University of Alberta

Optimal Spectrum Sensing and Resource Allocation in Cognitive Radio

by

Rongfei Fan

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Communications

Department of Electrical and Computer Engineering

Fall 2012
Edmonton, Alberta

Abstract

Cognitive radio is an emerging spectrum-agile technology to alleviate the upcoming spectrum shortage problem. In a cognitive radio network, unlicensed users (secondary users) can access the licensed spectrum if there is no transmission/reception activity of licensed users (primary users) in an overlay mode, or the interference of secondary users to primary users is below a threshold in an underlay mode. Spectrum sensing is essential in an overlay mode, while resource allocation is challenging in both modes due to the lower priority of secondary channel access. The focus of this thesis is on optimal spectrum sensing and resource allocation in cognitive radio networks, to provide necessary protection for primary users and achieve resource efficiency for secondary users.

Firstly, optimal sensing time allocation in multichannel cognitive radio network is studied, to maximize the average throughput of secondary users while protecting primary activities. The initially formulated optimization problems are non-convex, which are very hard to be solved optimally. By finding special properties of the problems, the problems are decomposed into bi-level convex optimization problems, which can be solved optimally.

Secondly, channel sensing order setting in a two-user multichannel cognitive radio network is investigated. Two sub-optimal algorithms are proposed and numerically verified to have comparable performances to optimal solutions. When adaptive modulation is adopted, it is shown that the stopping rule should be designed jointly with sensing order setting strategy of the two users.

Thirdly, joint sensing time setting and resource allocation in a multichannel cognitive radio network is studied. An optimization problem is formulated to maximize the weighted average throughput of secondary users. The problem is non-convex. With the aid of monotonic optimization and bi-level optimization, the non-convex problem is solved optimally. The research is also extended to cases maximizing the proportional or max-min fairness level of the users.

Last but not least, optimal resource allocation in an underlay mode is investigated. The average rate of secondary users is maximized while limiting the interference to primary users. Convex problems are formulated. By deriving special properties of the optimal solutions, simple online algorithms are given, with closed-form solutions.

Acknowledgments

First I would like to express my sincerest gratitude to my supervisor, Dr. Hai Jiang, for his patient guidance, constant encouragement and continuous support. His enthusiasm and inspiration kept motivating and encouraging me in all times of my research while his professional experiences in teaching and supervising helped direct me to my current achievements in academic research. I also thank him for setting up an excellent example to me in the areas besides research. I always benefit from learning from his dedication, his generosity and his amiability.

I am also grateful to Dr. Yindi Jing and Dr. Lifeng Lai for their help in my research works. Their inspiring suggestions helped me improve the quality of my research works.

I would like to thank the members of my defense committee, Dr. Chintha Tellambura, Dr. F. Richard Yu, Dr. Xinwei Yu and Dr. H. Vicky Zhao, for their time spent reviewing my thesis and providing insightful comments on my dissertation.

I wish to thank the colleges in my lab including Jie Gao, Zhou Zhang, Lijie Huang, Qian Cao, Zhongjiang Yan, Hongxing Guo, Saman Atapattu, Qiang Guo, etc. for the stimulating discussions and the warm friendships. Many thanks also go to Shuai Zhang, Meng Shi, Gongpu Wang, Gencheng Guo, Yang Cheng, Wei Zhang, Hua Shao and many others for their kind-hearted helps in the past years. Many good memories of mine were marked with them during these years.

Most importantly, I thank my parents and grandparents for their understanding, supporting and encouragement, and most of all their ever-lasting love.

Last but not the least, I appreciate the financial support from Alberta Innovates Technology Futures during my program.

List of Symbols

k_n	The number of mini-slots for sensing channel n
$l(\cdot)$	The length of a vector
p_n	Transmission power of secondary BS on channel n
$p_{n,m}$	Transmission power of m^{th} secondary user on channel n
$r^0_{n,m}$	Transmission rate of m^{th} secondary user on channel n when channel n is idle and sensed to be idle
$r^1_{n,m}$	Transmission rate of m^{th} secondary user on channel n when channel n is busy but sensed to be idle
t^m_n	Time spent by m^{th} secondary user on sensing channel n
$(x)^+$	$\max(x, 0)$
$[x]^b_a$	$\max(\min(b, x), a)$
x_n	Bandwidth of channel n used by the secondary BS
$x_{n,m}$	Portion of the bandwidth of channel n assigned to m^{th} secondary user
B	Total bandwidth used by the secondary BS
$\mathrm{Diag}(\cdot)$	Diagonal matrix formed by all the elements of a vector
I^{LT}_n	Maximal long-term interference power on channel n
I^{PEAK}_n	Maximal peak PSD of interference power on channel n
I^{ST}_n	Maximal short-term interference power on channel n
M	The total number of secondary users
N	The total number of channels
P_d	Detection probability
P_f	False alarm probability
P^{avg}_m	Maximal average transmission power of m^{th} secondary user
P^{peak}_m	Maximal peak transmission power of m^{th} secondary user
P^d_n	Detection probability on channel n

P_n^f	False alarm probability on channel n
P_{th}	Threshold of detection probability
P^{LT}	Maximal long-term transmission power
P^{PEAK}	Maximal peak PSD of transmission power
P^{ST}	Maximal short-term transmission power
$Q(\cdot)$	Q function
$Q^{-1}(\cdot)$	Inverse Q function
Q_n^d	Detection probability on channel n with wideband spectrum sensing
Q_n^f	False alarm probability on channel n with wideband spectrum sensing
R_m	The average throughput of m^{th} secondary user
R_n^0	Average transmission rate on channel n when channel n is idle and sensed to be idle
R_n^1	Average transmission rate on channel n when channel n is busy but sensed to be idle
T	The time duration of one time slot or the test statistic
$(\cdot)^T$	Transpose
$\mathcal{H}^0$	Hypothesis that primary user is idle
$\mathcal{H}^1$	Hypothesis that primary user is busy
$\mathcal{H}_n^0$	Hypothesis that the primary user on channel n is idle
$\mathcal{H}_n^1$	Hypothesis that the primary user on channel n is busy
$\mathcal{I}$	The set of all positive integers
$\mathcal{M}$	The set of M secondary users
$\mathcal{N}$	The set of N channels
γ_n	Average SNR of primary user signal received by the secondary user on channel n
δ	The time duration of one mini-slot
ε	Detection threshold
ε_n	Detection threshold for sensing channel n
θ_n	Free probability of the primary user on channel n
θ_n^{free}	Free probability of channel n
μ	Sampling rate
τ	Time duration of the sensing phase in a time slot

τ_n	Time for sensing channel n

List of Abbreviations

Acronyms	Definition
1G	first-generation
2G	second-generation
3G	third-generation
4G	fourth-generation
i.i.d	independent and identically distributed
BS	base station
CAT	constant access time
CDMA	code division multiple access
CDT	constant data time
CSCG	circularly symmetric complex Gaussian
CSI	channel state information
DSSS	direct-sequence spread spectrum
FCC	Federal Communications Commission
FFT	fast Fourier transform
ISM	industrial, scientific and medical
ITU	International Telecommunication Union
KKT	Karush-Kuhn-Tucker
LTE Advanced	long-term-evolution advanced
MIMO	multiple-input multiple-output
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiple access
PDF	probability density function
QoS	quality of service

SCC	standards coordinating committee
SMS	short message service
SNR	signal-to-noise ratio
TDMA	time division multiple access
UWB	ultra-wideband
WAP	wireless application protocol
WiMAX	worldwide interoperability for microwave access
WLAN	wireless local area networks
WSN	wireless sensor networks

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Spectrum Scarcity and Under-utilization Problems

The past decades have witnessed the rapid growth of wireless communication networks. Among them, cellular networks and wireless local area networks (WLANs) are the two most popular wireless communication networks, and are involved in people's daily activities, including voice conversation, messaging, web browsing, emails, etc.

A cellular network is a radio network distributed over a wide geographic area, which is divided into a number of smaller areas, referred to as cells. In each cell there is a base station that provides wireless access service to mobile users within its coverage. A mobile user can receive seamless service when it moves from one cell to another cell, or from one cellular network to another cellular network. In other words, the cellular networks provide ubiquitous wireless access coverage to mobile users. The first-generation (1G) cellular networks are based on analog techniques, and can support only voice conversations [1]. The second-generation (2G) cellular networks adopt digital communication techniques, and use time division multiple access (TDMA) or code division multiple access (CDMA) [2]. Data service, such as short message service (SMS) and wireless application protocol (WAP) service, is also provided, with data rate ranging from 56 kbps to 236.8 kbps. So far the 2G networks are popularly deployed in the world. To improve the data transmission rate and to better support multimedia service, the third-generation (3G) networks are based on advanced CDMA technology and are expected to replace 2G networks [3]. As defined by International Telecommunication Union (ITU), the 3G networks can provide a minimum data rate of 2 Mbps for stationary or walking users and 384 kbps for users with high velocity (for example, in a moving vehicle) [4]. High-speed Internet access and data transmission are also supported in 3G networks. Nowadays researchers are still working on standardization of the four-generation (4G) cellular networks, for which many issues are still open. 4G networks are expected to offer a peak data rate of 1 Gbps for static or low-speed users and 100 Mbps for high-speed mobile users [5]. Long-term-evolution advanced (LTE Advanced) and Worldwide Interoperability for Microwave Access (WiMAX) are two major candidate standards [6].

Different from cellular networks that can support ubiquitous coverage, WLANs provide high-speed wireless access service via an access point for wireless devices (such as laptops and tablet computers) in a local area, such as campus, airport, hotel, office, etc. [7]. The IEEE 802.11 is one most popular standard family for WLANs. According to IEEE 802.11, a WLAN can take a bandwidth of more than 20MHz [8] and can be effective within the range of approximately 38 m for indoor service or 140 m for outdoor service. With a large bandwidth in a local area, it is possible for a WLAN to provide high data rates for its users. The IEEE 802.11b, an early version of IEEE 802.11 standards, adopts the direct-sequence spread spectrum (DSSS) modulation technique, and is able to provide a maximum rate of 11 Mbps for one user. The subsequent IEEE 802.11a and IEEE 802.11g adopt the orthogonal frequency division multiplexing (OFDM) technique in physical layer and can support a data rate up to 54 Mbps. The most latest IEEE 802.11 standard, IEEE 802.11n, adopts OFDM and multiple input multiple output (MIMO) techniques, and can support a data rate up to 150 Mbps [9]. It can be seen that the users in a WLAN can transmit data at high rates.

