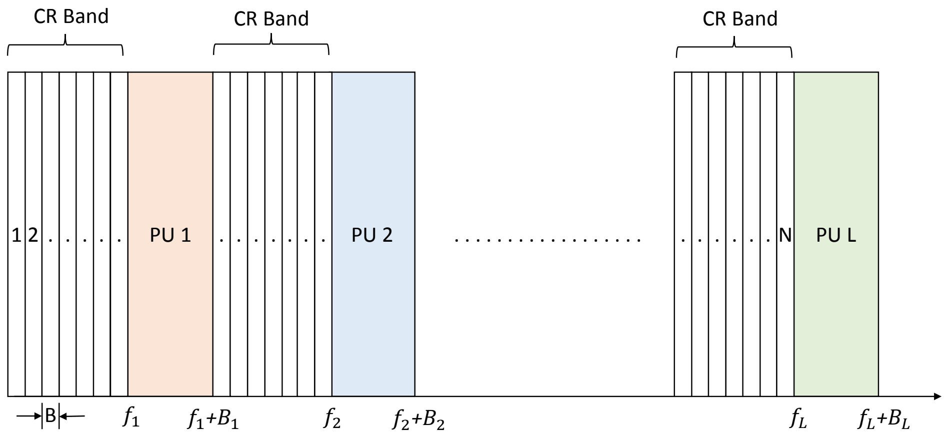


Figure 4.2: RF Spectrum sharing between PUs and SUs.

where $g_{k,n}^{SS}$ is the power gain from the SBS to the kth SU on the nth subchannel, $p_{k,n}$ stands for the transmitted power from SBS to the kth SU on the nth subchannel, and σ^2 denotes the noise power. Additionally, I_k^{PS} represents the interference caused by PUs on the kth SU, which can be measured as part of the noise by the receiver of the kth SU [56]. It is assumed that the power gains from the SBS to SUs ($g_{k,n}^{SS}$) and PUs ($g_{l,n}^{SP}$) are known by the SBS.

Towards this end, by denoting the SINR on the nth subchannel occupied by the kth SU with unit power as

$$s_{k,n} \doteq \frac{g_{k,n}^{SS}}{\sigma^2 + I_k^{PS}}, \tag{4.3}$$

the downlink resource allocation optimization problem in CR networks can be formalized as the problem (P_1) shown below:

$$\max_{x_{k,n},p_{k,n}} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n}\log(1+s_{k,n}p_{k,n}) \tag{P_1}$$

$$\text{s.t.} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n}p_{k,n} \le P_T \tag{C_1^1}$$

$$\sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n}p_{k,n}I_{l,n}^{SP} \le I_l^{\max}, \ \ l=1,\cdots,L, \tag{C_1^2}$$

$$\sum_{k=1}^{K} x_{k,n} \le 1, \ \ n=1,\cdots,N \tag{C_1^3}$$

$$p_{k,n} \ge 0, \ \ k=1,\cdots,K \ \ n=1,\cdots,N \tag{C_1^4}$$

$$x_{k,n} \in \{0,1\}, \ \ k=1,\cdots,K \ \ n=1,\cdots,N \tag{C_1^5}$$

where $r_{k,n} = \log(1+s_{k,n}p_{k,n})$, $x_{k,n}$ is a binary variable indicating whether the nth subchannel is occupied by the kth SU, P_T denotes the total transmit power by the SBS, $I_l^{\max}$ stands for the interference tolerance for the lth PU, constraint (C_1^3) requires that each subchannel can be only allocated to only one SU, and constraint (C_1^4) simply illustrates the nonnegativity property of power variables.

4.3 Optimal Subchannel and Power Allocation Algorithm

Based on the discussion in Subsection 2.3, the continuous relaxation of problem (P_1) is non-convex with respect to $x_{k,n}$ and $p_{k,n}$. Therefore, optimization problem (P_1) is a non-convex MINLP problem, which is NP-hard and computationally difficult to solve. To efficiently obtain the optimal solution for (P_1), we first introduce a new variable $y_{k,n}$, where $y_{k,n} = x_{k,n}p_{k,n}$, to replace the variable $p_{k,n}$ in (P_1). Thus, the resulting optimization problem is equivalent to (P_1) and it is termed as "P_1^E" shown below:

$$
\begin{aligned}
\max_{x_{k,n}, y_{k,n}} \quad & \sum_{k=1}^{K} \sum_{n=1}^{N} x_{k,n} \log(1 + \frac{s_{k,n} y_{k,n}}{x_{k,n}}) && (P_1^E) \\
\text{s.t.} \quad & \sum_{k=1}^{K} \sum_{n=1}^{N} y_{k,n} \le P_T \\
& \sum_{k=1}^{K} \sum_{n=1}^{N} y_{k,n} I_{l,n}^{SP} \le I_l^{\max}, \;\; l = 1, \cdots, L \\
& \sum_{k=1}^{K} x_{k,n} \le 1, \;\; n = 1, \cdots, N \\
& y_{k,n} \ge 0, \;\; k = 1, \cdots, K \;\; n = 1, \cdots, N \\
& x_{k,n} \in \{0, 1\}, \;\; k = 1, \cdots, K \;\; n = 1, \cdots, N
\end{aligned}
$$

It is proved in Appendix A that the continuous relaxation of the optimization problem (P_1^E), termed as (P_2) later, is a convex optimization problem. Therefore, problem (P_1^E) admits a formulation of convex MINLP. However, as illustrated in Subsection 2.3, convex MINLP problems are still NP-hard. To find the optimal solution of (P_1^E), the brute-force approach requires generating K^N permutations of binary variables $x_{k,n}$ and solving the power allocation problem for each permutation. Other convex MINLP algorithms, such as the branch-and-bound and the outer approximation method, can definitely reduce the searching time compared to the brute-force approach, but they are still computationally complex especially for the OFDM systems with a large number of subchannels. Therefore, we instead investigate the continuous relaxation of the problem by relaxing the binary constraint $x_{k,n} \in \{0, 1\}$ to $0 \le x_{k,n} \le 1$. This approach is also referred to as the time-sharing method in the literature of multicarrier wireless systems. The term "time-sharing" comes from the intuitive meaning of the binary variable. In multicarrier signaling scheme, relaxing the binary variable simply indicates that this carrier is occupied by a certain user in a fractional portion of time and the remaining time can also be shared by other users. The time-sharing approach was first presented in [57] and has been widely used in the multicarrier OFDM system to transform binary variables into continuous ones [58–61]. Thus, the following relaxed optimization problem (P_2) is formulated

$$\max_{x_{k,n}, y_{k,n}} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n} \log(1 + \frac{s_{k,n} y_{k,n}}{x_{k,n}}) \tag{P_2}$$

$$\text{s.t.} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} y_{k,n} \leq P_T \tag{C_2^1}$$

$$\sum_{k=1}^{K}\sum_{n=1}^{N} y_{k,n} I_{l,n}^{SP} \leq I_l^{\max}, \;\; l = 1, \cdots, L \tag{C_2^2}$$

$$\sum_{k=1}^{K} x_{k,n} \leq 1, \;\; n = 1, \cdots, N \tag{C_2^3}$$

$$y_{k,n} \geq 0, \;\; k = 1, \cdots, K \;\; n = 1, \cdots, N \tag{C_2^4}$$

$$x_{k,n} \geq 0, \;\; k = 1, \cdots, K \;\; n = 1, \cdots, N \tag{C_2^5}$$

Since (C_2^3) implicitly implies that $x_{k,n} \leq 1$, (C_2^5) is simply written as $x_{k,n} \geq 0$ instead of $0 \leq x_{k,n} \leq 1$.

In the remaining of this subsection, by making a general assumption, we show that problem (P_2) always achieve a binary optimal solution for the user assignment variable even though the variable $x_{k,n}$ can take any value from 0 to 1.

Since all the constraints of the convex problem (P_2) are affine, the regularity condition holds and (P_2) admits a zero duality gap [29]. Therefore, (P_2) can be solved in the dual domain using KKT conditions. The Lagrange function of (P_2) is expressed as

$$
\begin{aligned}
&L(x_{k,n}, y_{k,n}, \lambda, \mu_l, \eta_n) \\
=&x_{k,n} \log\left(1 + \frac{s_{k,n} y_{k,n}}{x_{k,n}}\right) + \lambda\left(P_T - \sum_{k=1}^{K}\sum_{n=1}^{N} y_{k,n}\right) + \sum_{l=1}^{L}\mu_l\left(I_l^{\max} - \sum_{k=1}^{K}\sum_{n=1}^{N} y_{k,n} I_{l,n}^{SP}\right) \\
&+ \sum_{n=1}^{N}\eta_n\left(1 - \sum_{k=1}^{K} x_{k,n}\right),
\end{aligned}
\tag{4.4}
$$

where constraints (C_2^4) and (C_2^5) are absorbed in the KKT conditions.

For simplicity, the logarithm functions are assumed to be natural logarithms. Due to the KKT conditions, it follows that

$$
\frac{\partial L}{\partial y_{k,n}} = \frac{s_{k,n} x_{k,n}}{x_{k,n} + s_{k,n} y_{k,n}} - \lambda - \sum_{l=1}^{L}\mu_l I_{l,n}^{SP} \begin{cases} = 0, \ y_{k,n} > 0 \\ < 0, \ y_{k,n} = 0 \end{cases}
\tag{4.5}
$$

which further leads to

$$
y_{k,n} = \left(\frac{1}{\lambda + \sum_{l=1}^{L}\mu_l I_{l,n}^{SP}} - \frac{1}{s_{k,n}}\right)^{+} x_{k,n},
\tag{4.6}
$$

with $(\cdot)^+$ is defined as $\max(\cdot, 0)$. In addition, taking the derivative of the Lagrange function with respect to $x_{k,n}$ leads to

$$
\frac{\partial L}{\partial x_{k,n}} = \log\left(1 + \frac{s_{k,n} y_{k,n}}{x_{k,n}}\right) - \frac{s_{k,n} y_{k,n}}{x_{k,n} + s_{k,n} y_{k,n}} - \eta_n
$$

For simplicity, we denote

$$\begin{aligned} J_{k,n} =& \log\left(1+\frac{s_{k,n}y_{k,n}}{x_{k,n}}\right) - \frac{s_{k,n}y_{k,n}}{x_{k,n}+s_{k,n}y_{k,n}} \\ =& \log\left[1+s_{k,n}\left(\frac{1}{\lambda+\sum\limits_{l=1}^{L}\mu_l I_{l,n}^{SP}} - \frac{1}{s_{k,n}}\right)^{+}\right] - \frac{s_{k,n}\left(\frac{1}{\lambda+\sum\limits_{l=1}^{L}\mu_l I_{l,n}^{SP}} - \frac{1}{s_{k,n}}\right)^{+}}{1+s_{k,n}\left(\frac{1}{\lambda+\sum\limits_{l=1}^{L}\mu_l I_{l,n}^{SP}} - \frac{1}{s_{k,n}}\right)^{+}}, \end{aligned} \tag{4.7}$$

where the second equality follows from (5.16). From the KKT conditions, it turns out that

$$J_{k,n} - \eta_n = \begin{cases} = 0, \ x_{k,n} > 0 \\ < 0, \ x_{k,n} = 0. \end{cases} \tag{4.8}$$

Moreover, due to the complementary slackness of KKT conditions, it follows that

$$\lambda\left(P_T - \sum_{k=1}^{K}\sum_{n=1}^{N} y_{k,n}\right) = 0 \tag{4.9}$$

$$\mu_l\left(I_l^{\max} - \sum_{k=1}^{K}\sum_{n=1}^{N} y_{k,n} I_{l,n}^{SP}\right) = 0, \quad l = 1, \cdots, L \tag{4.10}$$

$$\eta_n\left(1 - \sum_{k=1}^{k} x_{k,n}\right) = 0, \quad n = 1, \cdots, N \tag{4.11}$$

with $\lambda, \mu_l, \eta_n \geq 0$.

Based on the aforementioned equations, the following result can be established.

Lemma 1. *For* $n = 1, \cdots, N$, $\eta_n = \max_k J_{k,n}$.

Proof. See Appendix B.1. □

Based on Lemma 1 and the KKT conditions (4.5) - (4.11), the main results of our work are summarized by the following theorems.

Theorem 4. *If* $\max_k J_{k,n}$ *is achieved by only one SU, i.e.,* $J_{k_n^*,n} > J_{k,n},\ k \neq k_n^*$*, then the* n*th subchannel is assigned to the* k_n^**th SU only, i.e.,*

$$x_{k_n^*,n} = 1|_{k_n^*=\arg\max J_{k,n}},\ \forall n \tag{4.12}$$

If $\max_k J_{k,n}$ *is tied by multiple SUs, to optimize problem* (P_2)*, the* n*th subchannel must be shared by these SUs.*

Proof. See Appendix B.2 □

Theorem 5. *Under the assumption that for any given subchannel* n*,* $s_{k,n}$*'s are all distinct for* $k = 1, \cdots, K$*, the problem* (P_2) *can always achieve a binary optimal solution for the user assignment variable* $x_{k,n}$*. In particular, each subchannel is solely assigned to the SU with the largest* $s_{k,n}$*, i.e.,*

$$x_{k_n^*,n} = 1|_{k_n^*=\arg\max s_{k,n}} \text{ and } x_{k,n} = 0|_{k\neq k_n^*},\ \forall n \tag{4.13}$$

Proof. See Appendix B.3. □

In general, SUs are randomly located around the SBS and the probability that two SUs have exactly the same SINR per unit power is almost zero. Therefore, in practical scenarios, the optimization problem (P_2) can always achieve a binary optimal solution for the user assignment variables.

Towards this end, it can be claimed that under the assumption in Theorem 5, problem (P_2) can always achieve a binary optimal solution in (4.13). In this way, the dimension of problem (P_2) can be significantly reduced from $2KN$ to N. In particular, plugging (4.13) into problem (P_2) yields

$$\max_{y_{k_n^*,n}} \sum_{n=1}^{N} \log(1 + s_{k_n^*,n} y_{k_n^*,n}) \tag{P_3}$$

$$\text{s.t.} \sum_{n=1}^{N} y_{k_n^*,n} \le P_T \tag{C_3^1}$$

$$\sum_{n=1}^{N} y_{k_n^*,n} I_{l,n}^{SP} \le I_l^{\max}, \;\; l = 1, \cdots, L \tag{C_3^2}$$

$$y_{k_n^*,n} \ge 0, \;\; n = 1, \cdots, N, \tag{C_3^3}$$

which is termed as problem (P_3) in this subsection.

It can be observed that problem (P_3) represents a classic waterfilling problem without constraint (C_3^2). The existence of constraint (C_3^2) adds another level of difficulty to the problem, which does not admit an analytical solution. However, problem (P_3) still can be solved by any standard convex optimization numerical method, such as the interior-point method and the subgradient method. Since all the derivations aforementioned are discussed in the dual domain, we use the subgradient method to iteratively update the dual variables and recover the primal variable $y_{k_n^*,n}$. In particular, suppose the dual variables associated with constraints (C_3^1) and (C_3^2) to be β and γ_l, respectively. They can be updated by

$$\beta^{i+1} = \left(\beta^i - \alpha^i (P_T - \sum_{n=1}^{N} y_{k_n^*,n})\right)^+ \tag{4.14}$$

$$\gamma_l^{i+1} = \left(\gamma_l^i - \alpha^i (I_l^{\max} - \sum_{n=1}^{N} y_{k_n^*,n} I_{l,n}^{SP})\right)^+, \tag{4.15}$$

where i is the iteration number and α_i is the step size in the ith iteration. A common choice for the step size in subgradient methods is referred to as “square summable but not summable", in which the step sizes satisfy

$$\alpha_i \geq 0, \quad \sum_{i=1}^{\infty} \alpha_i = \infty, \quad \sum_{i=1}^{\infty} \alpha_i^2 < \infty.$$

One typical example is $\alpha^i = a/(i+b)$ with some $a > 0$ and $b \geq 0$. Once the optimal dual variables are obtained, the corresponding optimal primal variable can be recovered by the following relationship:

$$y_{k_n^*,n} = \left(\frac{1}{\beta + \sum_{l=1}^{L} \gamma_l I_{l,n}^{SP}} - \frac{1}{s_{k_n^*,n}} \right)^+, \tag{4.16}$$

which is derived by applying the KKT conditions to problem (P_3). The reader can refer to [62] for more details about the subgradient method.

To summarize, the outline of our proposed subgradient method for solving problem (P_3), or equivalently problem (P_2), is presented below in Algorithm 1. Once the optimal solution $x_{k,n}^*$ and $y_{k,n}^*$ for problem (P_2) is obtained, we can claim it is also optimal for the original problem (P_1). This argument can be proved in two steps: first, since (P_2) is a relaxed version of (P_1), the optimal objective value of (P_2) provides an upper bound for (P_1); second, since $x_{k,n}^*$'s are binary, the optimal solution of (P_2) is a feasible solution of (P_1) and consequently achieves a lower bound for the optimal objective value in (P_1).

In terms of the computational complexity, Algorithm 1 requires $\mathcal{O}(KN)$ operations to find the optimal user assignment using (4.13). Suppose the subgradient method converges in Δ iterations, and in each iteration, it needs $\mathcal{O}(N)$ to update $y_{k_n^*,n}$ and $\mathcal{O}(L)$ to update μ_l. Thus, the overall complexity for Algorithm 1 is $\mathcal{O}(N\Delta)$ since the subgradient method generally needs much more than K iterations to converge.

Algorithm 1 Subgradient method for problem (P_2)

Require: SINRs of SUs in (4.3) are distinct.

Obtain the optimal SU assignment using (4.13). Assume the k_n^*th SU occupies the nth subchannel.
For each subchannel n, set $y_{k,n} = 0$ for $k \neq k_n^*$.
Initialize β and γ_l. Set $i = 0$.
while $i \leq$ maximum number of iterations **do**
  for $n = 1$ to N **do**
    Calculate $y_{k_n^*,n}$ using (4.16)
  end for
  Update β^{i+1} and γ_l^{i+1} using (4.14) and (4.15), respectively
  i = i + 1
end while

4.4 Simulation Results

For the simulation results presented in this subsection, we assume the numbers of PUs to be 4, i.e., $L = 4$. The whole OFDM access scheme presents 44 subchannels with the bandwidth of each subchannel equal to $B = 15$kHz. The four PUs occupy the subchannels 1-3, 10-14, 20-27 and 30-40, respectively. Thus, the number of available CR subchannels is $N = 20$. It is also assumed that the symbol duration $T = 66.7\mu$s and the noise power $\sigma^2 = 10^{-3}$W. The interference introduced to SUs, I_k^{PS}, is assumed to be uniformly distributed from 10^{-3}W to 10^{-2}W. In addition, the channels from the SBS to SUs, as well as the channels from the SBS to PUs, are assumed to be Rayleigh fading with an average channel power gain 10dB. Thus, $g_{k,n}^{SS}$ and $g_{l,n}^{SP}$ are exponentially distributed with mean 10dB. Since some parameters are randomly generated, an average result of 1,000 independent Monte-Carlo runs is carried out for simulating the total SU throughput.

In Figure 4.3, it is assumed that the number of SUs is 5, i.e., $K = 5$, the total SU throughput for (P_1) is performed with a power limit P_T ranging from 1W to 1.6W and an interference threshold $I_l^{\max}$ ranging from 4×10^{-12}W to 7×10^{-12}W. It can be observed

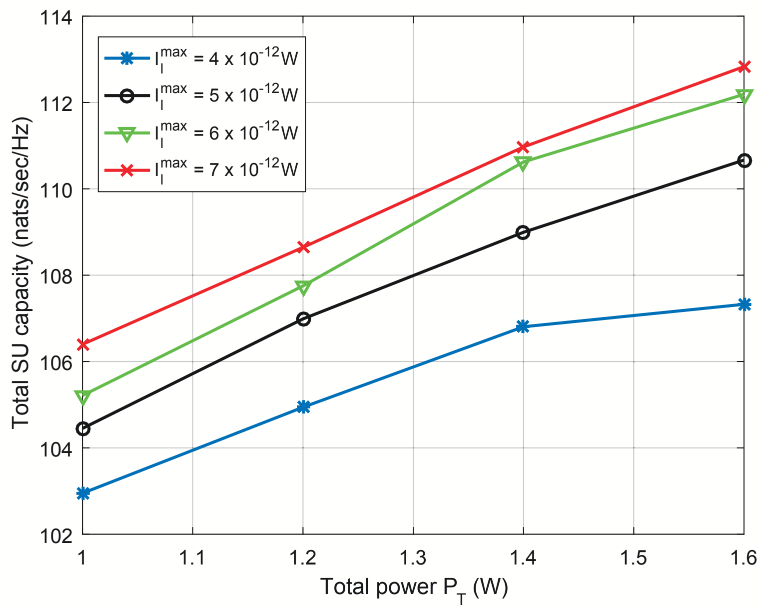


Figure 4.3: Total SU throughput as a function of P_T and $I_l^{\max}$ with $K = 5$.

that the total SU throughput increases with the increase of power limit and/or interference threshold.

In Figure 4.4 and 4.5, the interference threshold $I_l^{\max}$ is fixed to be 5×10^{-12}W. We alternatively assume that the number of SUs varies from 2 to 11. In this case, the total SU throughput for (P_1) is simulated in Figure 4.4 with the power limit P_T ranging from 1W to 1.6W. It can be seen that even with the same power limit and interference threshold, the CR system with more SUs can achieve a higher throughput. As expected, it is depicted in Figure 4.5 that the average SU throughput decreases as the number of SUs increases for a fixed power limit and interference threshold.

4.5 Conclusions

In this section, the resource allocation problem in a downlink CR network was investigated. Our goal was to maximize the total throughput of SUs subject to an interference constraint at PUs. The problem was formulated as a non-convex MINLP optimization

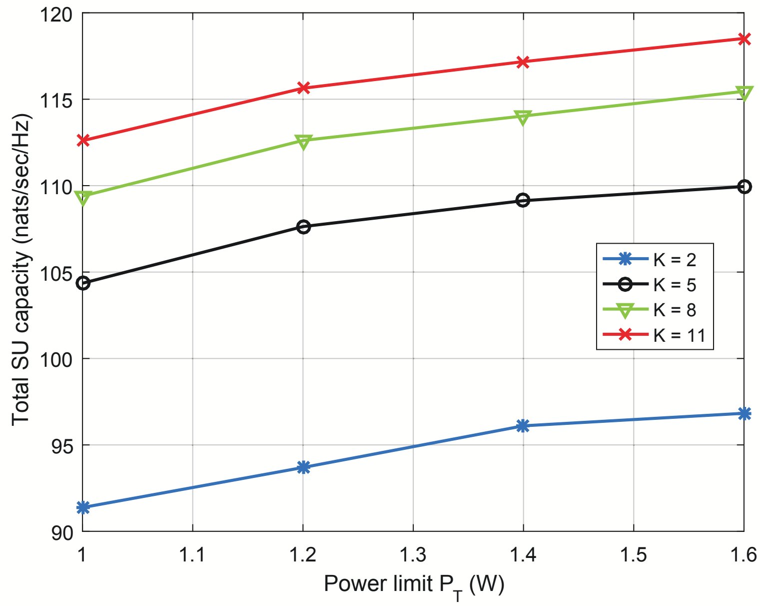


Figure 4.4: Total SU throughput as a function of P_T and K, with $I_l^{\max} = 5 \times 10^{-12}$W.

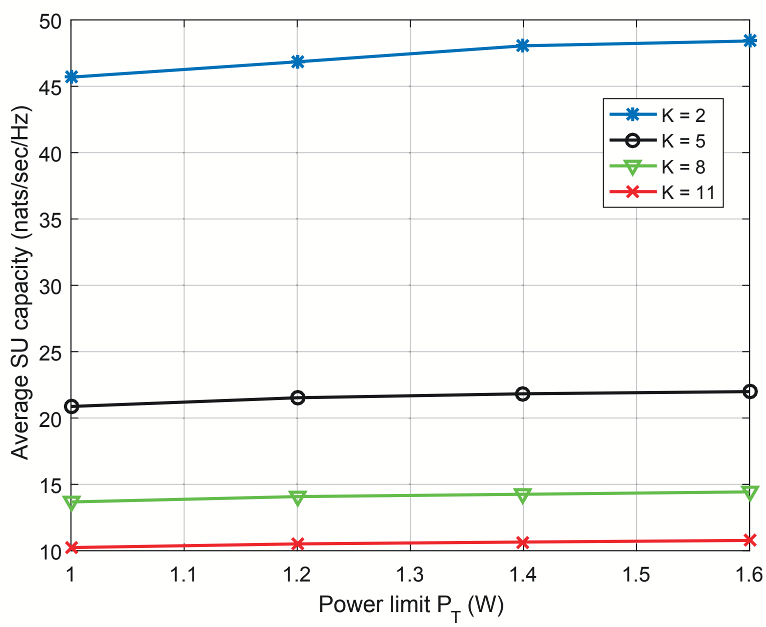


Figure 4.5: Average SU throughput as a function of P_T and K, with $I_l^{\max} = 5 \times 10^{-12}$W.

problem. Under a general and quite practical assumption, we demonstrated that the formulated problem can be globally maximized via standard convex optimization techniques. In particular, the subgradient method was adopted to address the problem in the dual domain. Numerical results were presented to discuss the change of total SU throughput as a function of power limit, interference threshold and the number of SUs.

5. JOINT SPECTRUM SENSING AND RESOURCE ALLOCATION IN MULTI-CHANNEL-MULTI-USER COGNITIVE RADIO NETWORKS

5.1 Introduction

In this section, the joint spectrum sensing and resource allocation problem is investigated in a multi-channel-multi-user CR network [63]. Assuming imperfect spectrum sensing information, our goal is to jointly optimize the sensing threshold and power allocation strategy such that the average total throughput of SUs is maximized. Additionally, the power of SUs is constrained to keep the interference introduced to PUs under certain limits, which gives rise to a non-convex MINLP optimization problem. Our contribution in this section is threefold. First, it is illustrated that the dimension of the non-convex MINLP problem can be significantly reduced, which helps to re-formulate the optimization problem without resorting to integer variables. Second, it is demonstrated that the simplified formulation admits the canonical form of a monotonic optimization and an ϵ-optimal solution can be achieved using the polyblock outer approximation algorithm. Third, a practical low-complexity spectrum sensing and resource allocation algorithm is developed to reduce the computational cost. Finally, the effectiveness of proposed algorithms is verified by simulations.

The rest of the section is structured as follows. The system model and problem formulation are described in Subsection 5.2. By exploring the special structure of the problem in Subsection 5.2, an equivalent but significantly reduced in dimension formulation of the optimization problem is presented in Subsection 5.3. The proposed algorithm exploits the knowledge of a monotonic optimization, and it is shown in Subsection 5.4 to achieve an ϵ-optimal solution. In Subsection 5.5, we develop a suboptimal spectrum sensing and power allocation algorithm with reduced computational complexity. Numerical results and discussions are presented in Subsection 5.6. Finally, conclusions are drawn in Subsection 5.7.

5.2 System Model and Problem Formulation

In order to make the rest of the section easy to follow, a list of some frequently-used terminologies and their corresponding notations in this section is illustrated in Table 5.1.

We consider a downlink multi-band CR interweave network with a SBS coexisting in the vicinity of a PBS. There are total L licensed PUs and K SUs, as shown in Figure 5.1. The CR model in the RF domain is depicted in Figure 5.2. It is assumed that the whole RF spectrum has N total channels with equal bandwidth B. In addition, all the channels are licensed to PUs and the bands of PUs are not overlapping. At a certain time instant, the SBS detects the spectrum and finds the spectrum holes which are licensed to but not utilized by PUs based on the joint optimization rule. These spectrum holes are referred to as the CR bands and are available for SUs' transmission. The key difference between the model proposed here and the one in Subsection 4.2 is that the imperfect spectrum sensing is taken into account here. For example, the PUs' occupancy and the available CR bands in Figure 5.2 are based on the detection information, which might be false. In other words, the occupied bands in Figure 5.2 are subject to the possibility to be idle and the vacant bands in Figure 5.2 might be occupied.