Various wireless communication networks enable users to communication anywhere and anytime. This also motivates new wireless applications, along with the increasing demands of users for faster and more stable wireless services. However, it seems that there is no much wireless spectrum left for future wireless applications. Fig. 1.1 shows the U.S. radio spectrum frequency allocation [10]. It can be seen that there is almost no spectrum available for new wireless applications. To alleviate the spectrum scarcity problem in the future, the spectrum resources should be more efficiently utilized.

One possible way to better utilize the spectrum is to use more advanced modulation and coding techniques. However, the current advanced transmission techniques nowadays have been able to reach the Shannon capacity closely, which is the theoretical upper bound on achievable rate. Therefore, there is no much room for increasing the data rate by more advanced modulation and coding techniques.

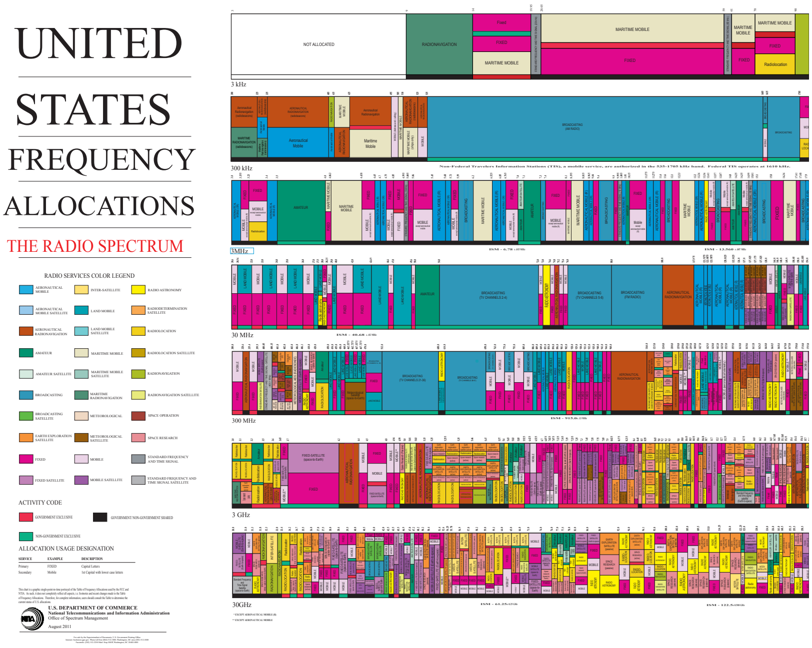

Fig. 1.1. U.S. frequency allocation chart as of October 2011.

On the other hand, a report from the U.S. Federal Communications Commission (FCC) indicates that the allocated spectrum is utilized at a percentage ranging from 15% to 85% in major US metropolitan areas, according to real measurements of the spectrum usage [11]. Therefore, the allocated spectrum is severely under-utilized, which is also supported by measurements from other organizations.

Both the spectrum scarcity and spectrum under-utilization motives the idea of dynamic spectrum access, which allows unlicensed users to temporarily access the spectrum in a period of time or at a particular region, which does not have harmful effects to licensed users. Cognitive radio is an emerging technique to realize the idea of dynamic spectrum access [12], [13].

1.2 Cognitive Radio: History and Progress

The term "cognitive radio" is first proposed by Joseph Mitola in 1999 [14], targeting at enhancing the flexibility of personal wireless devices. According to Joseph Mitola, cognitive radio is a smart radio which can observe the wireless communication environment, detect the user communication demands, and adjust its communication parameters intelligently. This definition of cognitive radio only describes its characteristics, such as environment awareness, intelligence, and adaptiveness. How the cognitive radio works is not specified. Recently, in [13], how cognitive radio works is described, by specifying the basic cognitive cycle with three fundamental tasks, namely radio-scene analysis, channel identification, and transmit-power control and dynamic spectrum management.

When cognitive radio is used to achieve dynamic spectrum access, licensed users and unlicensed users are referred to as primary users and secondary users, respectively. Primary users have priority in spectrum access. Secondary users can access the spectrum only when they do not affect primary activities. For spectrum sharing of secondary users, there are two modes: spectrum overlay mode and spectrum underlay mode [15]. In the overlay method, secondary user can access the spectrum only when primary users are not active. Therefore, secondary users first search, by spectrum sensing, for spectrum holes, i.e., the spectrum bands that are not utilized by primary users at a particular time or location, and then access those spectrum holes for a limited time. Since the primary users may return in some spectrum bands, the secondary users need to perform spectrum sensing periodically. Since the secondary users are not active simultaneously with primary users, there is few constraints on transmission power of secondary users. On the other hand, in underlay mode, secondary users and primary users can be simultaneously active at a location. To protect the activities of primary users, the interference from secondary users to primary users should be limited under a certain level.

Due to the ability of cognitive radio to alleviate the spectrum scarcity problem, many research institutions, radio management and standardization organizations have paid much attention to cognitive radio. In 2002, the FCC published a report [16], aiming at addressing ubiquitous spectrum issues such as spectral efficiency, interference protection and so on by improving the spectrum policies. To improve the utilization of spectrum resources, this report recommends the adoption of cognitive radio techniques. This report paves the way for the development of cognitive radio. In 2004, the IEEE started the standardization of the IEEE 802.22, which is the first worldwide air interface standard targeting at sharing of geographically unused spectrum assigned to the television broadcast service [17]. Other standards, such as European Computer Manufacturers Association 392 (ECMA 392) [18], IEEE Standards Coordinating Committee 41 (IEEE SCC41) [19] and IEEE 802.11af [20], also contribute to release cognitive radio's potential for commercial use. Some research institutions have also started projects on cognitive radio. For example, the neXt Generation (XG) program is to develop the technology to realize dynamic spectrum utilization; the European Dynamic Radio for IP Services in Vehicular Environments (DRiVE) project and the Spectrum Efficient Uni- and Multicast Over Dynamic Radio Networks in Vehicular Environments (OverDRiVE) project focus on the dynamic spectrum allocation in heterogeneous networks [21].

1.3 Thesis Motivations and Contributions

In cognitive radio, the two most important issues are channel sensing [22]–[31] and channel access [32]–[41]. Channel sensing is particularly for overlay mode, when secondary users first sense the availability of spectrum before transmitting over the spectrum. For channel access in either overlay mode or underlay mode, secondary users need to determine how to allocate limited resources optimally. This thesis focuses on optimal channel sensing, and optimal resource allocation in channel access.

1.3.1 Optimal Channel Sensing Setting

In a cognitive radio network with overlay mode, it is likely that there are multiple primary channels. And due to hardware constraint, a secondary user may be able to sense one channel at a time, and thus, each secondary user performs spectrum sensing sequentially, from one channel to another.

In addition, to increase the sensing accuracy, cooperative sensing is a good solution. In cooperative sensing, multiple secondary users first sense the channels, and send their sensing results to a coordinator, which makes decisions on presence or absence of primary activities. Details of cooperative spectrum sensing are to be discussed in Chapter 2.

1.3.1.1 Optimal Sensing Time Allocation in multichannel Cooperative Sensing

For cooperative sensing, the secondary users first sense the channels, and report results to the coordinator; the coordinator makes the decisions, and informs the secondary users how they can transmit their data over channels that are considered free. One important question is: how much time should be allocated for the secondary users to use in sensing? Generally, a larger portion of time for sensing leads to more accurate spectrum sensing, which gives primary activities more protection. However, a larger portion of sensing time also means a smaller portion of time for data transmission, which leads to a lower average throughput of secondary users. So there is a tradeoff between protection of primary activities and average throughput of secondary users. Optimal spectrum sensing strategy should be able to adjust the sensing time to balance two conflicting objectives: maximizing the average transmission rate of the secondary users, and protecting the activities of primary users at a certain level.

In Chapter 3, the research work on optimal spectrum sensing time setting in a multi-channel cognitive radio network with multiple secondary users is reported. The average throughput of multiple secondary users is targeted to be maximized while protecting the activities of primary users at a certain level. Optimization problems are formulated and proved to be non-convex, which are very hard to be solved optimally. A series of special properties of the formulated optimization problems are found. Then the non-convex optimization problems are decomposed into bi-level optimization problems, and each level can be converted to convex problems based on the special properties. Polynomial-complexity algorithms are developed to solve the formulated optimization problems optimally.

1.3.1.2 Channel Sensing Order with Noncooperative Spectrum Sensing

In this scenario, channels are sensed and accessed independently by multiple secondary users. If one channel is sensed to be free, a secondary user will stop sensing and start to access the channel. As the channel availabilities (which are usually measured by channel free probabilities) are always heterogeneous, how to set the spectrum sensing order is an open question. And when there are multiple secondary users, if two or more secondary users sense the same free channel simultaneously, they may access the channel at the same time, and cause a collision. Possible collisions should be considered when determining the sensing orders of the users, and actions need to be taken to avoid collisions.

As the first step to study the channel sensing order setting problem for multiple users, Chapter 4 considers a case with two secondary users. When adaptive modulation is not adopted by the secondary users, a secondary user stops sensing once a free channel is detected. Two sub-optimal algorithms are proposed, both of which are numerically verified to have comparable performance with the optimal solution by exhaustive search, and have much less computation complexity. When adaptive modulation is adopted by the secondary users, a secondary user stop sensing when a free channel with good channel quality is detected. Our results show that the stopping rule (when to stop sensing) should be designed jointly with the sensing order setting strategy of the two users.