5.2.1 Spectrum Sensing

When the SBS senses the spectrum, two kinds of sensing errors typically exist. The first is termed as false alarm, which occurs when the channel is detected to be occupied but it is actually vacant. The second kind is referred to as misdetection, and it represents the case when the SBS fails to detect the occupancy of PUs, i.e., the channel is detected to be vacant but it is actually occupied. We employ the sensing technique in [20], in which

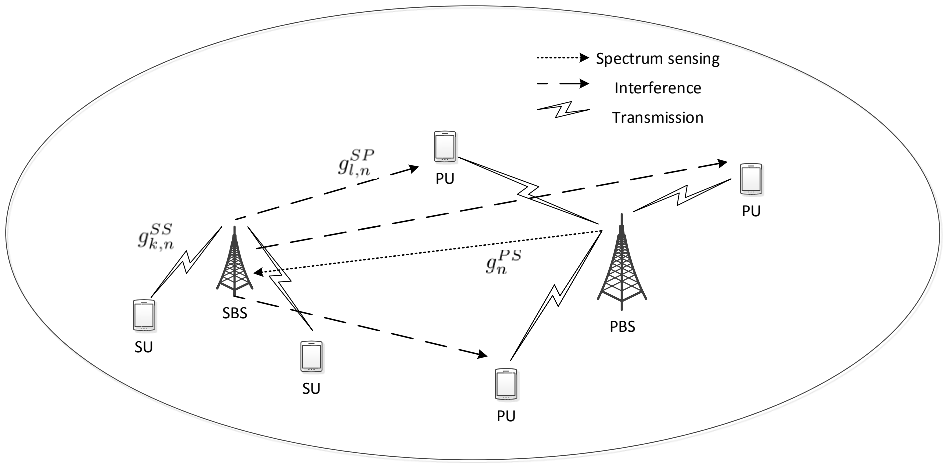


Figure 5.1: A downlink CR network assuming imperfect spectrum sensing.

the false alarm and detection probabilities in the nth channel are expressed as

$$P_n^F(\gamma_n) = P(d_n^1|\mathcal{H}_n^0) = Q\left(\frac{\gamma_n - M\sigma_\nu^2}{\sigma_\nu^2\sqrt{2M}}\right), \tag{5.1}$$

and

$$P_n^D(\gamma_n) = P(d_n^1|\mathcal{H}_n^1) = Q\left(\frac{\gamma_n - M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2)}{\sigma_\nu\sqrt{2M(\sigma_\nu^2 + 2g_n^{PS}\sigma_s^2)}}\right), \tag{5.2}$$

where d_n^0 denotes the event that the nth channel is detected to be vacant and d_n^1 otherwise, $\mathcal{H}_n^0$ represents the hypothesis that the nth channel is actually vacant and $\mathcal{H}_n^1$ otherwise, and γ_n is the sensing threshold. Basically, the hypothesis $\mathcal{H}_n^1$ is chosen if the received signal level is above γ_n on the nth channel. On the other hand, the hypothesis $\mathcal{H}_n^0$ is selected if the level is below γ_n. Additionally, $Q(\cdot)$ is the tail probability of the standard normal distribution, g_n^{PS} represents the power gain from the PBS to the SBS, and M stands for the number of sensing samples. Additionally, σ_s^2 denotes the average power of the transmitted signal from the PBS and σ_ν^2 represents the noise power at the receiver end of SBS, both of which are assumed to be equal at each channel. Similar to [20], we assume that $M, \sigma_s^2, \sigma_\nu^2$,

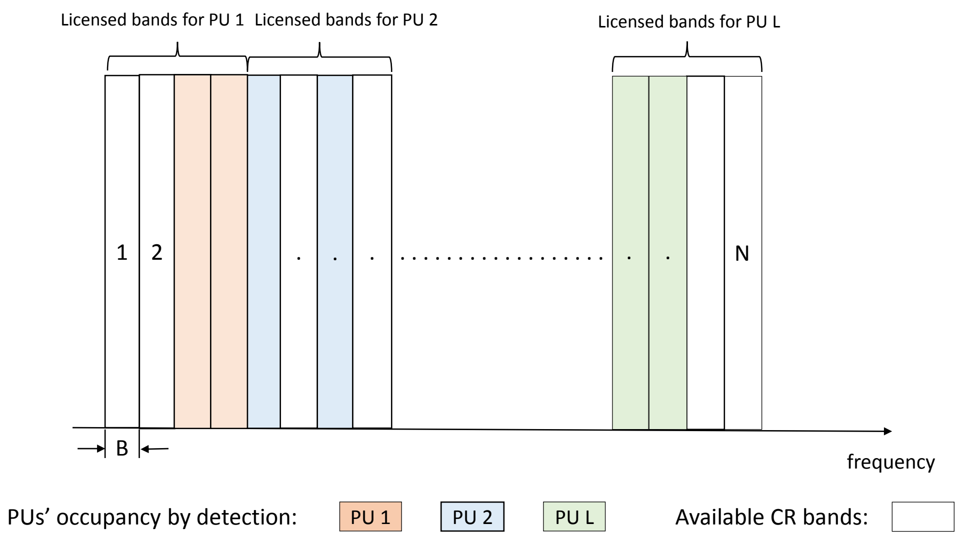


Figure 5.2: RF spectrum sharing between PUs and SUs assuming imperfect spectrum sensing.

and g_n^{PS} are known *a priori* at the SBS. Particularly, the power gain g_n^{PS} can be learned during a period when the primary transmitter is known to be working [16].

The probability of misdetection is then given by

$$P_n^{MD}(\gamma_n) = P(d_n^0|\mathcal{H}_n^1) = 1 - P_n^D(\gamma_n). \tag{5.3}$$

It can be observed that the choice of sensing threshold results in a trade-off between the probabilities of false alarm and misdetection. Due to the monotonically decreasing property of the Q function, a higher sensing threshold yields a smaller probability of false alarm but a larger probability of misdetection, and vice versa.

5.2.2 Resource Allocation

Denoting $p_{k,n}$ as the transmitted power from the SBS to the kth SU on the nth channel in the case when the nth channel is detected to be vacant, and $x_{k,n}$ as the binary variable

Table 5.1: List of terminologies and corresponding notations

L	Number of PUs
K	Number of SUs
N	Number of channels
B	Bandwidth of each channel
γ_n	Sensing threshold for the nth channel
M	Number of sensing samples
σ_ν^2	Noise power
g_n^{PS}	Power gain from the PBS to the SBS
σ_s^2	Average power of the transmitted signal from the PBS
$\mathcal{H}_n^0$	Hypothesis that the nth channel is vacant
$\mathcal{H}_n^1$	Hypothesis that the nth channel is occupied
d_n^0	Event that the nth channel is detected to be vacant
d_n^1	Event that the nth channel is detected to be occupied
P_n^D	Detection rate on the nth channel
P_n^F	False alarm rate on the nth channel
P_n^{MD}	Misdetection rate on the nth channel
P_n^0	Probability that the nth channel is vacant
P_n^1	Probability that the nth channel is occupied
$R_{k,n}$	Average throughput of the kth SU on the nth channel
$R_{k,n}$	Approximated average throughput of the kth SU on the nth channel
$y_{k,n}$	Transmitted power from the SBS to the kth SU on the nth channel
$x_{k,n}$	Binary variable with $x_{k,n} = 1$ indicating that the nth channel is occupied by the kth SU and $x_{k,n} = 0$ otherwise
$I_{k,n}$	Interference from PUs introduced to the kth SU on the nth channel
$g_{k,n}^{SS}$	Power gain from the SBS to the kth SU on the nth channel
$g_{l,n}^{SP}$	Power gain from the SBS to the lth PU on the nth channel
$I_l^{\max}$	Average interference constraint to the lth PU
S_l	Licensed channels for the lth PU
P_n^T	Peak power limit of the SBS on the nth channel

indicating whether the nth channel is occupied by the kth SU, the average throughput of the kth SU on the nth channel can be expressed via Shannon capacity formula as follows:

$$
\begin{aligned}
R_{k,n} = {} & P(d_n^0, \mathcal{H}_n^0) x_{k,n} \log\left(1 + \frac{g_{k,n}^{SS} p_{k,n}}{\sigma_\nu^2}\right) + P(d_n^0, \mathcal{H}_n^1) x_{k,n} \log\left(1 + \frac{g_{k,n}^{SS} p_{k,n}}{\sigma_\nu^2 + I_{k,n}}\right) \\
& + P(d_n^1, \mathcal{H}_n^0) \times 0 + P(d_n^1, \mathcal{H}_n^1) \times 0,
\end{aligned}
$$

where $g_{k,n}^{SS}$ stands for the power gain from the SBS to the kth SU on the nth channel, which is assumed to be known at the SBS. This information can be acquired at the SBS by measuring the uplink channel gain and utilizing the symmetry property between uplink and downlink channels [21]. Moreover, $I_{k,n}$ represents the interference from the PBS introduced to the kth SU on the nth channel. Basically, SUs are only allowed to transmit on the nth channel when it is detected to be vacant, i.e., at the scenario d_n^0. In this case, if the detection is correct, then only noise exists. On the other hand, when it fails to detect the occupancy of PUs, there is interference from PUs to the SU. For simplicity, the logarithm functions are assumed to be natural logarithms in this dissertation. By expressing

$$P(d_n^0, \mathcal{H}_n^0) = P(\mathcal{H}_n^0)P(d_n^0|\mathcal{H}_n^0) = P_n^0(1 - P_n^F),$$

and

$$P(d_n^0, \mathcal{H}_n^1) = P(\mathcal{H}_n^1)P(d_n^0|\mathcal{H}_n^1) = P_n^1(1 - P_n^D),$$

where $P_n^0 = P(\mathcal{H}_n^0)$ and $P_n^1 = P(\mathcal{H}_n^1)$ denote the priori probabilities and can be obtained via long-term estimation. The above equation resumes to

$$R_{k,n} = P_n^0(1 - P_n^F)x_{k,n} \log\left(1 + \frac{g_{k,n}^{SS}p_{k,n}}{\sigma_\nu^2}\right) + P_n^1(1 - P_n^D)x_{k,n} \log\left(1 + \frac{g_{k,n}^{SS}p_{k,n}}{\sigma_\nu^2 + I_{k,n}}\right). \tag{5.4}$$

In order to formulate the optimization problem, the following constraints are considered:

- QoS of PUs: To sufficiently protect the QoS of PUs, an interference threshold $I_l^{\max},\ l = 1, \cdots, L$, is enforced to constrain the average interference introduced at PUs by SUs:

$$\sum_{k=1}^{K}\sum_{n\in S_l}P_n^1(1-P_n^D)x_{k,n}p_{k,n}g_{l,n}^{SP}\leq I_l^{\max},\ l=1,\cdots,L, \tag{5.5}$$

where $g_{l,n}^{SP}$ denotes the power gain from the SBS to the lth PU on the nth channel, which can be periodically measured by a band manager or estimated by listening to a becon signal and then fed back to the SBS [55], and S_l represents the licensed channels for the lth PU.

- Transmit power constraint: Instead of considering the average power constraint, which only guarantees that the power limit is not exceeded in an average sense, we consider the peak power constraint herein dissertation. Let P_n^T represent the maximum transmit power of the SBS on the nth channel. The power constraint reduces to

$$\sum_{k=1}^{K}x_{k,n}p_{k,n}\leq P_n^T,\ n=1,\cdots,N. \tag{5.6}$$

Additionally, constraint (5.6) also prevents the ill-conditioned situation in constraint (5.5). Specifically, if $P_n^1(1-P_n^D)$ admits a very small value on some channel n, then $\sum_{k=1}^{K}x_{k,n}p_{k,n}$ can even take a large value close to infinity and still make (5.5) satisfied.

- Interference constraint among SUs: To prevent the interference among SUs, we simply add a constraint requiring that each channel can be occupied by at most one SU, i.e.,

$$\sum_{k=1}^{K}x_{k,n}\leq 1,\ n=1,\cdots,N. \tag{5.7}$$

- Sensing error constraint: In order to receive the best balance between false alarm and misdetection, the following sensing error constraints are added:

$$P_n^D(\gamma_n) \geq 0.5,\ P_n^F(\gamma_n) \leq 0.5,\ n = 1, \cdots, N. \tag{5.8}$$

Due to the monotonically decreasing property of Q function in (5.1) and (5.2), and the fact that $Q^{-1}(0.5) = 0$, constraints (5.8) can be further expressed as

$$M\sigma_\nu^2 \leq \gamma_n \leq M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2),\ n = 1, \cdots, N. \tag{5.9}$$

- Basic constraints for $x_{k,n}$ and $p_{k,n}$: Since $x_{k,n}$ is a binary variable indicating whether the nth channel is occupied by the kth SU, it follows that

$$x_{k,n} \in \{0, 1\},\ k = 1, \cdots, K,\ n = 1, \cdots, N. \tag{5.10}$$

In addition, the non-negative property of the power variables is expressed as

$$p_{k,n} \geq 0,\ k = 1, \cdots, K,\ n = 1, \cdots, N. \tag{5.11}$$

Since we are investigating a CR network where the licensed spectrum is inefficiently utilized, the PU activity probability P_n^1 is assumed to be very small, say less than 0.25, and it follows that $P_n^0 > P_n^1$. Moreover, based on (5.8), it turns out that $1 - P_n^F > 1 - P_n^D$. Therefore, combing these two observations and the fact that $\sigma_\nu^2 < \sigma_\nu^2 + I_{k,n}$, the first term in (5.4) dominates $R_{k,n}$, and $R_{k,n}$ can be further approximated as

$$\tilde{R}_{k,n} = P_n^0(1 - P_n^F)x_{k,n} \log\left(1 + \frac{g_{k,n}^{SS} p_{k,n}}{\sigma_\nu^2}\right). \tag{5.12}$$

The joint spectrum sensing and resource allocation problem, referred to as "P-Joint" herein dissertation, is formulated as

$$\max_{x_{k,n},p_{k,n},\gamma_n} \quad \sum_{k=1}^{K}\sum_{n=1}^{N}\tilde{R}_{k,n} \tag{P-Joint}$$

$$\text{s.t.} \;\; (5.5),(5.6),(5.7),(5.9),(5.10),(5.11).$$

5.3 An Equivalent Formulation

Since the objective function of P-Joint is non-convex with respect to $x_{k,n}, p_{k,n}$ and γ_n, it is a non-convex MINLP problem, which is NP-hard and computationally complex to solve [36]. In this subsection, we illustrate that under a general and practical assumption, the dimension of the original non-convex MINLP problem can be significantly reduced. Basically, the binary variable $x_{k,n}$ can be removed which results in an equivalent formulation only in terms of the power and sensing threshold variables.

Instead of directly analyzing P-Joint with respect to $x_{k,n}, p_{k,n}$ and γ_n, we alternatively tackle the problem with a fixed value of γ_n satisfying the constraint (5.9). In this way, the resulting problem can be expressed as

$$\max_{x_{k,n},p_{k,n}} \quad \sum_{k=1}^{K}\sum_{n=1}^{N}\tilde{R}_{k,n} \tag{5.13}$$

$$\text{s.t.} \;\; (5.5),(5.6),(5.7),(5.10),(5.11).$$

It can be observed (5.13) is still a non-convex MINLP problem. However, by exploring the special structure in the dual domain, it is demonstrated that the optimal solution of (5.13) can be derived using standard convex optimization tools under a general and practical assumption. The same procedure used in Subsection 4.2 is employed here. Specifically, we first introduce a new variable $y_{k,n}$, where $y_{k,n} = x_{k,n}p_{k,n}$, to replace the variable $p_{k,n}$. Thus, the resulting problem is equivalent to (5.13). The time-sharing method is then used by relaxing the binary variable $x_{k,n} \in \{0,1\}$ to $0 \leq x_{k,n} \leq 1$. Even though the time-sharing method is employed, unlike reference [26], we do not allow multiple SUs to share a common channel. The time-sharing method is just utilized to conduct the continuous relaxation of the MINLP problem such that the convex optimization tools can be implemented. Thus, the problem (5.13) is relaxed to

$$\begin{aligned}
&\max_{x_{k,n},y_{k,n}} \quad \sum_{k=1}^{K}\sum_{n=1}^{N} P_n^0(1-P_n^F)x_{k,n}\log\left(1+\frac{g_{k,n}^{SS}y_{k,n}}{\sigma_\nu^2 x_{k,n}}\right) \qquad (5.14)\\
&\text{s.t.} \ \sum_{k=1}^{K}\sum_{n\in S_l} P_n^1(1-P_n^D)y_{k,n}g_{l,n}^{SP} \leq I_l^{\max},\ l=1,\cdots,L\\
&\quad \sum_{k=1}^{K} y_{k,n} \leq P_n^T,\ n=1,\cdots,N\\
&\quad \sum_{k=1}^{K} x_{k,n} \leq 1,\ \ n=1,\cdots,N\\
&\quad x_{k,n}\geq 0, p_{k,n}\geq 0,\ \ k=1,\cdots,K,\ \ n=1,\cdots,N
\end{aligned}$$

Since constraint $\sum_{k=1}^{K} x_{k,n} \leq 1$ implicitly implies that $x_{k,n} \leq 1$, the relaxation condition is simply expressed as $x_{k,n} \geq 0$ instead of $0 \leq x_{k,n} \leq 1$. Analogous to the convexity proof of problem (P_2) shown in Appendix A, it can be readily inferred that $x_{k,n}\log(1+g_{k,n}^{SS}y_{k,n}/(\sigma_\nu^2 x_{k,n}))$ is a concave function with respect to $x_{k,n}$ and $y_{k,n}$, since it is the perspective function of the concave logarithm function. The objective function of (5.14), being the positive summation of concave functions, is also concave. Hence, the optimization problem (5.14) is a convex optimization problem due to the fact that all its constraints are affine.

We tackle the problem in the dual domain, which leads to the adoption of the Lagrange function

$$Lg = \sum_{k=1}^{K}\sum_{n=1}^{N} P_n^0(1-P_n^F)x_{k,n}\log\left(1+\frac{g_{k,n}^{SS}y_{k,n}}{\sigma_\nu^2 x_{k,n}}\right) + \sum_{n=1}^{N}\lambda_n\left(P_n^T - \sum_{k=1}^{K} y_{k,n}\right)$$
$$+ \sum_{l=1}^{L}\mu_l\left(I_l^{\max} - \sum_{k=1}^{k}\sum_{n\in S_l} P_n^1(1-P_n^D)y_{k,n}g_{l,n}^{SP}\right) + \sum_{n=1}^{n}\eta_n\left(1-\sum_{k=1}^{K}x_{k,n}\right), \tag{5.15}$$

where the last constraint in (5.14) is absorbed in KKT conditions. Applying KKT conditions [35] leads to

$$\frac{\partial Lg}{\partial y_{k,n}} = P_n^0(1-P_n^F)\frac{g_{k,n}^{SS}x_{k,n}}{\sigma_\nu^2 x_{k,n} + g_{k,n}^{SS}y_{k,n}} - \lambda_n - P_n^1(1-P_n^D)\mu_l g_{l,n}^{SP}\begin{cases} = 0, & y_{k,n} > 0, \\ < 0, & y_{k,n} = 0, \end{cases}$$

with $n \in S_l$. It turns out that

$$y_{k,n} = \left(\psi_n - \frac{\sigma_\nu^2}{g_{k,n}^{SS}}\right)^+ x_{k,n}, \tag{5.16}$$

where

$$\psi_n = \frac{P_n^0(1-P_n^F)}{\lambda_n + P_n^1(1-P_n^D)\mu_l g_{l,n}^{SP}},$$

with $n \in S_l$,. Moreover, taking the derivative of the Lagrange function with respect to $x_{k,n}$ and plugging (5.16) into the result yields

$$\frac{\partial Lg}{\partial x_{k,n}} = J_{k,n} - \eta_n\begin{cases} = 0, & x_{k,n} > 0, \\ < 0, & x_{k,n} = 0, \end{cases} \tag{5.17}$$

where

$$J_{k,n} = P_n^0(1-P_n^F)\left[\log\left(1+\frac{g_{k,n}^{SS}\left(\psi_n - \sigma_\nu^2/g_{k,n}^{SS}\right)^+}{\sigma_\nu^2}\right) - \frac{g_{k,n}^{SS}(\psi_n - \sigma_\nu^2/g_{k,n}^{SS})^+}{\sigma_\nu^2 + g_{k,n}^{SS}(\psi_n - \sigma_\nu^2/g_{k,n}^{SS})^+}\right]. \tag{5.18}$$

Based to the aforementioned equations, the following result can be established.

Lemma 2. *All the channels must be occupied by SUs. That is to say, for any channel n, there exist some SUs such that $x_{k,n} > 0$. In addition, $\eta_n = \max_k J_{k,n},\ n = 1, \cdots, N$.*

Proof. See Appendix C.1. □

Based on Lemma 2, the following more general result is derived.

Theorem 6. *Under the assumption that for any channel n, $g_{k,n}^{SS}$'s are all distinct for $k = 1, \cdots, K$, the problem* (5.14) *always achieve a binary optimal solution for the user assignment variable $x_{k,n}$. Particularly, each channel is solely assigned to the SU with the largest $g_{k,n}^{SS}$, i.e.,*

$$\begin{aligned} x_{k_n^*,n} &= 1|_{k_n^*=\arg\max_k g_{k,n}^{SS}},\ n = 1, \cdots, N, \\ x_{k,n} &= 0|_{k \neq k_n^*},\ n = 1, \cdots, N. \end{aligned} \tag{5.19}$$

Proof. See Appendix C.2. □

Generally, SUs are randomly located around the SBS and the probability that two SUs present exactly the same power gain $g_{k,n}^{SS}$ is almost zero. Therefore, the optimization problem (5.14) always achieve a binary optimal solution for the user assignment variable. It can be further claimed that the optimal binary solution in (5.19) for problem (5.14) is also optimal for problem (5.13). This argument can be proved in two steps: first, the optimal objective value of (5.14) serves as an upper-bound for (5.13) due to the fact that (5.14) is a relaxed version of (5.13); second, since the optimal solution (5.19) for (5.14) is binary, it represents a feasible solution of (5.13) and consequently achieves a lower-bound for the optimal objective value of (5.13).

Towards this end, it turns out that for a fixed γ_n, the optimization problem (5.13) admits an optimal solution (5.19) for the user assignment variable $x_{k,n}$. Since this statement holds for any feasible γ_n, by plugging (5.19) into the formulation of P-Joint, it can be concluded that the original joint optimization P-Joint can be simplified to

$$\max_{p_{k_n^*,n},\gamma_n} \quad \sum_{n=1}^{N} P_n^0(1-P_n^F)\log\left(1+\frac{g_{k_n^*,n}^{SS}p_{k_n^*,n}}{\sigma_\nu^2}\right) \tag{P-JointE}$$

$$\text{s.t.} \quad \sum_{n\in S_l} P_n^1(1-P_n^D)p_{k_n^*,n}g_{l,n}^{SP} \le I_l^{\max},\ l=1,\cdots,L \tag{C_1}$$

$$0 \le p_{k_n^*,n} \le P_n^T,\ n=1,\cdots,N \tag{C_2}$$

$$M\sigma_\nu^2 \le \gamma_n \le M(\sigma_\nu^2+g_n^{PS}\sigma_s^2),\ n=1,\cdots,N, \tag{C_3}$$

where $k_n^* = \arg\max_k g_{k,n}^{SS},\ n=1,\cdots,N$. Since the above optimization problem is equivalent to P-Joint, it is termed as "P-JointE". For the rest of this section, we focus on addressing P-JointE since its dimension is significantly less than that of P-Joint and the integer variable $x_{k,n}$ has been removed.

5.4 Optimal Joint Spectrum Sensing and Resource Allocation Algorithm

Due to the non-convex property of the objective function, P-JointE is still a non-convex optimization problem with respect to $p_{k_n^*,n}$ and γ_n. However, it can be readily found that P-JointE is a convex optimization problem with regard to $p_{k_n^*,n}$ while fixing γ_n. Moreover, it also admits the formulation of a convex optimization in terms of γ_n while fixing $p_{k_n^*,n}$. This is due to the fact that P_n^F and $1-P_n^D$ are convex functions with respect to γ_n under the constraint (5.8), or equivalently (C_3) [20]. In literature, the non-convex optimization with such kind of formulation is typically solved by means of the AO method [25, 26], where the sensing threshold and power variable are iteratively optimized while fixing the other. However, as discussed earlier, the convergence of the AO method heavily depends on the initial point and the iterative algorithm can be trapped at a local maximum near the starting point. Therefore, in this subsection, we propose an algorithm by exploiting the knowledge of a monotonic optimization to solve the joint optimization problem P-JointE. It is shown that the problem admits the canonical form of a monotonic optimization, which can be addressed in a finite steps with an ϵ-optimal solution[1].

5.4.1 Polyblock Outer Approximation Algorithm

Since the preliminaries of monotonic optimization have been presented in Subsection 3.1, in this subsubsection, the polyblock outer approximation algorithm, which aims to solving P-JointE, is described via a detailed illustration.

Let $[\mathbf{p}, \boldsymbol{\gamma}] = [p_{k_1^*,1}, \cdots, p_{k_N^*,N}, \gamma_1, \cdots, \gamma_N]$ be a vector with length $2N$, the objective function for P-JointE is an increasing function with respect to $[\mathbf{p}, \boldsymbol{\gamma}]$ since $1 - P_n^F$ is an increasing function of γ_n and the logarithm is an increasing function of $p_{k_n^*,n}$. Moreover, in the region

$$\left[\mathbf{0}, [P_1^T, \cdots, P_N^T, M(\sigma_\nu^2 + g_1^{PS}\sigma_s^2), \cdots, M(\sigma_\nu^2 + g_N^{PS}\sigma_s^2)]\right], \tag{5.20}$$

the normal set $\mathcal{G}$ and the conormal set $\mathcal{H}$ can be expressed as

$$\mathcal{G} = \left\{(\mathbf{p}, \boldsymbol{\gamma}) | (C_1), (C_2), 0 \leq \gamma_n \leq M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2)\right\},$$

and

$$\mathcal{H} = \left\{(\mathbf{p}, \boldsymbol{\gamma}) | p_{k_n^*,n} \geq 0, \gamma_n \geq M\sigma_\nu^2\right\},$$

respectively. A monotonic optimization problem can be solved using the polyblock outer

[1] $\bar{\mathbf{x}}$ is an ϵ-optimal solution if $f(\mathbf{x}^*) - \epsilon \leq f(\bar{\mathbf{x}}) \leq f(\mathbf{x}^*)$, where $\mathbf{x}^*$ represents the optimal solution.

approximation algorithm with an ϵ-optimal solution provided that the conormal set $\mathcal{H}$ is strictly bounded away from $\mathbf{0}$, i.e., there exists a positive vector $\mathbf{a}$ such that $\mathbf{0} < \mathbf{a} \leq \mathbf{x}$, $\forall \mathbf{x} \in \mathcal{H}$. It can be observed that this condition does not hold for our formulation due to the fact that $p_{k_n^*,n}$ can be zero in $\mathcal{H}$. In computer simulations, this issue can be addressed by setting a very small number δ, say 10^{-3}, as a lower-bound for $p_{k_n^*,n}$ without affecting the simulation results. Alternatively, one can shift the origin to a negative coordinate such that the conormal set $\mathcal{H}$ is strictly bounded away from the new origin [51].

The proposed polyblock outer approximation algorithm for P-JointE is summarized in Algorithm 2. The basic idea of polyblock outer approximation algorithm is to enclose the feasible set $\mathcal{G} \cap \mathcal{H}$ as closely as possible by a nested sequence of "polyblocks", starting with the polyblock $\mathcal{P}_1$ in (5.20). The starting polyblock $\mathcal{P}_1$ is illustrated in Figure 5.3(a) where a $2N$ dimension figure is plotted in a two-dimension plane for simplicity. Among all the vertices of the polyblock (in $\mathcal{P}_1$, we only have one vertex $\mathbf{z}_1$), we choose the one with the largest objective value (Line 3) and project this vertex to the upper-boundary of $\mathcal{G}$ (Line 4). The projection process is mathematically defined as the following single-variable optimization problem

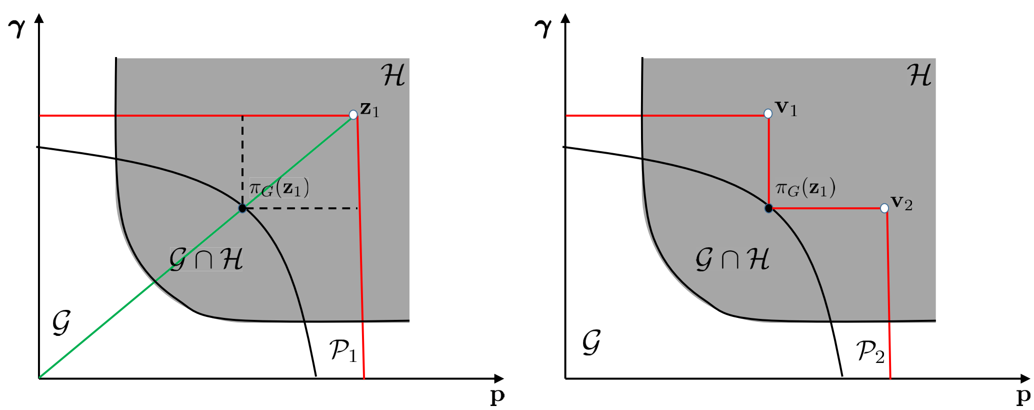
$$\alpha = \arg\max \alpha|_{\alpha>0, \alpha\mathbf{z}_i \in \mathcal{G}},$$

which can be solved efficiently by the bisection search algorithm due to the normality of $\mathcal{G}$. The projection of the best vertex is termed as $\pi_{\mathcal{G}}(\mathbf{z}_i)$ in Algorithm 2. From Line 5 to Line 11, the algorithm determines the current best solution and the current best value (CBV), and store them as $[\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]$ and CBV, respectively. In Line 12, a new polyblock $\mathcal{P}_{i+1}$ with vertices $\mathcal{T}_{i+1}$ is generated by cutting off a cone that is in the infeasible set. In particular, the best vertex $\mathbf{z}_i$ is removed and $2N$ new vertices $\{\mathbf{v}_i,\ i = 1, \cdots, 2N\}$ are added. The newly-added vertices together with the remaining vertices $\mathcal{T}_i \backslash \mathbf{z}_i$ generate the new polyblock. For the illustration purpose, the new polyblock $\mathcal{P}_2$, which comes from the polyblock $\mathcal{P}_1$, is depicted in Figure 5.3(b). In Line 13, improper vertices are removed to accelerate the convergence speed and save the memory. Due to the monotonically increasing property of the objective function, removing improper vertices does not affect the shape of the polyblock [52].

This procedure is repeated until the current best value is close to the best vertex value, i.e., $|f(\mathbf{z}_i) - \text{CBV}_i| \leq \epsilon$. In this way, a sequence of polyblocks

$$\mathcal{P}_1 \supset \mathcal{P}_2 \supset \cdots \supset \mathcal{P}_i \supset \cdots \supset \mathcal{G} \cap \mathcal{H}$$

is constructed and the feasible set is approximated more and more closely. The current best solution $[\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]$ is guaranteed to be ϵ-optimal.



(a) (b)

Figure 5.3: (a) The initialized polyblock $\mathcal{P}_1$; (b) The polyblock $\mathcal{P}_2$ generated from $\mathcal{P}_1$.