1.3.2 Optimal Resource Allocation

Resource allocation is an important issue for any wireless network, as the resources are always limited. Resource allocation problems have been well investigated in the literature for traditional wireless networks such as cellular networks, WLANs, relay network, wireless sensor networks (WSNs) and many others. The resource allocation generally includes power allocation and spectrum allocation. An effective resource allocation strategy should be able to provide a wireless network with high throughput, low delay, low energy consumption, etc. To achieve this, optimization methods are widely used. In a cognitive radio network, however, the resource allocation has to be subject to the priority of primary activities. Therefore, cognitive resource allocation is much more challenging than traditional resource allocation, and is one focus of this thesis.

In a cognitive radio network in overlay mode, both spectrum sensing time and resource allocation strategy have impact on the average throughput of a secondary user. To maximize the average throughput of secondary users, previous works in the literature focus only on how to set the spectrum sensing time optimally while ignoring optimal resource allocation. In Chapter 5 of this thesis, resource allocation and sensing time configuration are investigated jointly in a multichannel cognitive radio network. An optimization problem is formulated by optimizing transmission power, accessible bandwidth of the secondary users and the spectrum sensing time such that the weighted average throughput of secondary users is maximized. The problem is non-convex. By applying monotonic optimization and bi-level optimization techniques, the non-convex problem can be solved optimally. The research is also extended to cases when the objective is to maximize the proportional fairness level or the max-min fairness level of secondary users.

In a cognitive radio network in underlay mode, spectrum sensing is not needed; but the secondary users have to limit the generated interference to primary users. Therefore, secondary resource allocation is subject to the constraint on interference to primary users. In this case, Chapter 6 investigates how to maximize average rate (the average of the instant rates over a long period) of the cognitive network with limited transmission power constraint, by performing power and bandwidth resource allocation. In the problem, the constraint of the transmission power of the secondary users can be either long-term or short-term, and the constraint of the interference-power to primary users can also be long-term or short-term. Then there are four possible combinations of power constraints. For each combination, the average rate maximization problem is formulated as a convex problem, which can be solved by traditional methods, but with high complexity and without closed-form solutions. In this research, special properties for the optimal solutions are derived, and based on these properties, simple online algorithms are given to quickly find optimal closed-form solutions.

1.4 Thesis Outline

The thesis is organized as follows. In Chapter 2, basic concepts of spectrum sensing and resource allocation in cognitive radio are introduced, and related literature is surveyed. Chapter 3 reports the research results on optimal spectrum sensing time allocation in a multichannel cognitive radio network with cooperative spectrum sensing. In Chapter 4, the channel sensing order setting problem is studied for a two-user cognitive radio network with noncooperative sequential spectrum sensing. In Chapter 5, spectrum sensing time and resource allocation are optimized jointly in a multichannel cognitive radio network in overlay mode. Chapter 6 presents the research results on optimal resource allocation for a cognitive radio network in underlay mode. Chapter 7 concludes the thesis and presents future research topics.

Chapter 2

Background and Literature Review

2.1 Spectrum Sharing

As introduced in Chapter 1, there are two modes of spectrum sharing by secondary users: spectrum underlay and spectrum overlay. In underlay mode, the interference from secondary users to primary users should be limited under a certain level. When the statistical information of primary channels is unknown by the secondary users [42], the power spectrum density (PSD) of secondary transmission should be much lower than a normal level in a traditional network. To guarantee the communication quality of secondary users, a widely spread bandwidth is required, which is typically achieved by CDMA or ultra-wideband (UWB) communications. When the statistical information of primary channel is known by the secondary users, the PSD of secondary transmission only needs to be less than an acceptable level for primary users, which probably does not need to be much lower than a normal level. So the secondary users do not have to resort to CDMA or UWB communications.

In the overlay method, secondary users can be active only when and where primary users are idle. A secondary user may need to access multiple spectrum holes that are not continuous. OFDM is a good candidate for this requirement. On a busy frequency band with active primary user, transmission activity of any secondary user is not allowed, which can be easily realized by OFDM through switching off the corresponding subcarriers. The goal of this method is to make the best use of opportunities in time and space.

In the framework of overlay sharing, after a secondary user has sensed idle channels, it can access the channels. However, if primary users become active subsequently, collisions may happen between secondary and primary users. One intuitive solution to this problem is to require that a secondary user terminate its transmission after a period of time and re-sense the channel availability. Thus, the dynamic spectrum access becomes periodical. In other words, the channel sensing and access is performed in each unit of time, e.g., in a slotted time structure, where each time slot can be divided into two parts: sensing and transmission. The slot structure can be categorized into two classes: constant access time (CAT) and constant data time (CDT) [43]. In the CAT class, the length of every time slot is fixed, i.e., the transmitter of a secondary user has a fixed amount of time in total for both sensing and data transmission. In the CDT class, the secondary transmitter has a constant amount of time for data transmission no matter how many channels it has sensed and how much time it has spent in sensing.

In the sensing part, multiple channels can be sensed sequentially (in the situation of limited hardware [21]) or in parallel (with the aid of wide-band spectrum sensing [44]). When the channel is sensed sequentially, the sensing part can be further divided into smaller time intervals. Each interval represents the time required for sensing one channel. In the transmission part, secondary user will transmit data if the target channel is sensed to be idle.

2.2 Spectrum Sensing in Cognitive Radio

2.2.1 Spectrum Sensing Technologies

Spectrum sensing plays an important role in cognitive radio. With the ability of spectrum sensing, secondary user can detect the activity of primary users and hence make a decision on whether or not to access spectrum. Three typical spectrum sensing techniques are: matched filter detection, cyclostationary detection and energy detection. The energy detection method is the most popular one for cognitive radio, due to its low complexity.

Matched Filter Detection: Matched filter is widely used in wireless communications for signal detection at the receiver as it can offer higher signal-to-noise ratio (SNR) [42]. Matched filter is realized by convolving unknown input signal with a conjugated time-reversed version of a known template signal to detect the presence of the template signal in the unknown input signal. When a secondary user has *a priori* knowledge of the primary user signal and uses it as the template signal, matched filter can be implemented to detect the presence of primary signal. However, knowing information of primary user signal at the secondary user is generally impractical.

Cyclostationary Detection: A cyclostationary signal is the signal with statistical properties varying cyclically with time [45]. In a wide sense, a signal can be classified to be cyclostationary if its autocorrelation function is a periodic function of time t. Specifically, for a random signal $a(t)$, the autocorrelation function is

$$R_a(t, \Delta t) = \mathbb{E}\left[a\left(t - \frac{\Delta t}{2}\right) a^*\left(t + \frac{\Delta t}{2}\right)\right] \tag{2.1}$$

where $\mathbb{E}[\cdot]$ means expectation. The signal $a(t)$ can be said to be wide-sense cyclostationary with period T_0, if

$$R_a(t, \Delta t) = R_a(t + T_0, \Delta t). \tag{2.2}$$

When a primary user signal has a strong cyclostationary property, the cyclostationary detection method can be used for spectrum sensing in cognitive radio. In addition, if a primary user signal is cyclostationary, the cyclostationary features are usually different from those of noise signal or secondary user signal. So cyclostationary detection method is robust to noise uncertainty and the interference signal from other secondary users.

Energy Detection: When secondary user knows little information of primary user signal, it is better to use energy detection. In energy detection approaches, the energy of interested spectrum is collected and compared with a predefined threshold ε. When the collected energy is larger than ε, the spectrum is estimated as busy; when the collected energy is lower than ε, the spectrum is estimated to be idle.

There are also some other spectrum sensing technologies, such as wavelet detection and covariance detection [46]. Interested readers please refer to [47] and [48].

Due to the simplicity and no requirement on primary signal information, energy detection is the most popular spectrum sensing technique, and is adopted in our research when spectrum sensing is needed.

2.2.2 Energy Detection

The collected energy can be expressed as the energy of sampled signals in time domain. Next is one illustration from [49]. Denote μ as the sampling rate and τ as sensing time, the binary hypothesis of spectrum sensing is given as:

$$\begin{aligned} \mathcal{H}^0: \quad & y(i) = w(i), && i = 1, 2, ..., \mu\tau \\ \mathcal{H}^1: \quad & y(i) = r(i) + w(i), && i = 1, 2, ..., \mu\tau \end{aligned} \tag{2.3}$$

where $\mathcal{H}^0$ and $\mathcal{H}^1$ mean that the primary user is idle and busy, respectively, i is the sample index, $y(\cdot)$ is the received signal at the secondary user, $w(\cdot)$ is background noise which is assumed to be circular symmetric complex gaussian (CSCG), independent and identically distributed (i.i.d) random process with mean being 0 and variance being σ^2, $r(\cdot)$ is the received primary user signal at the secondary user which is assumed to be an i.i.d random process with mean being zero and variance being σ_r^2. The *test statistic* of received signal energy at the secondary user is defined as [49]

$$T = \frac{1}{\mu\tau} \sum_{i=1}^{\mu\tau} |y(i)|^2. \tag{2.4}$$

The primary user will be estimated to be busy when $T \geq \varepsilon$, or idle when $T < \varepsilon$.

The accuracy of spectrum sensing can be measured by two metrics: detection probability (i.e., the probability that, if there is primary activity, the secondary user can detect it successfully) and false alarm probability (i.e., the probability that, if there is no primary activity, the secondary user falsely concludes that primary user is active). Denote the detection probability as P_d and false alarm probability as P_f. When $r(i)$ is also a CSCG random variable, reference [49] shows that P_d and P_f can be expressed approximately as functions of detection threshold ε and number of samples $\mu\tau$ (or sensing time τ when sampling rate μ is fixed), as

$$P_d(\tau, \varepsilon) = \Pr\left(T \geq \varepsilon | \mathcal{H}^1\right) = Q\left(\left(\frac{\varepsilon}{\sigma^2} - \gamma - 1\right) \sqrt{\mu\tau/(\gamma+1)^2}\right) \tag{2.5}$$

and

$$P_f(\tau, \varepsilon) = \Pr\left(T \geq \varepsilon | \mathcal{H}^0\right) = Q\left(\left(\frac{\varepsilon}{\sigma^2} - 1\right) \sqrt{\mu\tau}\right) \tag{2.6}$$

respectively, where γ is the average SNR of primary user signal received by the secondary user, defined as $\gamma \triangleq \frac{(\sigma^r)^2}{\sigma^2}$, $Q(\cdot)$ is the Q function, defined as $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{z^2}{2}\right) dz$ and $\Pr(\cdot)$ stands for the probability of an event. From (2.5) and (2.6), the P_f can be also expressed as a function of P_d and τ as

$$P_f\left(\tau, P_d\right) = Q\left((\gamma+1)Q^{-1}\left(P_d\right) + \gamma\sqrt{\mu\tau}\right). \tag{2.7}$$

Alternatively, another way to express the collected energy is to write it as the energy of received signal in frequency domain, which can be obtained by performing fast Fourier transform (FFT) over sampled signal in time domain. In particular, reference [44] proposes a wideband spectrum sensing technique that can detect the activities of primary users spanning over multiple frequency bands, namely *multiband joint detection*. To realize multiband joint detection, a capability of high sampling rate at the secondary user is assumed. Suppose there are N frequency bands to be sensed. In a multipath fading environment, denote $h(l), l = 0, 1, ..., L-1$ as the discrete-time channel impulse response between the primary transmitter and the secondary user, where L is the number of resolvable paths. The received signal at the secondary user can be expressed as

$$y(i) = \sum_{l=0}^{L-1} h(l)s(i-l) + w(i) \tag{2.8}$$

where $s(i)$ denotes the i^{th} sample of transmitted signal of primary user, and $w(\cdot)$ is i.i.d and CSCG random process with mean being 0 and variance being σ^2. The N-point FFT ($N \geq L$) of received signal $y(i)$ is given as:

$$\begin{aligned} Y_n &= \tfrac{1}{\sqrt{N}} \sum_{i=0}^{N-1} y(i) e^{\frac{-j2\pi in}{N}} \\ &= H_n S_n + W_n, n = 1, 2, ..., N \end{aligned} \tag{2.9}$$

where H_n, S_n and W_n are N-point FFT of the discrete-time channel impulse response $h(l)$, the primary signal $s(i)$ and noise signal $w(i)$ respectively. The binary hypotheses of the activity on channel n can be given as:

$$\begin{aligned} &\mathcal{H}_n^0 : Y_n = W_n \\ &\mathcal{H}_n^1 : Y_n = H_n S_n + W_n \end{aligned} \tag{2.10}$$

where $\mathcal{H}_n^0$ and $\mathcal{H}_n^1$ mean that the primary user in channel n is idle and busy respectively. Within a duration of τ in the sensing phase, the secondary user collects $\mu\tau$ samples of received signal. By doing the FFT on every sampled N points, we have a number, $\frac{\mu\tau}{N}$, of values for Y_n, denoted $Y_n(1), Y_n(2), ..., Y_n(\frac{\mu\tau}{N})$. By employing energy detection, test statistic T_n is defined and the decision rule is given as

$$T_n \triangleq \frac{N}{\mu\tau} \sum_{k=1}^{\mu\tau/N} |Y_n(k)|^2 \begin{cases} \geq \varepsilon_n, \mathcal{H}_n^1 \\ < \varepsilon_n, \mathcal{H}_n^0 \end{cases} \tag{2.11}$$

for $n = 1, 2, ..., N$, where ε_n is the threshold on n^{th} frequency band.

For energy detection in frequency domain, detection probability and false alarm probability on channel n, denoted as P_n^d and P_n^f respectively, can be also approximately expressed as a function of τ and ε_n similarly to equations (2.5) and (2.6) [44]. In addition, P_n^f can be also expressed as a function of τ and P_n^d, like equation (2.7).

An energy detection method is easy to be implemented and requires little information of primary user signal. An obvious drawback of an energy detection method is that the energy detector cannot distinguish primary user signal from the interference signal coming from other secondary transmitters. This drawback can be addressed by coordinating the transmissions of all secondary users.

2.2.3 Cooperative Spectrum Sensing

When the channel from a primary transmitter to a secondary user experiences shadowing or multipath fading, it is difficult for the secondary user to detect the presence of the primary transmitter. In this case, the secondary user may mistakenly access the licensed spectrum with a primary user being active on it, and bring significant interference to the primary receiver that might be near the secondary user. This issue can be addressed by employing cooperative spectrum sensing, in which multiple spatially disjoint secondary users cooperate to sense interested frequency bands together [46]. Cooperative spectrum sensing is usually performed in the following procedure: In the first stage, every secondary user senses the interested spectrum independently. In the second stage, all the secondary users forward their sensing data or local decisions to a coordinator. In the final stage, the coordinator makes a final decision on whether a primary user is present or absent.

Depending on the information forwarded from every secondary user to the coordinator, there are two kinds of cooperative spectrum sensing: decision fusion cooperative spectrum sensing and data fusion cooperative spectrum sensing [50]. In decision fusion cooperative spectrum sensing, every secondary user makes a local binary decision on the presence or absence of primary user over the interested spectrum, and forwards the one-bit decision information to the coordinator. In data fusion cooperative spectrum sensing, the coordinator collects sensing data from each secondary user, which are generally test statistics obtained independently by each secondary user. After collecting these test statistics, the coordinator linearly combines the test statistics to obtain an *overall test statistic* for deciding on the presence or absent of primary user [44].

2.2.4 Tradeoff between Sensing Accuracy and Data Transmission

In the slot structure for sensing and data transmission, a longer spectrum sensing time can produce a higher detection probability and a lower false alarm probability as shown in equations (2.5) and (2.6), respectively, and therefore, more protection for primary users is provided, and more real spectral opportunities for secondary users are detected. On the other hand, longer spectrum sensing time will lead to less portion of time for data transmission of secondary users. How to set spectrum sensing time to balance sensing accuracy and data transmission is an interesting topic.

In a single-channel scenario, optimal sensing time duration is derived in [49], to maximize the throughput of secondary user while guaranteeing the detection probability above a threshold. CAT slot structure is assumed with total slot duration T and sensing duration τ. A secondary user can transmit data in two cases: when the primary user is idle and is sensed to be idle, the secondary user will transmit at rate C_0; when the primary user is busy but is sensed to be idle, the secondary user will transmit at a lower rate C_1, due to the interference introduced by busy primary user. The first case happens with probability $\Pr(\mathcal{H}^0)(1-P_f)$, while the second case happens with probability $\Pr(\mathcal{H}^1)(1-P_d)$. The average capacity, or average throughput, in a slot is given as $(1-\frac{\tau}{T})\left[\Pr(\mathcal{H}^0)(1-P_f)C_0+\Pr(\mathcal{H}^1)(1-P_d)C_1\right]$, which is maximized while guaranteeing the detection probability P_d to be larger than a threshold.

In [51], a different perspective is taken for the spectrum sensing time configuration. The primary user activity is described as a two-state (ON and OFF) birth-death process. The On state and OFF state represent the period when primary user is active and idle, respectively. Both the duration of ON and OFF state are independent and exponentially distributed. CAT slot structure is also used with total slot duration T and sensing time τ. Both T and τ are optimized to maximize $\frac{T-\tau}{T}$ (the portion of time used for data transmission) while limiting the potential interference to primary users.

2.2.5 Sequential Sensing

In one time slot, sequentially sensing multiple channels is also possible. In this case, there is a gain coming from the fact that more and/or higher-quality available channels may be sensed and accessed. On the other hand, there is also a loss coming from the less data transmission time. So when sensing multiple channels sequentially in a slot, a tradeoff exists in whether to sense more channels or to stop at a sensed-free channel. This problem falls into the definition of well known *optimal stopping problem* [52], [53].

In [54], optimal stopping rule is applied to maximize expected usable bandwidth when a secondary user senses multiple channels in one slot. CDT slot structure is employed with data transmission duration T_f and one-channel sensing duration δ. N channels are sensed sequentially from channel 1 to channel N. In each step of sequential spectrum sensing, the secondary user makes a decision on whether to proceed to sense next channel or to stop sensing and start data transmission over channels that are sensed free.Take the i^{th} step of decision for example. Assume sensing is perfect. Let X_n represent the random state and x_n represent the sensing result of n^{th} channel. X_n is a random variable but x_n is not. $X_n = 0$ and $X_n = 1$ means the n^{th} channel is busy and idle respectively. After performing spectrum sensing over the i^{th} channel, the instantaneous reward $y_i(x_1, x_2, ..., x_i)$ is expressed as

$$y_i(x_1, x_2, ..., x_i) = \frac{T_f}{T_f + i\delta} b_i(x_1, x_2, ..., x_i), \tag{2.12}$$

where $b_i(x_1, x_2, ..., x_i)$ is the maximal usable bandwidth over the channel set $\{1, 2, ..., i\}$. The expected reward if secondary user proceeds to sense the $(i+1)^{th}$ channel, $\mathbb{E}\left[V_{i+1}^N\right]$, can be obtained by following the iteration formulas recursively, which are given as

$$V_N^N(x_1, ..., x_N) = y_N(x_1, ..., x_N) = \frac{T_f}{T_f + N\delta} b_N(x_1, ..., x_N) \tag{2.13}$$

$$\begin{aligned} &V_{i+1}^N(x_1, ..., x_{i+1}) = \\ &\max\left(y_{i+1}(x_1, ..., x_{i+1}), \mathbb{E}\left[V_{i+2}^N(x_1, ..., x_{i+1}, X_{i+2}) | X_1 = x_1, ..., X_{i+1} = x_{i+1}\right]\right) \end{aligned} \tag{2.14}$$

where V_{i+1}^N means the maximum throughput when the secondary user proceeds to sense the $(i+1)^{th}$ channel. In the i^{th} step of sensing procedure, if y_i is larger than $\mathbb{E}\left[V_{i+1}^N(x_1, ..., x_i, X_{i+1})\right]$, sequential channel probing should be stopped. Otherwise the secondary user should proceed to sense next channel.

Reference [55] also works on optimal stopping problem to maximize average throughput in cognitive radio. Different from [54], reference [55] exploits the potential gain when knowing the SNR between secondary transmitter and secondary receiver. In addition, reference [55] also analyzes the maximal average throughput as a function of spectrum sensing time and SNR probing time.

In real applications, it is very likely that the multiple channels are not homogeneous, especially in sense of available probability. The heterogeneity among multiple channels introduces another dimension of flexibility by selecting the sensing order of multiple channels.