The dimension of the problem, which directly determines the size of polyblock vertex set, is a key factor for the computational complexity of monotonic optimization. Additionally, the form of constraints also affects the complexity since the feasibility is required to

Algorithm 2 Polyblock outer approximation algorithm for problem P-JointE

Input: Function f: The objective function in P-JointE.

$$\mathcal{G} = \left\{(\mathbf{p}, \boldsymbol{\gamma}) | (C_1), (C_2), 0 \leq \gamma_n \leq M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2)\right\},$$

$$\mathcal{H} = \left\{(\mathbf{p}, \boldsymbol{\gamma}) | p_{k_n^*,n} \geq \delta, \gamma_n \geq M\sigma_\nu^2\right\}$$

Output: An ϵ-optimal solution $[\mathbf{p}^*, \boldsymbol{\gamma}^*]$

Initialization: Error tolerance $\epsilon > 0$. $\text{CBV}_0 = -\infty$. $i = 0$. $[\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]_0 = \mathbf{0}$. The vertex set $\mathcal{T}_1 = [P_1^T, \cdots, P_N^T, M(\sigma_\nu^2 + g_1^{PS}\sigma_s^2), \cdots, M(\sigma_\nu^2 + g_N^{PS}\sigma_s^2)]$.

while $|f(\mathbf{z}_i) - \text{CBV}_i| \leq \epsilon$ **do**
  $i = i + 1$
  From $\mathcal{T}_i$, select $\mathbf{z}_i = \arg\max f(\mathbf{z})|_{\mathbf{z} \in \mathcal{T}_i}$.
  Find $\pi_{\mathcal{G}}(\mathbf{z}_i) = \alpha \mathbf{z}_i,\ \alpha = \arg\max \alpha|_{\alpha > 0, \alpha \mathbf{z}_i \in \mathcal{G}}$.
  if $\pi_{\mathcal{G}}(\mathbf{z}_i) = \mathbf{z}_i$, i.e., $\mathbf{z}_i \in \mathcal{G}$ **then**
    $[\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]_i = \mathbf{z}_i$ and $\text{CBV}_i = f(\mathbf{z}_i)$.
  else if $\pi_{\mathcal{G}}(\mathbf{z}_i) \in \mathcal{H}$ and $f(\pi_{\mathcal{G}}(\mathbf{z}_i)) \geq \text{CBV}_{i-1}$ **then**
    $[\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]_i = \pi_{\mathcal{G}}(\mathbf{z}_i)$ and $\text{CBV}_i = f(\pi_{\mathcal{G}}(\mathbf{z}_i))$.
  else
    $[\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]_i = [\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]_{i-1}$ and $\text{CBV}_i = \text{CBV}_{i-1}$.
  end if
  $\mathcal{T}_{i+1} = (\mathcal{T}_i \backslash \mathbf{z}_i) \cup \{\mathbf{v}_i,\ i = 1, \cdots, 2N\}$, where $\mathbf{v}_i$ is obtained by replacing the ith entry of $\mathbf{z}_i$ by the ith entry of $\pi_{\mathcal{G}}(\mathbf{z}_i)$.
  Remove from $\mathcal{T}_{i+1}$ all the improper vertices.
end while
Let $[\mathbf{p}^*, \boldsymbol{\gamma}^*] = [\bar{\mathbf{p}}, \bar{\boldsymbol{\gamma}}]_i$ and terminate the algorithm.

be checked in each iteration. As an under-explored research area, limited computational experience has shown that polyblock algorithms work very well for monotonic optimization problems with a small dimension m, generally $m \leq 10$ [24]. The reader can find more details about monotonic optimization as well as its applications in communication and networking systems in [52] and references therein.

The problem P-JointE, at first glance, has a dimension $2N$. However, since the licensed channels for PUs are non-overlapping, i.e., $S_i \cap S_j = 0,\ i \neq j$ and all the channels are licensed to PUs, i.e., $S_1 \cup \cdots \cup S_l = \{1, \cdots, N\}$, P-JointE actually can be decomposed into

L independent optimization problems whose dimensions are $|S_1|, \cdots, |S_L|$, respectively. Specifically, these L independent optimization problems are

$$\begin{aligned}
\max_{p_{k_n^*,n}, \gamma_n} \quad & \sum_{n \in S_l} P_n^0 (1 - P_n^F) \log \left(1 + \frac{g_{k_n^*,n}^{SS} p_{k_n^*,n}}{\sigma_\nu^2} \right) \\
\text{s.t.} \quad & \sum_{n \in S_l} P_n^1 (1 - P_n^D) p_{k_n^*,n} g_{l,n}^{SP} \le I_l^{\max} \\
& 0 \le p_{k_n^*,n} \le P_n^T, \; n \in S_l \\
& M\sigma_\nu^2 \le \gamma_n \le M(\sigma_\nu^2 + g_n^{PS} \sigma_s^2), \; n \in S_l,
\end{aligned}$$

with $l = 1, \cdots, L$. In this way, as long as the largest number of channels occupied by licensed PUs is small, the monotonic optimization technique can be used to achieve an ϵ-optimal solution.

5.5 Low-Complexity Suboptimal Algorithm

Algorithm 2 provides an ϵ-optimal solution for the joint spectrum sensing and resource allocation problem P-JointE. However, the proposed polyblock outer approximation algorithm is typically suitable for a small-dimension monotonic optimization problem and the computational complexity is very high in a CR network with a large number of channels, such as the OFDM wireless system. In this subsection, we propose a low-complexity suboptimal algorithm by first modifying the constraint of P-JointE and then solving a list of single-variable optimization problems.

Instead of limiting the total interference to PUs, we alternatively impose a threshold for the interference to each licensed channel of PUs. Mathematically, the modified interference constraint is expressed as

$$P_n^1 (1 - P_n^D) p_{k_n^*,n} g_{l,n}^{SP} \le I_l^{\max} / |S_l|, \; \forall n \in S_l, \; l = 1, \cdots, L. \qquad (C_1')$$

Thus, the joint spectrum sensing and resource allocation problem can be reformulated as

$$\max_{p_{k_n^*,n},\gamma_n} \sum_{n=1}^{N} P_n^0(1-P_n^F)\log\left(1+\frac{g_{k_n^*,n}^{SS}p_{k_n^*,n}}{\sigma_\nu^2}\right) \tag{P-JointEM}$$
$$\text{s.t.} \ (C_1'),(C_2),(C_3)$$

It can be inferred that constraint (C_1') is more stringent compared to constraint (C_1). Thus, P-JointEM, being a modified version of P-JointE, provides a lower-bound for P-JointE. Moreover, with the help of the modified constraint (C_1'), P-JointEM can be decomposed into N independent subproblems as shown below

$$\max_{p_{k_n^*,n},\gamma_n} P_n^0(1-P_n^F)\log\left(1+\frac{g_{k_n^*,n}^{SS}p_{k_n^*,n}}{\sigma_\nu^2}\right) \tag{P-JointEM-n}$$
$$\text{s.t.} \ (C_1'),(C_2),(C_3)$$

with $n=1,\cdots,N$. In this way, we only need to solve N optimization problems, each of which presents two variables $p_{k_n^*,n}$ and γ_n.

To solve P-JointEM-n, it is first claimed that either constraint (C_1') or (C_2) achieves the equality at the optimal solution. This claim can be easily verified via a proof by contradiction. First, it is assumed that neither constraint achieves the equality at the optimal solution. Then, we can definitely add more power to this channel to achieve the equality at either constraint while obtaining a larger objective value, which completes the proof by contradiction. Consequently, $p_{k_n^*,n}$ can be expressed in terms of γ_n as follows:

$$p_{k_n^*,n} = \min\left(\frac{I_l^{\max}}{|S_l|P_n^1(1-P_n^D)g_{l,n}^{SP}}, P_n^T\right) = \frac{I_l^{\max}}{\max\left(|S_l|P_n^1(1-P_n^D)g_{l,n}^{SP}, I_l^{\max}/P_n^T\right)}, \tag{5.21}$$

and P-JointEM-n resumes to a single-variable optimization problem

$$\begin{aligned} \max_{\gamma_n} \ & f(\gamma_n) - g(\gamma_n) \\ \text{s.t.} \ & M\sigma_\nu^2 \leq \gamma_n \leq M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2), \end{aligned} \tag{5.22}$$

with

$$f(\gamma_n) = P_n^1(1 - P_n^F)\log(\sigma_\nu^2 + g_{k,n}^{SS} I_l^{\max}),$$

and

$$g(\gamma_n) = P_n^1(1 - P_n^F) \times \log\left[\sigma_\nu^2 \cdot \max\left(|S_l| P_n^1(1 - P_n^D) g_{l,n}^{SP}, \frac{I_l^{\max}}{P_n^T}\right)\right].$$

Problem (5.22) can be solved using a one-dimensional brute-force algorithm. To avoid the exhaustive search, in the rest of this subsection, we show that it can also be addressed using the monotonic optimization technique described in Subsection 5.4.

Both P_n^D and P_n^F are monotonically decreasing functions in terms of γ_n. As a result, the cost function in (5.22) is the difference between two increasing functions $f(\gamma_n)$ and $g(\gamma_n)$. It is shown in Appendix E that problem (5.22) can be rewritten as the canonical form of a monotonic optimization by introducing a new parameter t:

$$\begin{aligned} \max \ & f(\gamma_n) + t \\ \text{s.t.} \ & (\gamma_n, t) \in \mathcal{G} \cap \mathcal{H}, \end{aligned} \tag{5.23}$$

with

$$\begin{aligned} \mathcal{G} = \Big\{ (\gamma_n, t) \Big| & 0 \leq \gamma_n \leq M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2), 0 \leq t \leq g\left(M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2)\right) - g\left(M\sigma_\nu^2\right), \\ & g(\gamma_n) + t \leq g\left(M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2)\right) \Big\}, \end{aligned} \tag{5.24}$$

and

$$\mathcal{H} = \{(\gamma_n, t) | \gamma_n \geq M\sigma_\nu^2, t \geq 0\}. \tag{5.25}$$

We can implement the same technique as in Subsection 5.4 to achieve an ϵ-optimal solution for variable γ_n in problem (5.23), or equivalently P-JointEM-n. Then the power solution can be obtained using (5.21).

5.6 Simulations and Discussions

For the numerical results presented in this subsection, we consider two different kinds of scenarios, namely a CR network with a small number of channels and one with a large number of channels. For the former scenario, we compare the simulation performance among three algorithms: i) The proposed optimal monotonic optimization algorithm in Subsection 5.4; ii) The proposed suboptimal monotonic optimization algorithm in Subsection 5.5; and iii) The state-of-the-art AO method in literature, which is briefly discussed in the beginning of Subsection 5.4 and fully illustrated in Appendix E. For the later scenario, only the simulation for the second and third algorithms aforementioned is carried out, since the optimal monotonic optimization algorithm does not work very well for a large size problem. However, notice that the suboptimal solution for P-JointEM-n can be further enhanced by plugging it back into P-JointE as an initial point of the AO method. Therefore, this scheme is termed as “suboptimal-enhanced” and included as a benchmark for the simulation in a large-dimension CR network. As discussed earlier in the section, some references studied the CR problem with a fixed sensing parameter or a fixed physical layer resource layout. We do not consider these set-ups in this subsection since the joint spectrum sensing and resource allocation problem obviously outperforms these set-ups from a mathematical point of view.

5.6.1 A Small-Dimension CR Network

In this subsubsection, the numbers of PUs and SUs are assumed to be 3, i.e., $L = 3$ and $K = 3$. The whole RF spectrum presents $N = 6$ channels. The licensed channels for these three PUs are 1-2, 3-4 and 5-6, respectively. The PU activity probability P_n^1 is set to be 0.2 for all the channels. For the sensing parameters, it is assumed that the number of sensing samples $M = 10$ and the average power of the transmitted signal at the PBS $\sigma_s^2 = 10W$. For the resource allocation parameters, the channels from the SBS to SUs, as well as the channels from the SBS to PUs, are assumed to be Rayleigh fading with an average channel power gain 5dB. Thus, $g_{k,n}^{SS}$ and $g_{l,n}^{SP}$ are both distributed exponentially with mean 5dB. The noise power σ_ν^2 is assumed to be 1W.

Since the channel gains are randomly generated, an average result of 100 independent Monte-Carlo runs is carried out for calculating the average total throughput of secondary network. In Figure 5.4, the average total throughput of SUs is simulated with a power limit P_n^T ranging from 10W to 30W for all channels and an interference threshold $I_l^{\max} = 1$W, $l = 1, \cdots, L$. Moreover, we consider the sensing SNR at the SBS in two cases. Specifically, the sensing channels from the PBS to the SBS, g_n^{PS}, are assumed to be Rayleigh fading with an average channel power gain -10dB and -4dB for all channels. Thus, the sensing SNRs, $\Gamma = \mathrm{E}[g_n^{PS}]\sigma_s^2/\sigma_\nu^2$, at the SBS are 0dB and 6dB for all channels. It can be seen that with the power limit increasing, the average total SU throughput also increases. In addition, the suboptimal monotonic optimization algorithm outperforms the AO method and is very close to the optimal solution in both sensing SNRs. The same trend can be observed in Figure 5.5 with $P_n^T = 25$W and $I_l^{\max}$ ranging from 0.4W to 1.2W.

In Figure 5.6 and Figure 5.7, the simulation with one realization of channel gains is performed coupled with parameters $P_n^T = 25$W, $I_l^{\max} = 1$W and $\Gamma = 6$dB. Particularly, in Figure 5.6, the probabilities of misdetection and false alarm, together with the allocated power, for each individual channel are illustrated. For the 1st, 2nd, 5th and 6th channel where the probabilities of misdetection and false alarm are low, the maximum power P_n^T is assigned. In contrast, for the 4th channel where the sensing errors are relatively high, a very low power is allocated to protect the QoS of PUs. In Figure 5.7, 10 independent simulations are performed of these algorithms with a fixed realization of channel gains. Since both the optimal monotonic and the suboptimal monotonic algorithms achieve a solution at most ϵ away from its own optimal solution, their performances keep consistent. On the other hand, the performance of AO method varies a lot due to different selections of the initial guess.

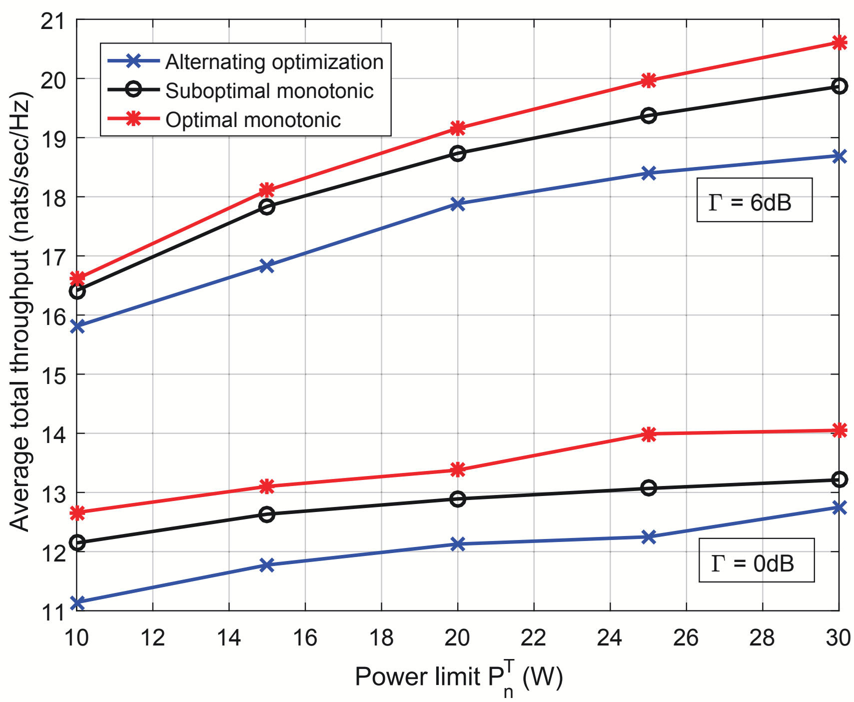


Figure 5.4: Average total throughput of SUs in terms of the power limit and the sensing SNR ($N = 6$).