Reference [56] considers both optimal stopping and spectrum sensing order problem. CAT slot structure is employed with total slot length as T and single channel sensing time as δ. N channels are assumed with free probabilities $\theta_1, \theta_2, ..., \theta_N$, respectively. The sensing order is denoted as $(s_1, ..., s_N)$. In channel sensing, if channel s_i is not free, the secondary user continues to sense channel s_{i+1}; if channel s_i is sensed to be free with channel gain SNR_{s_i}, the secondary user will transmit on channel s_i when the average transmission rate $\left(1 - \frac{i\delta}{T}\right)\ln(1 + \mathrm{SNR}_{s_i})$ is larger than the expected average rate at the $(i+1)^{th}$ sensing, U_{i+1}, which can be also obtained by following certain iterative formulas. Intuitively, the optimal sensing order should be sorted by θ_i in a descending order. However, it is proved that the intuitive rule is not optimal in general. In addition, special properties are proved to provide simple sensing order searching methods when some conditions are met.

2.3 Resource Allocation in Cognitive Radio

As the resources (generally regarded as transmission power and spectrum) are always limited, it is important to allocate the resources efficiently for any wireless network, so as to improve the performance of the wireless network. An efficient and effective resource allocation strategy should be able to provide a wireless network with longer lifetime [57], higher throughput [58]–[62], more fairness [63] or better quality of service (QoS) [64].

When studying the cognitive radio, both underlay mode and overlay mode should be considered. In overlay mode, the secondary users first sense the availability of the spectrum. Then resource allocation can be performed on idle channels. In this case, resource allocation strategies in traditional networks can be followed. In underlay mode the interference to primary users should be limited [65]–[67]. The request for restricting the interference to primary users imposes additional constraint on the transmission power of secondary users. In the following, the resource allocation (including power allocation and spectrum allocation) in cognitive radio is introduced. Without the loss of generality the power allocation is discussed in the setup of underlay mode.

2.3.1 Power Resource Allocation

As of power allocation, it usually means the power allocation over multiple channels. Consider a general cognitive radio network with N channels and M secondary users. The set of channels is denoted as $\mathcal{N}$ and the set of secondary users is denoted as $\mathcal{M}$. The transmission power of the m^{th} user on the n^{th} channel is denoted as $p_{n,m}$. The channel gain from the m^{th} secondary transmitter to the primary receiver on the n^{th} channel is denoted as $h_{n,m}$. The set of transmission powers have to satisfy the following two kinds of constraints simultaneously:

- Transmission-power constraint: For a secondary user, the transmission power over all the channels should be not larger than the maximal transmission power of the secondary transmitter, i.e., $\sum_{n=1}^{N} p_{n,m} \leq P_m, \forall m \in \mathcal{M}$, where P_m is the maximal transmission power of the m^{th} secondary transmitter.
- Interference-power constraint: To protect the data transmission of the primary user on the n^{th} channel, the interference generated by all the secondary users to the primary receiver on the n^{th} channel should be no larger than a threshold I_n, i.e., $\sum_{m=1}^{M} p_{n,m} h_{n,m} \leq I_n, n \in \mathcal{N}$ [68].

As usually presented in literatures, the power constraints can also fall into either one or both of the following two types of constraints [67], [69].

- Average (Long-term) power constraint: Average power constraint is the kind of constraint that restricts transmission power on average over a period of time, during which the channel state varies and is usually assumed to be an ergodic process.

- Instantaneous (Short-term) power constraint: The instantaneous power constraint is the kind of constraint that holds at every instant. In the perspective of power supply of the transmitter, the instantaneous power constraint can be also interpreted as peak power constraint.

Restricted by the power constraints, optimization problems are formulated in [67], [70], [71], to maximize (or minimize) a certain utility, e.g., data rate, power consumption, of the secondary users.

2.3.2 Spectrum Resource Allocation

In a cognitive radio system, the spectrum can be divided into multiple small bands, called channels for simplicity of presentation. Aided by orthogonal frequency division multiple access (OFDMA) technique, a secondary user can access a subset of all available channels. In this context, the spectrum allocation can be realized by performing channel allocation, which determines which set of channels is allocated to the secondary user. Whether or not to allocate a channel can be indicated by an integer variable with the value being 0 or 1. In this case, the channel allocation problem can be formulated as a mixed-integer optimization problem, which is usually difficult to solve optimally. Suboptimal algorithms are always developed in literature [72], [73].

Chapter 3

Sensing Time Allocation in Multi-channel Cooperative Sensing

For spectrum sensing, a slotted time frame structure is widely used [56], [74], [75]. In each time slot, the first portion is used for spectrum sensing, and the second portion is used for packet transmission (if the channel is detected idle). A longer sensing time in a time slot will lead to a higher detection probability and a lower false alarm probability, which are desired. But it also results in less time in actual information transmission (assuming the duration of each time slot is fixed). Therefore, a tradeoff exists in the sensing time setting. The optimal tradeoff in sensing time setting is investigated in [49], [76] so as to optimally utilize the transmission opportunities in a single channel.

Different from the work for a single channel, in this chapter, we investigate the sensing time setting for a multi-user multi-channel case with cooperative sensing. A cognitive radio network with multiple potential channels is considered. Secondary users cooperatively sense the channels and send the sensing results to a coordinator, in which energy detection with a soft decision rule is employed to estimate whether there are primary activities in the channels. An optimization problem is formulated, which maximizes the throughput of secondary users while keeping detection probability for each channel above a pre-defined threshold. In particular, two sensing modes are investigated: slotted-time sensing mode and continuous-time sensing mode. With a slotted-time sensing mode, the sensing time of each secondary user consists of a number of mini-slots, each of which can be used to sense one channel. The initial optimization problem is shown to be a nonconvex mixed-integer problem. A polynomial-complexity algorithm is proposed to solve the problem optimally. With a continuous-time sensing mode, the sensing time of each secondary user for a channel can be any arbitrary continuous value. The initial nonconvex problem is converted into a convex bilevel problem, which can be successfully solved by existing methods. Numerical results are presented to demonstrate the effectiveness of our proposed algorithms. [1]

3.1 System Model

We consider a cognitive radio network with N frequency bands and M secondary users. In each channel, a primary user exists (which may not be active all the time). CAT slot structure is employed in the cognitive radio network. In each CAT slot, the primary user in a channel is either active for the whole slot, or idle for the whole slot. In the sensing phase of a CAT slot, each secondary user senses a number of channels sequentially by energy detection in time domain as introduced in Chapter 2.2.1 with sampling rate μ and threshold $\varepsilon_n, n \in \mathcal{N}$ where $\mathcal{N} = \{1, 2, ..., N\}$. The duration of the sensing phase is a design parameter. It is assumed that the channel gains in each channel (from the primary user to secondary users or between secondary users) keep fixed within the duration of a CAT time slot.

There is a coordinator in the cognitive radio network to perform data fusion cooperative spectrum sensing. If the coordinator estimates a channel, say channel n, to be idle, it notifies the secondary user assigned in the channel to transmit. The transmission power is p_n^s, and the transmission rate is given by $\log(1 + \text{SNR}_n^{ss})$, where SNR_n^{ss} means the SNR from the secondary user to its receiver at channel n.

Let t_n^m denote the time duration that secondary user m spends in sampling channel n. So within duration t_n^m, user m has μt_n^m samples for channel n, which follow a binary hypothesis:

$$\begin{aligned} \mathcal{H}_n^0 :&\quad y_n^m(i) = w_n(i), & i = 1, 2, ..., \mu t_n^m \\ \mathcal{H}_n^1 :&\quad y_n^m(i) = r_n^m(i) + w_n(i), & i = 1, 2, ..., \mu t_n^m \end{aligned} \tag{3.1}$$

where $\mathcal{H}_n^0$ and $\mathcal{H}_n^1$ mean that the primary user in channel n is idle and busy respectively, i is the sample index, $y_n^m(\cdot)$ is the received signal of channel n at secondary user m, $w_n(\cdot)$ is background noise in channel n, which is assumed to be CSCG with mean being zero and

[1] *A version of this chapter has been published in IEEE Trans. Wireless Commun., 9: 1128-1138 (2010).*

variance being σ^2, and $r_n^m(\cdot)$ is the signal of primary user in channel n received at secondary user m. As in Chapter 2.2.1, $r_n^m(i)$ is assumed to be a zero mean CSCG random variable. Furthermore, we also suppose $r_n^m(i)$ are independent and identically distributed random variables with variance $(\sigma_n^r)^2$ for $m \in \mathcal{M}$, where $\mathcal{M} = \{1, 2, ..., M\}$. This assumption is valid for a small-sized cognitive network (i.e., distance between the secondary users is much less than the distance from the primary user to the secondary users).

Then, the test statistic of secondary user m's received signal energy in channel n is calculated as

$$T_n^m(y) = \frac{1}{\mu t_n^m} \sum\nolimits_{i=1}^{\mu t_n^m} |y_n^m(i)|^2. \tag{3.2}$$

By performing data fusion cooperative sensing, secondary user $m, \forall m \in \mathcal{M}$ will transmit the number $T_n^m(y), \forall n \in \mathcal{N}$ to the coordinator, and the overall test statistic for channel n is calculated at the coordinator as

$$T_n^{\text{all}}(y) = \frac{\sum_{m=1}^{M} \mu t_n^m \cdot T_n^m(y)}{\sum\limits_{m=1}^{M} \mu t_n^m}. \tag{3.3}$$

With the overall test statistic $T_n^{all}(y)$, the detection probability and false alarm probability on channel n, P_n^d and P_n^f, can be expressed as the way in Chapter 2.2.1

$$\begin{aligned} P_n^d\left(\sum\nolimits_{m=1}^{M} t_n^m, \varepsilon_n\right) &= \Pr\left(T_n^{\text{all}}(y) > \varepsilon_n | \mathcal{H}_n^1\right) \\ &= Q\left(\left(\tfrac{\varepsilon_n}{\sigma^2} - \gamma_n - 1\right)\sqrt{\mu \sum\limits_{m=1}^{M} t_n^m / (\gamma_n + 1)^2}\right) \end{aligned} \tag{3.4}$$

and

$$P_n^f\left(\sum_{m=1}^{M} t_n^m, \varepsilon_n\right) = \Pr\left(T_n^{\text{all}}(y) > \varepsilon_n | \mathcal{H}_n^0\right) = Q\left(\left(\frac{\varepsilon_n}{\sigma^2} - 1\right)\sqrt{\mu \sum_{m=1}^{M} t_n^m}\right) \tag{3.5}$$

respectively, where γ_n is the average SNR of primary user signal received by a secondary user in channel n, defined as $\gamma_n \triangleq \frac{(\sigma_n^r)^2}{\sigma^2}$.