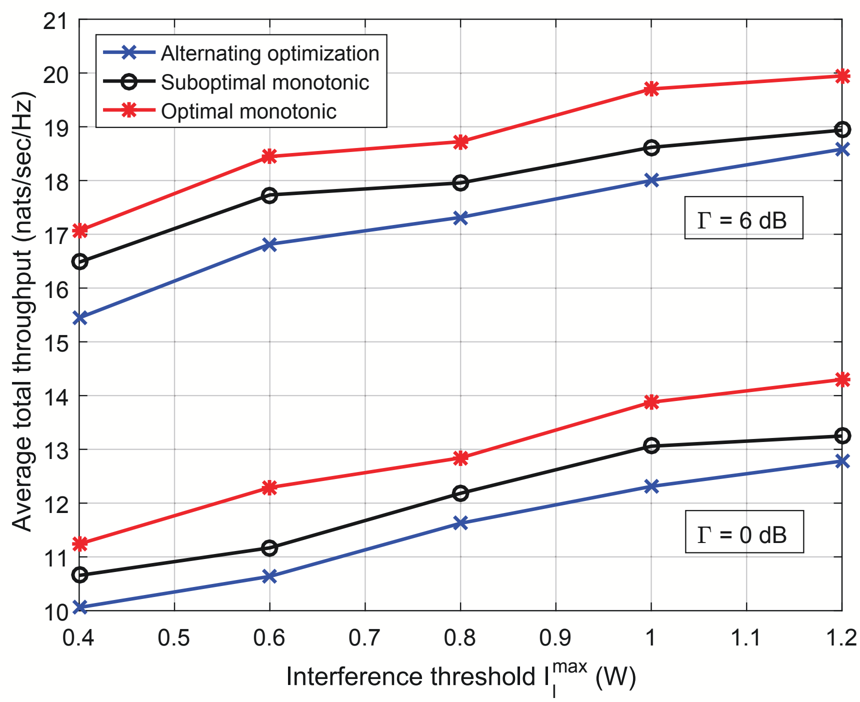


Figure 5.5: Average total throughput of SUs in terms of the interference threshold and the sensing SNR ($N = 6$).

5.6.2 A Large-Dimension CR Network

A simulation parameter setup similar to the one in Subsubsection 5.6.1 is implemented for a large-dimension CR network with $N = 40$. Specifically, $L = 4$, $K = 10$ and the licensed channels for four PUs are 1-10, 11-20, 21-30 and 31-40, respectively. Moreover, $P_n^1 = 0.2,\ n = 1, \cdots, N$, $M = 10$, $\sigma_s^2 = 10W$ and $\sigma_\nu^2 = 1$W. The channel power gains $g_{k,n}^{SS}$ and $g_{l,n}^{SP}$ are assumed to be exponentially distributed with mean 5dB.

In Figure 5.8, the average total SU throughput is simulated with a power limit P_n^T ranging from 10W to 30W for $n = 1, \cdots, N$ and an interference limit $I_l^{\max} = 5$W, $l = 1, \cdots, L$. Similar to Subsubsection 5.6.1, we consider the scenarios of $\Gamma = 0$dB and 6dB. As expected, the increase of power limit results in a higher average throughput for the secondary network. It is also found that the proposed suboptimal algorithm is

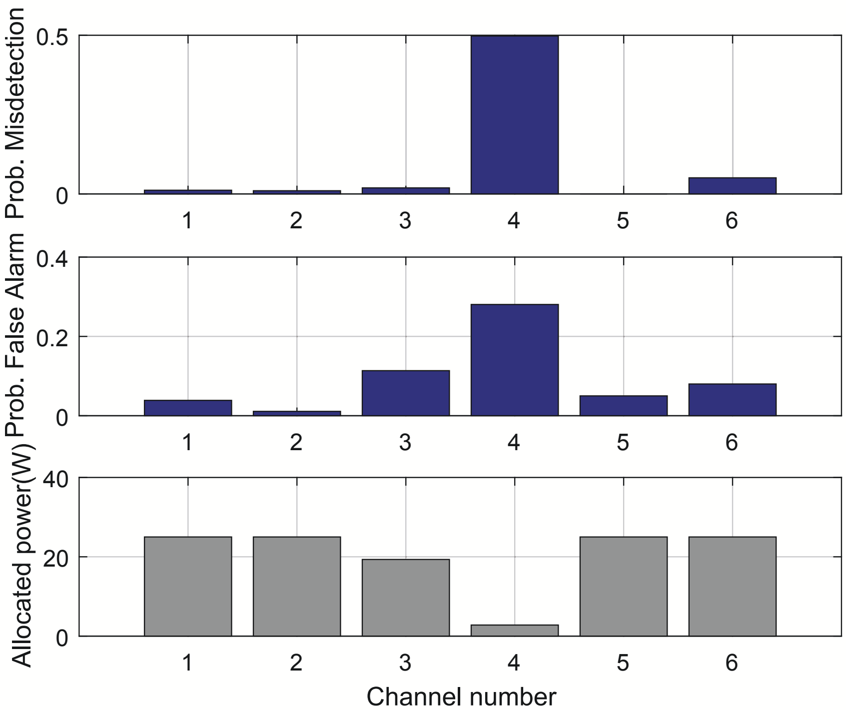


Figure 5.6: The probabilities of misdetection and false alarm, and the allocated power on individual channels.

comparable to the AO method with the enhanced-suboptimal scheme outperforming the others. The same trend can be found in Figure 5.9 with $P_n^T = 25\text{W}$ and $I_l^{\max}$ ranging from 1W to 5W. The convergence speed of the suboptimal-enhanced scheme is shown in Figure 5.10 with parameters $P_n^T = 25\text{W}$ and $I_l^{\max} = 5\text{W}$. It can be seen that these two independent simulations converge only in 3 iterations. More simulation results reveal that with the suboptimal monotonic solution as the starting point, the AO method generally converges in less than 5 iterations.

5.6.3 Discussions

Based on the numerical results presented in this subsection, it can be claimed that the proposed optimal and suboptimal monotonic optimization algorithms outperform the AO

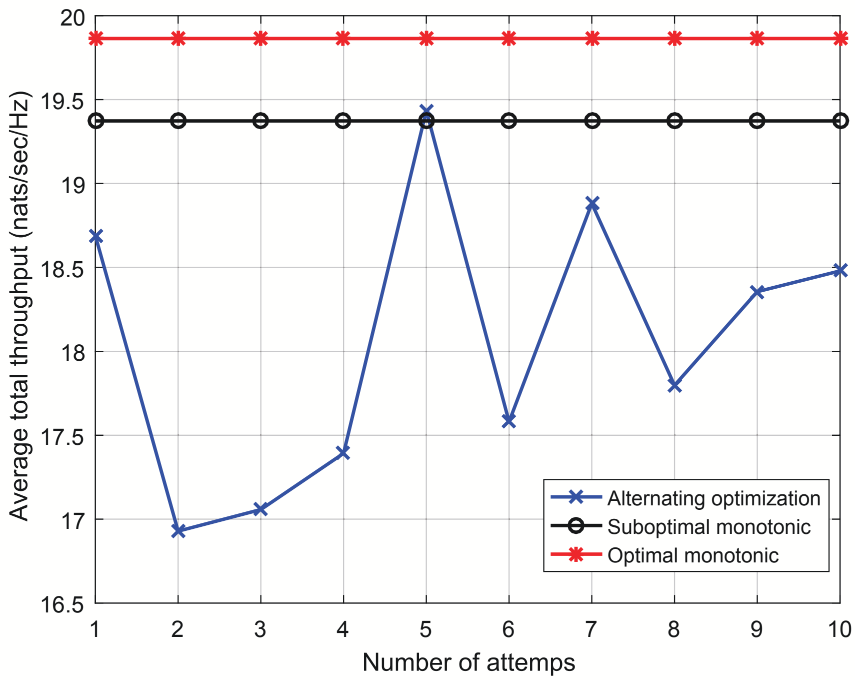


Figure 5.7: Stability of algorithms.

method in terms of accuracy and stability in the presence of a small-dimension CR network. For a large-dimension CR network, the performance of the suboptimal algorithm is comparable to the AO method. Using the solution provided by the suboptimal monotonic algorithm as the starting point, the suboptimal-enhanced algorithm achieves a higher objective value than the AO method. Moreover, it is also shown that the suboptimal-enhanced scheme converges in only a few iterations.

In terms of computational complexity, the optimal monotonic optimization algorithm is generally applicable for a problem whose size is less than 10. Such a framework is typically used as a benchmark for other heuristic algorithms or employed in a CR network with a very small number of channels. The state-of-the-art AO method is an iterative optimization method where spectrum sensing and resource allocation are optimized in turns. However, in each turn, the optimization problem presents a variable size same to

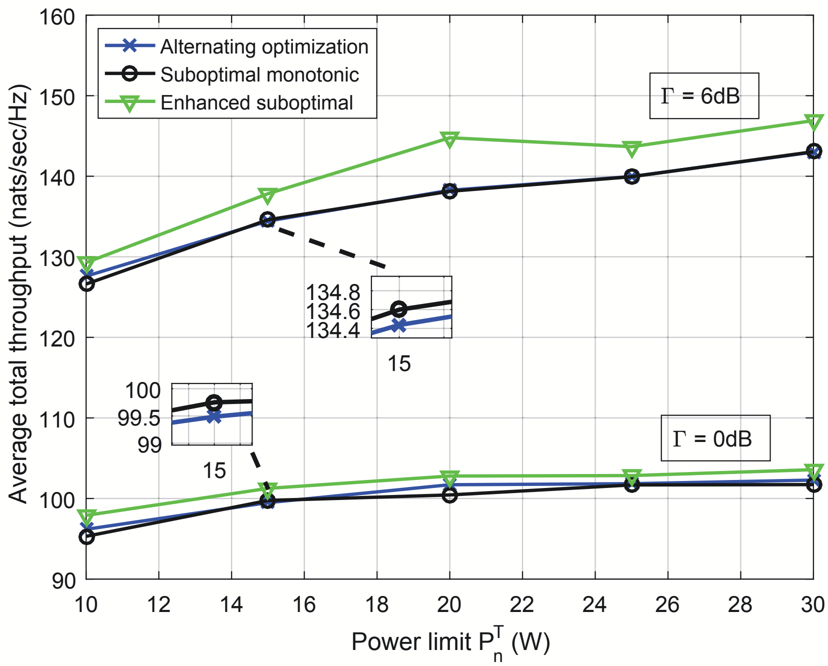


Figure 5.8: Average total throughput of SUs in terms of the power limit and the sensing SNR ($N = 40$).

the number of channels in the CR network, which results in a high computational burden in each iteration. Finally, the suboptimal algorithm, being a list of single-variable exhaustive search problems, or equivalently, a two-dimension monotonic optimization problem, generally exhibits the lowest computational complexity.

In a nutshell, for a small-dimension CR network, the suboptimal monotonic algorithm yields a near-optimal solution and converges very fast. If the computational time is not considered significant, the optimal monotonic algorithm can also be chosen. In a large-dimension CR network, the proposed suboptimal monotonic algorithm achieves a reasonable performance with a low computational complexity. The suboptimal-enhanced scheme can also be employed to further improve the performance.

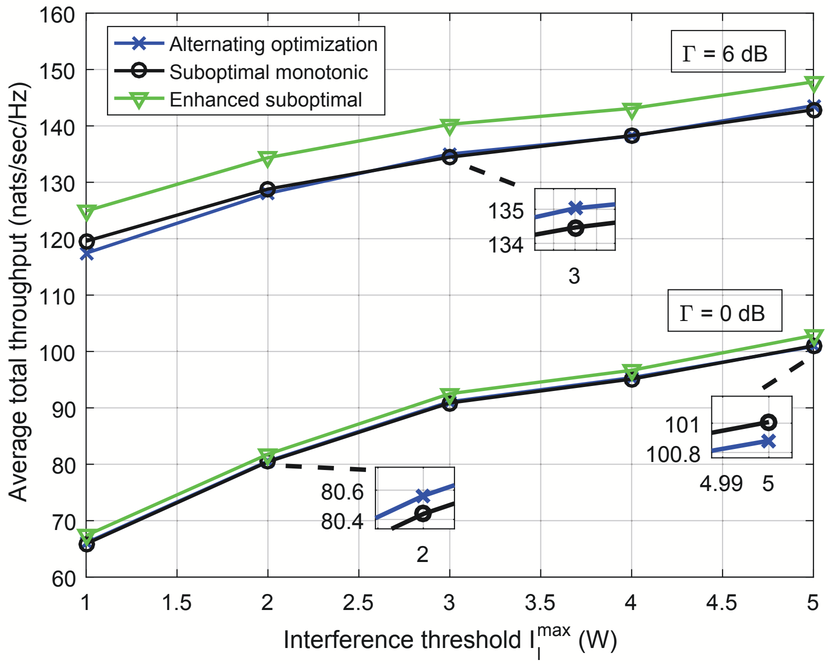


Figure 5.9: Average total throughput of SUs in terms of the interference threshold and the sensing SNR ($N = 40$).

5.7 Conclusions

This section studied the problem of designing the sensing threshold and power allocation strategy that maximize the average throughput of a multi-channel-multi-user CR network. The average interference and the peak power constraints, together with the sensing error constraint, are imposed to seek for the best balance between the protection of PUs and the throughput of CR network. Mathematically, the joint spectrum sensing and resource allocation problem is formulated as a non-convex MINLP optimization problem. The convex optimization technique is first employed to simplify the MINLP problem, which results in an equivalent formulation without any integer variables. Based on the knowledge of a monotonic optimization, we propose an algorithm to acquire the optimal sensing threshold and power allocation strategy for a CR network with a small number of

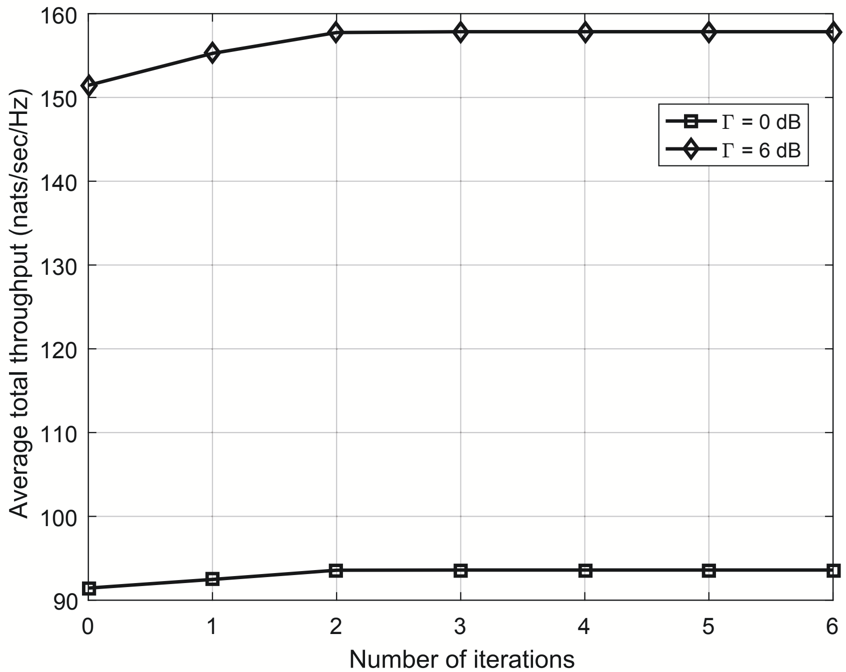


Figure 5.10: Convergence of the suboptimal-enhanced algorithm.

channels. In addition, a low-complexity suboptimal algorithm is developed to reduce the computational cost, and it is applicable to both small and large dimension CR networks. Numerical results indicate that the proposed optimal and suboptimal algorithms consistently outperform the start-of-the-art AO method in a CR network with a small number of channels. For a large-dimension CR network, the suboptimal algorithm also represents a good choice due to its low computational complexity.

6. CONCLUSIONS AND FUTURE DIRECTIONS

In this dissertation, the resource allocation problem was investigated in an interweave multi-channel-multi-user CR network where SUs can access opportunistically the licensed spectrum detected to be idle. Two powerful optimization techniques, namely the convex optimization and monotonic optimization, were employed to obtain the global optimal solutions.

First, some preliminaries about convex optimization and monotonic optimization were introduced. Since convex optimization techniques have been widely used in different kinds of areas for decades, only the essential concepts and theorems were given in Section 2 without any proof. On the other hand, in Section 3, the concept of monotonic optimization was introduced together with some basic features of monotonic optimization including increasing functions, normal sets, conormal sets, upper boundary, polyblocks, etc. As an under-explored research topic, monotonic optimization represents a powerful tool to achieve an optimal solution for non-convex optimization problems. In particular, by exploring the monotonic or hidden monotonic property of the problem, monotonic optimization enables to accelerate the searching speed for the optimal solution. Monotonic optimization problems can be addressed by the polyblock outer approximation method, which is analogous to the polyhedral outer approximation method for convex optimization problems.

The aforementioned optimization techniques were then exploited for the resource allocation problem in CR networks. In Section 4, the OFDMA signaling scheme was considered in a downlink interweave CR network. By exploring the special structure of the formulated MINLP problem using convex optimization techniques, we showed that the optimal subchannel and power allocation strategy, which aims at maximizing the total throughput of SUs, can be obtained using any standard convex optimization algorithm.

In Section 5, by taking into account the spectrum sensing errors, a joint spectrum sensing and resource allocation problem was studied in a multi-channel-multi-user interweave CR network. The joint optimization problem, with additional spectrum sensing variables, can be regarded as an extension of the resource allocation problem in Section 4 with an extra level of difficulty. Monotonic optimization techniques were implemented to address the problem, which represents an optimal solution. In addition, a low-complexity algorithm was also developed to reduce the computational cost.

Some of the potential future research directions that require further investigations include:

- For the joint optimization problem in Section 5, the sensing data is only collected and measured at the SBS. However, fading, shadowing and hidden primary user issues may dramatically affect the sensing accuracy in the case of a single sensing point. The concept of cooperative sensing was recently reported to collect data from different sensing points and make the decision by comprehensively considering all the sensing data collected [64, 65]. In this regard, our work of joint optimization might be extended by adopting the cooperative sensing technique. However, the resulting problem is more complicated since the false alarm and misdetection rates may admit a very complicated expression.

- Only interweave CR networks were studied in this dissertation by assuming the opportunistic access of SUs. Our work might be generalized to both underlay and overlay CR networks, where SUs are allowed to coexist with PUs in the same channel by ensuring that no harmful interference is caused to PUs. In these scenarios, the total throughput of SUs should increase since more channels are available to SUs.

- In this dissertation, the perfect CSI is assumed. However, in practical wireless systems, it is very hard to obtain perfect information about channel gains. Therefore, the imperfectness of CSI should be considered in our future research works.

- In both the resource allocation and joint optimization problems, we focused on instantaneous optimization. In other words, the network resource scheduling and decision making are based on the instantaneous channel gains obtained. In this way, the algorithms need to be calculated very frequently to follow up the rapid change of channel gains in CR networks. Such fast adaption may cause high computational burden and excessive signaling overhead. A slow adaptive network scheduling scheme based on long-term channel state parameters represents an important future direction for our study of CR networks. In this case, the channel coefficients are random rather than deterministic and the problem can be formulated as a stochastic optimization problem. The reader can refer to [66] for an existing research work which focused on optimizing the throughput of an OFDMA system stochastically.

- From a mathematical viewpoint, the theory of monotonic optimization as well as its applications in communications, signal processing and networking, is still at its infancy stage. Moreover, monotonic optimization with a large number of variables is still computationally complex. Developing more efficient algorithms for monotonic optimization and implementing them into practical communications systems represent important open research problems.

REFERENCES

[1] M. A. McHenry, P. A. Tenhula, D. McCloskey, D. A. Roberson, and C. S. Hood, "Chicago spectrum occupancy measurements & analysis and a long-term studies proposal," in *Proceedings of the first international workshop on Technology and policy for accessing spectrum*, p. 1, ACM, 2006.

[2] Federal Communications Commission, "Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies," *Et Docket*, no. 3-108, pp. 5–57, 2003.

[3] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, 1999.

[4] Federal Communications Commission, "Notice of proposed rule making and order: Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies," *ET Docket*, no. 03-108, p. 73, 2005.

[5] E. Biglieri, "An overview of cognitive radio for satellite communications," in *Satellite Telecommunications (ESTEL), 2012 IEEE First AESS European Conference on*, pp. 1–3, IEEE, 2012.

[6] H. A. Mahmoud, T. Yucek, and H. Arslan, "OFDM for cognitive radio: Merits and challenges," *IEEE Wireless Communications*, vol. 16, no. 2, pp. 6–14, 2009.

[7] IEEE Computer Society LAN MAN Standards Committee *et al.*, "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," *IEEE Standard 802.11-1997*, 1997.

[8] IEEE 802.16 Working Group *et al.*, "IEEE standard for local and metropolitan area networks. part 16: Air interface for fixed broadband wireless access systems," *IEEE*

Std, vol. 802, pp. 16–2004, 2004.

[9] N. Han, S. Shon, J. H. Chung, and J. M. Kim, "Spectral correlation based signal detection method for spectrum sensing in IEEE 802.22 WRAN systems," in *Advanced Communication Technology, 2006 IEEE International Conference on*, vol. 3, pp. 6–pp, IEEE, 2006.

[10] J. Zyren and W. McCoy, "Overview of the 3GPP long term evolution physical layer," *Freescale Semiconductor, Inc., white paper*, 2007.

[11] G. Bansal, M. J. Hossain, and V. K. Bhargava, "Optimal and suboptimal power allocation schemes for OFDM-based cognitive radio systems," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4710–4718, 2008.

[12] S. Wang, "Efficient resource allocation algorithm for cognitive OFDM systems," *IEEE Communications Letters*, vol. 14, no. 8, pp. 725–727, 2010.

[13] Y. Zhang and C. Leung, "Resource allocation in an OFDM-based cognitive radio system," *IEEE Transactions on Communications*, vol. 57, no. 7, pp. 1928–1931, 2009.

[14] M. Ge and S. Wang, "Fast optimal resource allocation is possible for multiuser OFDM-based cognitive radio networks with heterogeneous services," *IEEE Transactions on Wireless Communications*, vol. 11, no. 4, pp. 1500–1509, 2012.

[15] S. Wang, M. Ge, and W. Zhao, "Energy-efficient resource allocation for OFDM-based cognitive radio networks," *IEEE Transactions on Communications*, vol. 61, no. 8, pp. 3181–3191, 2013.

[16] S. Wang, Z.-H. Zhou, M. Ge, and C. Wang, "Resource allocation for heterogeneous cognitive radio networks with imperfect spectrum sensing," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 3, pp. 464–475, 2013.

[17] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 116–130, 2009.

[18] Y.-C. Liang, Y. Zeng, E. C. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1326–1337, 2008.

[19] E. C. Y. Peh, Y.-C. Liang, Y. L. Guan, and Y. Zeng, "Optimization of cooperative sensing in cognitive radio networks: A sensing-throughput tradeoff view," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 9, pp. 5294–5299, 2009.

[20] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 1128–1140, 2009.

[21] H. Zhang, C. Jiang, N. C. Beaulieu, X. Chu, X. Wen, and M. Tao, "Resource allocation in spectrum-sharing OFDMA femtocells with heterogeneous services," *IEEE Transactions on Communications*, vol. 62, no. 7, pp. 2366–2377, 2014.

[22] H. Zhang, C. Jiang, X. Mao, and H.-H. Chen, "Interference-limited resource optimization in cognitive femtocells with fairness and imperfect spectrum sensing," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1761–1771, 2016.

[23] H. Yu, W. Tang, and S. Li, "Joint optimal sensing and power allocation for cooperative relay in cognitive radio networks," in *Communications, 2012 IEEE International Conference on*, pp. 1635–1640, IEEE, 2012.

[24] N. Janatian, S. Sun, and M. Modarres-Hashemi, "Joint optimal spectrum sensing and power allocation in CDMA-based cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 9, pp. 3990–3998, 2015.

[25] X. Huang and B. Beferull-Lozano, "Joint optimization of detection and power allocation for OFDM-based cognitive radios," in *GLOBECOM, 2010 IEEE*, pp. 1–5, IEEE, 2010.

[26] R. Fan, H. Jiang, Q. Guo, and Z. Zhang, "Joint optimal cooperative sensing and resource allocation in multichannel cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 2, pp. 722–729, 2011.

[27] J. C. Bezdek and R. J. Hathaway, "Some notes on alternating optimization," in *AFSS International Conference on Fuzzy Systems*, pp. 288–300, Springer, 2002.

[28] W. Yu and J. M. Cioffi, "FDMA capacity of gaussian multiple-access channels with ISI," *IEEE Transactions on Communications*, vol. 50, no. 1, pp. 102–111, 2002.

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

[30] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994.

[31] Z. John Lu, "The elements of statistical learning: Data mining, inference, and prediction," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 173, no. 3, pp. 693–694, 2010.

[32] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific Belmont, 1999.

[33] J. R. Banga, "Optimization in computational systems biology," *BMC Systems Biology*, vol. 2, no. 1, p. 47, 2008.

[34] H. Hindi, "A tutorial on convex optimization," in *Proceedings of the 23th American Control Conference*, pp. 3252–3265, IEEE, 2004.

[35] H. Hindi, "A tutorial on convex optimization ii: Duality and interior point methods," in *American Control Conference, 2006*, pp. 11–pp, IEEE, 2006.

[36] S. Burer and A. N. Letchford, "Non-convex mixed-integer nonlinear programming: A survey," *Surveys in Operations Research and Management Science*, vol. 17, no. 2, pp. 97–106, 2012.

[37] O. K. Gupta and A. Ravindran, "Branch and bound experiments in convex nonlinear integer programming," *Management Science*, vol. 31, no. 12, pp. 1533–1546, 1985.

[38] A. M. Geoffrion, "Generalized benders decomposition," *Journal of Optimization Theory and Applications*, vol. 10, no. 4, pp. 237–260, 1972.

[39] M. A. Duran and I. E. Grossmann, "An outer-approximation algorithm for a class of mixed-integer nonlinear programs," *Mathematical Programming*, vol. 36, no. 3, pp. 307–339, 1986.

[40] R. A. Stubbs and S. Mehrotra, "A branch-and-cut method for 0-1 mixed convex programming," *Mathematical Programming*, vol. 86, no. 3, pp. 515–532, 1999.

[41] T. Westerlund and F. Pettersson, "An extended cutting plane method for solving convex minlp problems," *Computers & Chemical Engineering*, vol. 19, pp. 131–136, 1995.

[42] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 2012.

[43] H. Tuy, "Monotonic optimization: Problems and solution approaches," *SIAM Journal on Optimization*, vol. 11, no. 2, pp. 464–494, 2000.

[44] H. Tuy, F. Al-Khayyal, and P. T. Thach, "Monotonic optimization: Branch and cut methods," in *Essays and Surveys in Global Optimization*, pp. 39–78, Springer, 2005.

[45] H. Tuy and N. Hoai-Phuong, "Optimization under composite monotonic constraints and constrained optimization over the efficient set," in *Global Optimization*, pp. 3–31, Springer, 2006.

[46] H. Tuy, M. Minoux, and N. Hoai-Phuong, "Discrete monotonic optimization with application to a discrete location problem," *SIAM Journal on Optimization*, vol. 17, no. 1, pp. 78–97, 2006.

[47] L. P. Qian, Y. J. Zhang, and J. Huang, "MAPEL: Achieving global optimality for a non-convex wireless power control problem," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1553–1563, 2009.

[48] L. P. Qian and Y. Zhang, "Monotonic optimization for non-concave power control in multiuser multicarrier network systems," in *INFOCOM 2009, IEEE*, pp. 172–180, IEEE, 2009.

[49] E. A. Jorswieck and E. G. Larsson, "Monotonic optimization framework for the two-user MISO interference channel," *IEEE Transactions on Communications*, vol. 58, no. 7, pp. 2159–2168, 2010.

[50] C. Xing, S. Ma, and Y. Zhou, "Matrix-monotonic optimization for MIMO systems," *IEEE Transactions on Signal Processing*, vol. 63, no. 2, pp. 334–348, 2015.

[51] L. P. Qian and Y. J. Zhang, "S-MAPEL: Monotonic optimization for non-convex joint power control and scheduling problems," *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1708–1719, 2010.

[52] Y. J. A. Zhang, L. Qian, and J. Huang, "Monotonic optimization in communication and networking systems," *Foundations and Trends in Networking*, vol. 7, no. 1, pp. 1–75, 2013.

[53] R. Horst, N. V. Thoai, and H. Tuy, "Outer approximation by polyhedral convex sets," *OR Spectrum*, vol. 9, no. 3, pp. 153–159, 1987.

[54] X. Wang, S. Ekin, and E. Serpedin, "Optimal resource allocation for downlink OFDM-based cognitive radio networks," in *Networks, Computers and Communica-*

tions, 2017 IEEE International Symposium on, IEEE, 2017. Accepted for publication.

[55] S. M. Almalfouh and G. L. Stuber, "Interference-aware radio resource allocation in OFDMA-based cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 4, pp. 1699–1713, 2011.

[56] P. Setoodeh and S. Haykin, "Robust transmit power control for cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 915–939, 2009.

[57] C. Y. Wong, R. S. Cheng, K. B. Lataief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 10, pp. 1747–1758, 1999.

[58] Z. Shen, J. G. Andrews, and B. L. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Transactions on Wireless Communications*, vol. 4, no. 6, pp. 2726–2737, 2005.

[59] M. Tao, Y.-C. Liang, and F. Zhang, "Resource allocation for delay differentiated traffic in multiuser OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, 2008.

[60] A. Akbari, R. Hoshyar, and R. Tafazolli, "Energy-efficient resource allocation in wireless OFDMA systems," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*, pp. 1731–1735, IEEE, 2010.

[61] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in multiuser OFDM systems with wireless information and power transfer," in *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*, pp. 3823–3828, IEEE, 2013.

[62] S. Boyd, L. Xiao, and A. Mutapcic, "Subgradient methods," *Lecture Notes of EE392o, Stanford University, Autumn Quarter*, 2003.