In a real system, the detection probability P_n^d should be no less than 0.5 and the false alarm probability should be no larger than 0.5. From (3.4) and (3.5), the constraints $P_n^d \geq$

0.5 and $P_n^f \leq 0.5$ are equivalent to the following inequality.

$$\sigma^2 \leq \varepsilon_n \leq \sigma^2(1+\gamma_n), n \in \mathcal{N}. \tag{3.6}$$

The following equation rewrites the false alarm probability as a function of the detection probability P_n^d like the way in (2.7).

$$P_n^f \left(\sum_{m=1}^{M} t_n^m, P_n^d \right) = Q\left((\gamma_n+1)Q^{-1}\left(P_n^d\right) + \gamma_n \sqrt{\mu \sum_{m=1}^{M} t_n^m} \right). \tag{3.7}$$

Within this expression, the constraint $P_n^f \leq 0.5$ is equivalent to

$$(\gamma_n+1)Q^{-1}\left(P_n^d\right) + \gamma_n \sqrt{\mu \sum_{m=1}^{M} t_n^m} \geq 0, n \in \mathcal{N}. \tag{3.8}$$

From (3.4) and (3.5) it can be seen that only the total time used to sense channel n, i.e., $\sum_{m=1}^{M} t_n^m$, affects the detection performance for channel n, regardless of how this total time is distributed among the secondary users.

For the sensing in channel n, we have the following four scenarios:

- If channel n is idle and is estimated by the coordinator to be idle, then the secondary user assigned to channel n will transmit in the associated transmission phase of the slot, with the average transmission rate given by $R_n^0 = \mathbb{E}\left[\log\left(1+\frac{|h_n^{ss}|^2 p_n^s}{\sigma^2}\right)\right]$, where h_n^{ss} is the channel coefficient from the secondary user assigned to channel n to its receiver.

- If channel n is idle and is estimated by the coordinator to be busy (i.e., a false alarm happens), the secondary user assigned to channel n will not transmit in the associated transmission phase of the slot.

- If channel n is busy and is estimated by the coordinator to be busy, the secondary user assigned to channel n will not transmit in the associated transmission phase of the slot.

- If channel n is busy and is estimated by the coordinator to be idle (i.e., a missed detection happens), then the secondary user assigned to channel n will transmit in

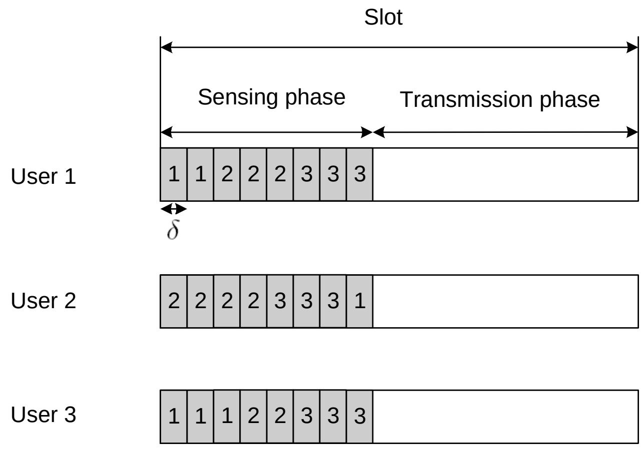

Fig. 3.1. The slotted-time sensing mode with 3 channels and 3 secondary users.

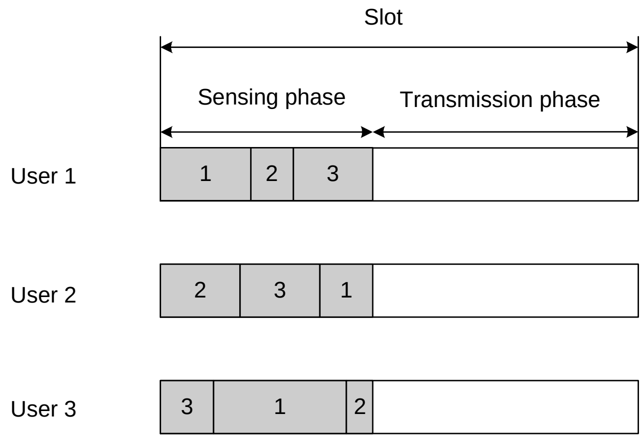

Fig. 3.2. The continuous-time sensing mode with 3 channels and 3 secondary users.

the associated transmission phase of the slot. As the primary user's signal will serve as an interference to the secondary transmission, the average transmission rate of the secondary user is given by $R_n^1 = \mathbb{E}\left[\log\left(1 + \frac{|h_n^{ss}|^2 p_n^s}{|h_n^{ps}|^2 p_n^p + \sigma^2}\right)\right]$, where p_n^p is the transmission power of the primary user in channel n, and h_n^{ps} is the channel coefficient from the primary user to the secondary receiver in channel n.

It can be seen that $R_n^0 > R_n^1$.

In the sensing phase, a secondary user needs to sequentially sense a number of channels, using one of the two modes: slotted-time sensing mode and continuous-time sensing mode.

- Slotted-time sensing mode: The sensing phase in each slot is further divided into a number, k, of mini-slots, each with duration δ. Here the value of k is a parameter to be optimized. Each mini-slot can be used by a secondary user to sense any channel. Fig. 3.1 shows an example for $N = 3$ channels, $M = 3$ secondary users, and $k = 8$ mini-slots in a sensing phase. The number inside each mini-slot means the channel to be sensed. By counting how many mini-slots are used to sense every channel by each secondary user, the sensing time $t_n^m, \forall n \in \mathcal{N}, m \in \mathcal{M}$ can be obtained. Take User 1 for example, 2 mini-slots are used to sense channel 1, and 3 mini-slots are used to sense channel 2 and channel 3 respectively, so $t_1^1 = 2\delta$, $t_2^1 = 3\delta$ and $t_3^1 = 3\delta$. Note that in the example, user 2 first senses channel 2, then channel 3, and finally channel 1. Actually the sensing order of channels 1, 2 and 3 does not affect the detection performance.

- Continuous-time sensing mode: In the sensing phase, the time to sense a channel can be any arbitrary-length duration bounded by the total duration in the sensing phase, while the duration of the sensing phase is also a parameter to be optimized. Fig. 3.2 shows an example for $N = 3$ channels and $M = 3$ secondary users.

In the following two sections, optimal sensing time settings in the slotted-time sensing mode and continuous-time sensing mode are investigated respectively to maximize the total throughput of the secondary network.

3.2 Optimal Sensing Time Setting in the Slotted-Time Sensing Mode

In the system, the sensing phase in a slot has k mini-slots. The value of k is a parameter to be optimized. Each mini-slot can be used by a secondary user to sense a channel. So there are totally kM mini-slots among the M secondary users to sense the N channels. Let $k_n > 0$ denote the number of mini-slots (among the kM mini-slots) that are used for sensing channel $n \in \mathcal{N}$. Then we have

$$\sum_{n=1}^{N} k_n = kM. \tag{3.9}$$

Let T denote the length of a time slot. Then the average throughput of channel n can be expressed [49] as

$$C_n(k,k_n,\varepsilon_n)=\frac{T-k\delta}{T}\Big(\operatorname{Pr}(\mathcal{H}_n^0)\left(1-P_n^f(k_n,\varepsilon_n)\right)R_n^0+\operatorname{Pr}(\mathcal{H}_n^1)\left(1-P_n^d(k_n,\varepsilon_n)\right)R_n^1\Big) \tag{3.10}$$

where

$$P_n^d(k_n,\varepsilon_n)=Q\left(\left(\frac{\varepsilon_n}{\sigma^2}-\gamma_n-1\right)\sqrt{\frac{k_n\delta\mu}{(\gamma_n+1)^2}}\right) \tag{3.11}$$

and

$$P_n^f(k_n,\varepsilon_n)=Q\left(\left(\frac{\varepsilon_n}{\sigma^2}-1\right)\sqrt{k_n\delta\mu}\right) \tag{3.12}$$

are from (3.4) and (3.5), respectively, $\operatorname{Pr}(\mathcal{H}_n^0)\geq 0$ is the available probability of channel n, and $\operatorname{Pr}(\mathcal{H}_n^1)=1-\operatorname{Pr}(\mathcal{H}_n^0)\geq 0$ is the busy probability of channel n.

Our goal is to maximize the throughput of secondary users in all the channels, denoted $C(k,\{k_n\},\{\varepsilon_n\})$,[2] while keeping the detection probability of any channel, P_n^d, above a pre-specified threshold $P_{th}(P_{th}>0.5)$ and the false alarm probability of any channel, P_n^f, no larger than 0.5. [3] So the problem can be defined as follows.

<u>Problem</u> 3.1.

$$\begin{aligned}
\max_{k,\{k_n\},\{\varepsilon_n\}} \quad & C(k,\{k_n\},\{\varepsilon_n\})=\tfrac{T-k\delta}{T}\cdot\textstyle\sum_{n=1}^{N}\Big(\operatorname{Pr}(\mathcal{H}_n^0)\left(1-P_n^f(k_n,\varepsilon_n)\right)R_n^0\\
&\qquad\qquad+\operatorname{Pr}(\mathcal{H}_n^1)\left(1-P_n^d(k_n,\varepsilon_n)\right)R_n^1\Big)\\
\text{s.t.} \quad & P_n^d(k_n,\varepsilon_n)\geq P_{th}, n\in\mathcal{N}\\
& \sigma^2\leq\varepsilon_n\leq\sigma^2(1+\gamma_n), n\in\mathcal{N}\\
& \sum_{n=1}^{N}k_n=kM\\
& k_n>0, k_n\in\mathcal{I}, n\in\mathcal{N}\\
& 0<k\leq\lfloor\tfrac{T}{\delta}\rfloor,\ \ k\in\mathcal{I}.
\end{aligned}$$

Here $\mathcal{I}$ is the set of all positive integers and $\lfloor\cdot\rfloor$ is the flooring function. Note that the constraints $\sigma^2\leq\varepsilon_n\leq\sigma^2(1+\gamma_n)$ are equivalent to the constraints $P_n^d\geq 0.5$ and $P_n^f\leq 0.5$ according to equations (3.11) and (3.12).