[63] X. Wang, S. Ekin, and E. Serpedin, "Joint spectrum sensing and resource allocation in multi-band-multi-user cognitive radio networks," *IEEE Transactions on Communications*, 2017. Submitted.

[64] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio, part I: Two user networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 6, pp. 2204–2213, 2007.

[65] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio, part II: multiuser networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 6, pp. 2214–2222, 2007.

[66] W. W.-L. Li, Y. J. Zhang, A. M.-C. So, and M. Z. Win, "Slow adaptive OFDMA systems through chance constrained programming," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3858–3869, 2010.

APPENDIX A

PROOF OF THE CONVEXITY FOR PROBLEM (P_2) IN SECTION 4

The continuous relaxation of problem (P_1^E) is represented as problem (P_2) in Subsection 4.3. Suppose

$$f_{k,n}(x_{k,n}, y_{k,n}) = x_{k,n} \log(1 + \frac{s_{k,n} y_{k,n}}{x_{k,n}}),$$

it can be claimed that $f_{k,n}$ is a concave function with respect to $x_{k,n}$ and $y_{k,n}$. First, it can be checked that $f_{k,n}$ is the perspective of the logarithm function $\log(1 + s_{k,n} y_{k,n})$. Since the logarithm function $\log(1 + s_{k,n} y_{k,n})$ is concave with respect to $y_{k,n}$. Due to Definition 4, the perspective $f_{k,n}$ preserves the concavity.

Because $f_{k,n}$ is a concave function for any k and n, the objective function of (P_2), being the summation of concave functions, is a concave function with respect to $x_{k,n}$ and $y_{k,n}$. Therefore, it can be concluded that the optimization problem (P_2) is a convex optimization problem due to the fact that all its constraints are affine.

APPENDIX B

PROOF OF LEMMAS AND THEOREMS IN SECTION 4

B.1 Proof of Lemma 1

On one hand, if $\eta_n < \max_k J_{k,n}$, for $k_n^* = \arg\max_k J_{k,n}$, it follows that $J_{k_n^*,n} - \eta_n > 0$, which violates (4.8). On the other hand, if $\eta_n > \max_k J_{k,n}$, it yields that $J_{k,n} - \eta_n < 0, \forall k$. Based on (4.8), it can be inferred that $x_{k,n} = 0, \forall k$, i.e., no SU occupies the nth subchannel. From (4.11), we have $\eta_n = 0$ and thus $\max_k J_{k,n} < 0$. However, this statement is contradictory to the formulation of $J_{k,n}$ in (4.7). Specifically, $J_{k,n}$ can be regarded as a function $J(w_{k,n}) = \log(1 + w_{k,n}) - w_{k,n}/(1 + w_{k,n})$ with $w_{k,n} = s_{k,n} \left(\frac{1}{\lambda + \sum_{l=1}^{L} \mu_l I_{l,n}^{SP}} - \frac{1}{s_{k,n}} \right)^+$. The function is a monotonically increasing function for $w_{k,n} \geq 0$, and its minimum value 0 is achieved at $w_{k,n} = 0$. Considering both sides, it yields that $\eta_n = \max_k J_{k,n}$, which completes the proof.

B.2 Proof of Theorem 4

If $\max_k J_{k,n}$ is achieved by only one SU, say the k_n^*th SU, from Lemma 1, we have

$$J_{k_n^*,n} - \eta_n = 0 \text{ and } J_{k,n} - \eta_n < 0,\ k \neq k_n^*.$$

From (4.8), it follows that $x_{k_n^*,n} > 0$ and $x_{k,n} = 0,\ k \neq k_n^*$. Moreover, since $J_{k_n^*,n} > J_{k,n},\ k \neq k_n^*$ and $J_{k,n} \geq 0$, it follows that $\eta_n = J_{k_n^*,n} > 0$ and furthermore $x_{k_n^*,n} = 1$ due to (4.11). Since KKT conditions are necessary and sufficient conditions for problem (P_2), $x_{k_n^*,n} = 1$ and $x_{k,n} = 0,\ k \neq k_n^*$ is the only solution satisfying the KKT conditions (4.5), (4.8) and (4.11), we can claim this is the optimal solution. However, if $\max_k J_{k,n}$ is tied for multiple SUs, these SUs all achieve positive $x_{k,n}$ due to (4.8), and the nth subchannel must be shared among these SUs. The complementary slackness conditions (4.9) and (4.10) also have to be incorporated to decide how to share the subchannel.

B.3 Proof of Theorem 5

Based on Theorem 4, each subchannel n must be allocated to at least one SU, i.e., there exists at least one positive $x_{k,n}$ for each n. Positive $x_{k,n}$ leads to positive $y_{k,n}$ (otherwise we can just set $x_{k,n} = 0$), and thus the set $\{k \mid 1/(\lambda + \sum_{l=1}^{L} \mu_l I_{l,n}^{SP}) - 1/s_{k,n} > 0\}$ is not empty due to (5.16). Moreover, the maximum of $J_{k,n}$ must be achieved among the SUs in this set. In this way, the $(\cdot)^+$ operator can be removed and $J_{k,n}$ in (4.7) can be simplified as

$$J_{k,n} = \log s_{k,n} + \frac{\lambda + \sum_{l=1}^{L} \mu_l I_{l,n}^{SP}}{s_{k,n}} - \log(\lambda + \sum_{l=1}^{L} \mu_l I_{l,n}^{SP}) - 1.$$

It turns out that

$$\frac{\partial J_{k,n}}{\partial s_{k,n}} = \frac{s_{k,n} - (\lambda + \sum_{l=1}^{L} \mu_l I_{l,n}^{SP})}{s_{k,n}^2},$$

and $J_{k,n}$ is a monotonically increasing function of $s_{k,n}$ in the set $\{k \mid 1/(\lambda + \sum_{l=1}^{L} \mu_l I_{l,n}^{SP}) - 1/s_{k,n} > 0\}$. Therefore, if $s_{k,n}$'s are all distinct, according to Theorem 4, the nth subchannel is assigned to the k_n^*th SU with $k_n^* = \arg\max_k s_{k,n}$.

APPENDIX C

PROOFS OF LEMMAS AND THEOREMS IN SECTION 5

C.1 Proof of Lemma 2

Assume that the nth channel is not occupied by any SU, i.e., $x_{k,n} = 0, \forall k$. Due to (5.17), it follows that $J_{k,n} < \eta_n, \forall k$. It can be observed that $J_{k,n}$ is a monotonically increasing function with respect to $(\psi_n - \sigma_\nu^2 / g_{k,n}^{SS})^+$ with its minimum value 0 attained at $(\psi_n - \sigma_\nu^2 / g_{k,n}^{SS})^+ = 0$. Thus, $\eta_n > 0$. Based on this fact and the complementary slackness condition

$$\eta_n \left(1 - \sum_{k=1}^{K} x_{k,n} \right) = 0, \tag{C.1}$$

it turns out that $\sum_{k=1}^{K} x_{k,n} = 1$, which contradicts the original assumption that $x_{k,n} = 0, \forall k$. Therefore, it can be claimed that all the channels must be occupied by SUs. Furthermore, the statement $J_{k,n} < \eta_n, \forall k$, does not hold, which implies that $\max_k J_{k,n} \geq \eta_n$. From (5.17), we have $\max_k J_{k,n} \leq \eta_n$, and thus $\eta_n = \max_k J_{k,n}$, which completes the proof.

C.2 Proof of Theorem 6

The theorem has to be proved in two scenarios. For a channel n, on one hand, it might occur that $\psi_n - \sigma_\nu^2 / g_{k,n}^{SS} \leq 0, \forall k$. In this case, $y_{k,n} = 0,\ \forall k$, and we can simply assign the channel to the SU with the largest $g_{k,n}^{SS}$ without affecting the objective value. On the other hand, if the set $\{k | \psi_n - \sigma_\nu^2 / g_{k,n}^{SS} > 0\}$ is not empty, $\max_k J_{k,n}$ must be achieved among SUs in this set. In this way, the $(\cdot)^+$ can be removed and it follows that $J_{k,n}$ is a monotonically increasing function of $g_{k,n}^{SS}$ in the set $\{k | \psi_n - \sigma_\nu^2 / g_{k,n}^{SS} > 0\}$. Thus, $\max_k J_{k,n}$ is solely achieved at the SU with the largest $g_{k,n}^{SS}$, say the k_n^*th SU, provided that $g_{k,n}^{SS}$'s are all distinct. Then, based on Lemma 2, $\eta = \max_k J_{k,n} > 0$, and it follows that $x_{k_n^*,n} > 0$ and $x_{k,n} = 0,\ k \neq k_n^*$ due to (5.17). Finally, the complementary slackness condition in (C.1) yields the result that $x_{k_n^*,n} = 1$.

APPENDIX D

TRANSFORMATION OF PROBLEM (5.22) in SECTION 5

To transform problem (5.22) into the canonical form of a monotonic optimization, notice that for every $\{\gamma_n | M\sigma_\nu^2 \leq \gamma_n \leq M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2)\}$ there exists a nonnegative parameter t such that $g(\gamma_n) + t = g(M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2))$. In this way, problem (5.22) can be rewritten as

$$\begin{aligned} \max_{\gamma_n} \ & f(\gamma_n) + t \\ \text{s.t.} \ & M\sigma_\nu^2 \leq \gamma_n \leq M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2) \\ & g(\gamma_n) + t = g(M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2)). \end{aligned}$$

Let $F(\gamma_n, t) = f(\gamma_n) + t$, the formulation above admits the canonical form

$$\begin{aligned} \max \ & F(\gamma_n, t) \\ \text{s.t.} \ & (\gamma_n, t) \in \mathcal{G} \cap \mathcal{H}, \end{aligned}$$

with $\mathcal{G}$ and $\mathcal{H}$ defined in (5.24) and (5.25), respectively. It is easy to check that in the region

$$\left[\mathbf{0}, [M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2), g(M(\sigma_\nu^2 + g_n^{PS}\sigma_s^2)) - g(M\sigma_\nu^2)]\right],$$

$F(\gamma_n, t)$ is an increasing function, $\mathcal{G}$ and $\mathcal{H}$ are the normal and conormal set, respectively.

APPENDIX E

ALTERNATING OPTIMIZATION METHOD FOR JOINT SPECTRUM SENSING AND RESOURCE ALLOCATION PROBLEM

In this appendix, the state-of-the-art AO method for solving the joint spectrum sensing and resource allocation in CR networks is described as a benchmark for the proposed optimal and suboptimal monotonic optimization methods in Section 5. We focus on applying the AO method to address P-JointE, which is equivalent to the original problem P-Joint.

For the AO method, a random but feasible guess of $\gamma_n^{(0)}$ is selected at the initialization step. Given a fixed value of γ_n, the first step is to calculate the optimal power allocation strategy $p_{k_n^*,n}$, which is provided by "Alternating optimization - power (AO-P)":

$$\begin{aligned}
&\text{AO-P :} \\
&\max_{p_{k_n^*,n}} \ \sum_{n=1}^{N} P_n^0(1-P_n^F)\log\left(1+\frac{g_{k_n^*,n}^{SS}p_{k_n^*,n}}{\sigma_\nu^2}\right) \\
&\quad \text{s.t.} \ \ (C_1),(C_2)
\end{aligned}$$

AO-P can be implemented via any standard convex optimization technique, such as the interior point method [29]. However, by re-formulating and analyzing the problem in the dual domain, it is shown that AO-P presents a closed-form solution for the power allocation variables $p_{k_n^*,n}$.

Considering the dual variables $\mu_l,\ l=1,\cdots,L$ for the constraint (C_1) leads to the

Lagrangian:

$$Lg(p_{k_n^*,n},\mu_l) = \sum_{n=1}^{N} P_n^0(1-P_n^F)\log\left(1+\frac{g_{k,n}^{SS}p_{k_n^*,n}}{\sigma_\nu^2}\right) + \sum_{l=1}^{l}\mu_l\left(I_l^{\max} - \sum_{n\in S_l} P_n^1(1-P_n^D)p_{k_n^*,n}g_{l,n}^{SP}\right),$$

where constraint (C_2) is absorbed in KKT conditions. The KKT conditions yield that

$$\frac{\partial Lg}{\partial p_{k_n^*,n}} = P_n^0(1-P_n^F)\frac{g_{k,n}^{SS}}{\sigma_\nu^2 + g_{k,n}^{SS}p_{k_n^*,n}} - \mu_l P_n^1(1-P_n^D)g_{l,n}^{SP}\begin{cases} <0, & p_{k_n^*,n}=0,\\ =0, & 0<p_{k_n^*,n}<P_n^T,\\ >0, & p_{k_n^*,n}>0.\end{cases}$$

It turns out that the optimal power variables admit the expression:

$$p_{k_n^*,n} = \min\left\{\left(\frac{P_n^0(1-P_n^F)}{\mu_l P_n^1(1-P_n^D)g_{l,n}^{SP}} - \frac{\sigma_\nu^2}{g_{k,n}^{SS}}\right)^+, P_n^T\right\} \tag{E.1}$$

Setting the power variables $p_{k_n^*,n}$ to fixed values in (E.1), the next step is to find the optimal sensing threshold, termed as the “Alternating optimization - sensing threshold (AO-ST)” in this paper:

$$\begin{aligned}&\text{AO-ST}:\\ &\max_{\gamma_n}\ \sum_{n=1}^{N} P_n^0(1-P_n^F)\log\left(1+\frac{g_{k_n^*,n}^{SS}p_{k_n^*,n}}{\sigma_\nu^2}\right)\\ &\text{s.t.}\ (C_2),(C_3)\end{aligned}$$

As aforementioned, P_n^F and P_n^D are convex and concave functions with regard to γ_n, respectively. Consequently, AO-ST represents a convex optimization problem, which can be solved efficiently via the interior point method.

Combining these two steps, namely, AO-P and AO-ST, the proposed enhanced AO method is summarized in Algorithm 3, where ϵ represents the stopping criterion and assumes a sufficiently small positive number. It turns out that the objective value is a non-decreasing function with

$$f^{(i)} = f(\gamma_n^{(i)}, p_{k_n^*,n}^{(i)}) \leq f(\gamma_n^{(i)}, p_{k_n^*,n}^{(i+1)}) \leq f(\gamma_n^{(i+1)}, p_{k_n^*,n}^{(i+1)}) = f^{(i+1)}.$$

Since γ_n and $p_{k_n^*,n}$ are upper-bounded, the convergence is ensured.

Algorithm 3 AO method for P-JE

Initialization: $i = 0$, ϵ, and $\gamma_n^{(0)}$

1: **repeat**
2: Given $\gamma_n^{(0)}$, calculate $p_{k_n^*,n}^*$ that solves AO-P from (E.1).
3: $p_{k_n^*,n}^{(i+1)} = p_{k_n^*,n}^*$
4: Given $p_{k_n^*,n}^{(i+1)}$, find the solution of AO-ST: γ_n^*.
5: $\gamma_n^{(i+1)} = \gamma_n^*$
6: Calculate the objective value of P-JE in terms of $\gamma_n^{(i+1)}$ and $p_{k_n^*,n}^{(i+1)}$, and denote it as $f^{(i+1)}$.
7: $i = i + 1$
8: **until** $f^{(i)} - f^{(i-1)} \leq \epsilon$