[2]Note that here $\{k_n\}$ means the set of $\{k_1,k_2,...,k_N\}$, and $\{\varepsilon_n\}$ means the set of $\{\varepsilon_1,\varepsilon_2,...,\varepsilon_N\}$.

[3]When additional constraints $P_n^f\leq P_{th}^f$ (where P_{th}^f is the pre-specified threshold such that $P_{th}^f\leq 0.5$), $\forall n\in\mathcal{N}$ are imposed for the proposed optimization problems in slotted-time sensing mode and continuous-time sensing mode, the methods presented in this chapter also work.

3.2.1 Nonconvexity of Problem 3.1

Lemma 3.1. Problem 3.1 is not a convex problem.

Proof. Assume that problem 3.1 is a convex problem. Thus, the objective function $C(k, \{k_n\}, \{\varepsilon_n\})$ is a concave function with respect to $k, \{k_n\}$, and $\{\varepsilon_n\}$. Based on (3.11) and (3.12), we have

$$\frac{\partial^2 P_n^f(k_n, \varepsilon_n)}{\partial {k_n}^2} = \frac{1}{4\sqrt{2\pi}} \left(\frac{\varepsilon_n}{\sigma^2} - 1\right) e^{-\frac{\left(\left(\frac{\varepsilon_n}{\sigma^2}-1\right)\sqrt{k_n \delta\mu}\right)^2}{2}} \cdot \left(\sqrt{\frac{\delta\mu}{k_n^3}} + (\delta\mu)^{3/2}\left(\frac{\varepsilon_n}{\sigma^2} - 1\right)^2 \frac{1}{\sqrt{k_n}}\right) \tag{3.13}$$

and

$$\frac{\partial^2 P_n^d(k_n, \varepsilon_n)}{\partial {k_n}^2} = \frac{1}{4\sqrt{2\pi}} \left(\frac{\varepsilon_n}{\sigma^2} - \gamma_n - 1\right) e^{-\frac{\left(\left(\frac{\varepsilon_n}{\sigma^2}-\gamma_n-1\right)\frac{\sqrt{k_n \delta\mu}}{(\gamma_n+1)}\right)^2}{2}} \cdot \left(\sqrt{\frac{\delta\mu}{(\gamma_n+1)^2}} \frac{1}{\sqrt{k_n^3}} + \frac{(\delta\mu)^{3/2}}{(\gamma_n+1)^3}\left(\frac{\varepsilon_n}{\sigma^2} - \gamma_n - 1\right)^2 \frac{1}{\sqrt{k_n}}\right). \tag{3.14}$$

Among the constraints of problem 3.1, we have $\sigma^2 \le \varepsilon_n \le \sigma^2(1+\gamma_n)$. So in the feasible region of problem 3.1, we have

$$\frac{\partial^2 P_n^f(k_n, \varepsilon_n)}{\partial {k_n}^2} > 0 \tag{3.15}$$

and

$$\frac{\partial^2 P_n^d(k_n, \varepsilon_n)}{\partial {k_n}^2} < 0. \tag{3.16}$$

Since $\Pr(\mathcal{H}_n^1) = 1 - \Pr(\mathcal{H}_n^0)$, it can be obtained that

$$\frac{\partial^2 C(k, \{k_n\}, \{\varepsilon_n\})}{\partial {k_n}^2} = \frac{T - k\delta}{T} \cdot \Bigg(- \frac{\partial^2 P_n^d(k_n, \varepsilon_n)}{\partial {k_n}^2} R_n^1 - \Pr(\mathcal{H}_n^0)\left(\frac{\partial^2 P_n^f(k_n, \varepsilon_n)}{\partial {k_n}^2} R_n^0 - \frac{\partial^2 P_n^d(k_n, \varepsilon_n)}{\partial {k_n}^2} R_n^1\right)\Bigg).$$

Therefore, if

$$\Pr(\mathcal{H}_n^0) < \frac{-\frac{\partial^2 P_n^d(k_n, \varepsilon_n)}{\partial {k_n}^2} R_n^1}{\frac{\partial^2 P_n^f(k_n, \varepsilon_n)}{\partial {k_n}^2} R_n^0 - \frac{\partial^2 P_n^d(k_n, \varepsilon_n)}{\partial {k_n}^2} R_n^1} \tag{3.17}$$

then we have

$$\frac{\partial^2 C(k, \{k_n\}, \{\varepsilon_n\})}{\partial {k_n}^2} > 0 \tag{3.18}$$

which contradicts the assumption that $C(k, \{k_n\}, \{\varepsilon_n\})$ is a concave function with respect to variable k_n. Note that, from (3.15) and (3.16), it can be seen that the right-hand side of (3.17) is within $(0, 1)$. Therefore, there exists $\Pr(\mathcal{H}_n^0)$ such that inequality (3.17) holds.

This completes the proof. $\square$

Besides being nonconvex, problem 3.1 is also a mixed-integer problem, which is usually NP-hard to be solved directly [77]. In order to solve problem 3.1, we resort to transformation of the problem into subproblems with low complexity, as follows.

In the expression of the objective function $C(k, \{k_n\}, \{\varepsilon_n\})$ in problem 3.1, it can be seen that variable k appears only in the term $\frac{T-k\delta}{T}$. So problem 3.1 can be transformed to

$$\begin{aligned} \max_{k} \quad & C(k) = \tfrac{T-k\delta}{T} \cdot U^*(k) \\ \text{s.t.} \quad & 0 < k \leq \lfloor \tfrac{T}{\delta} \rfloor, k \in \mathcal{I} \end{aligned} \tag{3.19}$$

where $U^*(k)$ is the optimal objective value of the following problem with a specific k value:

<u>Problem</u> 3.2. (with a specific k value)

$$\begin{aligned} \max_{\{k_n\},\{\varepsilon_n\}} \quad & U(\{k_n\}, \{\varepsilon_n\}) = \sum_{n=1}^{N} \Big(\Pr(\mathcal{H}_n^0)\big(1 - P_n^f(k_n, \varepsilon_n)\big) R_n^0 \\ & \qquad + \Pr(\mathcal{H}_n^1)\big(1 - P_n^d(k_n, \varepsilon_n)\big) R_n^1 \Big) \\ \text{s.t.} \quad & P_n^d(k_n, \varepsilon_n) \geq P_{th}, n \in \mathcal{N} \\ & \sigma^2 \leq \varepsilon_n \leq \sigma^2(1 + \gamma_n), n \in \mathcal{N} \\ & \sum_{n=1}^{N} k_n = kM \\ & k_n > 0, k_n \in \mathcal{I}, n \in \mathcal{N} \\ & k \in \mathcal{I}. \end{aligned}$$

Problem 3.2 is actually a subproblem of problem 3.1. In the following, we first discuss the properties of problem 3.2 and provide an optimal solution to it in Chapter 3.2.2, and then provide an optimal algorithm to solve problem 3.1 in Chapter 3.2.3 based on the optimal solution of problem 3.2.

3.2.2 Properties and Optimal Solution of Problem 3.2

Lemma 3.2. With the condition $P_{th} > 0.5$, the objective function $U(\{k_n\}, \{\varepsilon_n\})$ in problem 3.2 achieves the maximal value when $P_n^d(k_n, \varepsilon_n) = P_{th}, n \in \mathcal{N}$.

Proof. Denote

$$U_n(k_n, \varepsilon_n) = \Pr(\mathcal{H}_n^0)\big(1 - P_n^f(k_n, \varepsilon_n)\big)R_n^0 + \Pr(\mathcal{H}_n^1)\big(1 - P_n^d(k_n, \varepsilon_n)\big)R_n^1 \tag{3.20}$$

and therefore we have $U(\{k_n\}, \{\varepsilon_n\}) = \sum_{n=1}^{N} U_n(k_n, \varepsilon_n)$.

From equations (3.11) and (3.12), it can be seen that both $\big(1 - P_n^f(k_n, \varepsilon_n)\big)$ and $\big(1 - P_n^d(k_n, \varepsilon_n)\big)$ in $U_n(k_n, \varepsilon_n)$ grow with the increase of ε_n. On the other hand, the term $\big(1 - P_n^d(k_n, \varepsilon_n)\big)$ should be bounded by $1 - P_{th}$. Therefore, $U_n(k_n, \varepsilon_n)$ achieves its maximal value when $\big(1 - P_n^d(k_n, \varepsilon_n)\big)$ reaches its upper bound $(1 - P_{th})$, which happens when $P_n^d(k_n, \varepsilon_n) = P_{th}$.

This completes the proof.

□

Define

$$S(\{k_n\}) = \sum_{n=1}^{N} \Pr(\mathcal{H}_n^0)\left(1 - P_n^f(k_n, P_n^d = P_{th})\right) R_n^0 \tag{3.21}$$

where $P_n^f(k_n, P_n^d)$ is from (3.7), i.e.,

$$P_n^f\left(k_n, P_n^d\right) = Q\left((\gamma_n + 1)Q^{-1}(P_n^d) + \gamma_n\sqrt{\mu k_n \delta}\right). \tag{3.22}$$

Based on Lemma 3.2, we substitute P_n^d with P_{th} in the objective function $U(\{k_n\}, \{\varepsilon_n\})$ in problem 3.2, and we have

$$U(\{k_n\}, \{\varepsilon_n\})|_{P_n^d(k_n,\varepsilon_n)=P_{th}} = S(\{k_n\}) + \sum_{n=1}^{N} \Pr(\mathcal{H}_n^1)(1 - P_{th})R_n^1$$

in which the second term on the right-hand side of the equality is a fixed value.

Consider the constraints in problem 3.2. From equation (3.11), the constraint $\varepsilon_n \leq \sigma^2(1+\gamma_n)$ in problem 3.2 corresponds to the constraint $P_n^d \geq 0.5$, which can be guaranteed by setting $P_n^d = P_{th} > 0.5$. From equation (3.12), the constraint $\varepsilon_n \geq \sigma^2$ in problem 3.2 corresponds to the constraint $P_n^f \leq 0.5$, which can be expressed similarly to equation (3.8) as

$$(\gamma_n + 1)Q^{-1}\left(P_{th}\right) + \gamma_n\sqrt{\mu k_n \delta} \geq 0, n \in \mathcal{N} \tag{3.23}$$

and is equivalent to

$$k_n \geq \left(\frac{-(\gamma_n+1)Q^{-1}(P_{th})}{\sqrt{\delta\mu}\gamma_n}\right)^2, n \in \mathcal{N}. \tag{3.24}$$

Then by defining $z_n = \left\lceil \left(\frac{-(\gamma_n+1)Q^{-1}(P_{th})}{\sqrt{\delta\mu}\gamma_n}\right)^2 \right\rceil$ (where $\lceil \cdot \rceil$ is the ceiling function) and $q_n = k_n - z_n$ ($n \in \mathcal{N}$), and rewriting $S(\{k_n\})$ in the form of $S(\{q_n\})$, problem 3.2 is equivalent to the following problem.

Problem 3.3. (with a specific k value)

$$\begin{aligned}
\max_{\{q_n\}} \quad & S(\{q_n\}) = \sum_{n=1}^{N} \Pr(\mathcal{H}_n^0)\big(1 - P_n^f(q_n, P_n^d = P_{th})\big)R_n^0 \\
\text{s.t.} \quad & \sum_{n=1}^{N} q_n = kM - \sum_{n=1}^{N} z_n \\
& q_n \geq 0, q_n + 1 \in \mathcal{I}, n \in \mathcal{N} \\
& k \in \mathcal{I}
\end{aligned} \tag{3.25}$$

in which

$$P_n^f(q_n, P_n^d = P_{th}) = Q\big((\gamma_n + 1)Q^{-1}(P_{th}) + \sqrt{(q_n + z_n)\delta\mu}\gamma_n\big). \tag{3.26}$$

Note that to make the constraints in problem 3.3 feasible, k should satisfy $k \geq \left\lceil \frac{\sum_{n=1}^{N} z_n}{M} \right\rceil$.

The following lemma is in order.

Lemma 3.3. The function $S(\{q_n\})$ in problem 3.3 is an increasing concave function with respect to q_n ($n \in \mathcal{N}$) within the region $\sum_{n=1}^{N} q_n = kM - \sum_{n=1}^{N} z_n,\ q_n \geq 0$ ($n \in \mathcal{N}$).

Proof. The first order derivative of $S(\{q_n\})$ over q_n is

$$\begin{aligned}
& \frac{\partial S(\{q_n\})}{\partial q_n} \\
= \quad & -\Pr(\mathcal{H}_n^0)R_n^0 \cdot \frac{\partial Q\Big((\gamma_n+1)Q^{-1}(P_{th}) + \sqrt{\delta(q_n+z_n)\mu}\gamma_n\Big)}{\partial q_n} \\
= \quad & \frac{\Pr(\mathcal{H}_n^0)R_n^0}{2\sqrt{2\pi}} e^{-\frac{\left((\gamma_n+1)Q^{-1}(P_{th}) + \sqrt{\delta(q_n+z_n)\mu}\gamma_n\right)^2}{2}} \sqrt{\frac{\mu\delta}{(q_n+z_n)}}\gamma_n \\
> \quad & 0.
\end{aligned} \tag{3.27}$$

So $S(\{q_n\})$ is an increasing function with respect to q_n.

We define $s_n(q_n) \stackrel{\triangle}{=} \Pr(\mathcal{H}_n^0)\big(1 - P_n^f(q_n, P_n^d = P_{th})\big)R_n^0$. Therefore, $S(\{q_n\}) = \sum_{n=1}^{N} s_n(q_n)$. The second order derivative of $s_n(q_n)$ over q_n is given in (3.28).

$$\begin{aligned}\frac{\partial^2 s_n(q_n)}{\partial {q_n}^2} &= -\frac{\Pr(\mathcal{H}_n^0)R_n^0}{4\sqrt{2\pi}} \cdot \left(e^{-\frac{\left((\gamma_n+1)Q^{-1}(P_{th})+\sqrt{\delta(q_n+z_n)\mu}\gamma_n\right)^2}{2}} \sqrt{\frac{\mu\delta}{(q_n+z_n)^3}}\gamma_n + \frac{\mu\delta}{(q_n+z_n)}\gamma_n^2 \right. \\ &\left. \cdot e^{-\frac{\left((\gamma_n+1)Q^{-1}(P_{th})+\sqrt{\delta(q_n+z_n)\mu}\gamma_n\right)^2}{2}} \cdot \left((\gamma_n+1)Q^{-1}(P_{th}) + \sqrt{\delta(q_n+z_n)\mu}\gamma_n\right)\right) \\ &< 0. \end{aligned} \tag{3.28}$$

Note that the inequality in the derivation in (3.28) is based on $\big((\gamma_n+1)Q^{-1}(P_{th}) + \sqrt{\delta(q_n+z_n)\mu}\gamma_n\big) \geq 0$, which is from (3.23) based on the fact that $P_n^f \leq 0.5$.

Define $\boldsymbol{q}^a \stackrel{\triangle}{=} (q_1^a, q_2^a, ..., q_N^a)$ and $\boldsymbol{q}^b \stackrel{\triangle}{=} (q_1^b, q_2^b, ..., q_N^b)$ such that

$$\begin{array}{l} q_n^a \geq 0, q_n^b \geq 0,\ 1 \leq n \leq N; \\ \sum_{n=1}^{N} q_n^a = \sum_{n=1}^{N} q_n^b = kM - \sum_{n=1}^{N} z_n. \end{array} \tag{3.29}$$

For any $\lambda \in [0,1]$, it is easy to see that $\boldsymbol{q}^c \stackrel{\triangle}{=} \lambda \boldsymbol{q}^a + (1-\lambda)\boldsymbol{q}^b$ satisfies

$$q_n^c \geq 0, 1 \leq n \leq N; \quad \sum_{n=1}^{N} q_n^c = kM - \sum_{n=1}^{N} z_n. \tag{3.30}$$

So in the region $\sum_{n=1}^{N} q_n = kM - \sum_{n=1}^{N} z_n, q_n \geq 0\ (n \in \mathcal{N})$, we have

$$\begin{aligned} S(\{q_n^c\}) &= \sum_{n=1}^{N} s_n(q_n^c) \\ &= \sum_{n=1}^{N} s_n(\lambda q_n^a + (1-\lambda)q_n^b) \\ &\overset{\text{from (3.28)}}{\geq} \sum_{n=1}^{N} \lambda s_n(q_n^a) + \sum_{n=1}^{N} (1-\lambda)s_n(q_n^b) \\ &= \lambda S(\{q_n^a\}) + (1-\lambda)S(\{q_n^b\}) \end{aligned} \tag{3.31}$$

which means that the function $S(\{q_n\})$ is concave within the region $\sum_{n=1}^{N} q_n = kM - \sum_{n=1}^{N} z_n, q_n \geq 0\ (n \in \mathcal{N})$.

This completes the proof.

□

From Lemma 3.3, problem 3.3 becomes a convex mixed-integer problem. Generally a mixed-integer problem is NP-hard [77]. However, for problem 3.3, thanks to the separability and concavity of the objective function and the linear constraints with variable coefficients all being 1's, an *incremental algorithm* [78, page 384] can be used to solve problem 3.3 with polynomial complexity upper bound $O(kMN)$, which can converge to the global optimal point as proved in reference [79, pages 53-54].

The procedure of the incremental algorithm for problem 3.3, referred to as algorithm 3.1, is as follows.

Algorithm 3.1 Incremental algorithm for problem 3.3.

1: If $kM - \sum_{n=1}^{N} z_n < 0$, problem 3.3 is infeasible for the given k, return. Otherwise, set $q_n = 0, n \in \mathcal{N}$.
2: If $\sum_{n=1}^{N} q_n < kM - \sum_{n=1}^{N} z_n$, find $n^* = \arg\max_{n \in \mathcal{N}} \big(s_n(q_n + 1) - s_n(q_n)\big)$, and proceed to Step 3; Otherwise, proceed to Step 4.
3: $q_{n^*} = q_{n^*} + 1$, proceed to Step 2.
4: Output $\{q_n, n \in \mathcal{N}\}$.

Since problems 3.2 and 3.3 are equivalent, the optimal objective value of problem 3.2 (with a specific k value), i.e., $U^*(k)$, can be obtained by setting $k_n = q_n + z_n$ and $P_n^d(k_n, \varepsilon_n) = P_{th}, n \in \mathcal{N}$.

Lemma 3.4. $U^*(k)$ has two properties: (1) $U^*(k)$ is an increasing function; (2) $U^*(k) - U^*(k-1) \geq U^*(k+1) - U^*(k)$ for $k \geq \left\lceil \frac{\sum_{n=1}^{N} z_n}{M} \right\rceil, k \in \mathcal{I}$.

The proof of Lemma 3.4 is given in the Appendix A.

3.2.3 Optimal Solution to Problem 3.1

In Chapter 3.2.2, we have provided an optimal solution to problem 3.2, which is a sub-problem of problem 3.1. Now, we can proceed to solve problem 3.1. With the solution of problem 3.2, denoted $U^*(k)$ for a specific k, problem 3.1 is equivalent to

$$\begin{aligned} \max_{k} \quad & C(k) = \tfrac{T-k\delta}{T} \cdot U^*(k) \\ \text{s.t.} \quad & \left\lceil \frac{\sum_{n=1}^{N} z_n}{M} \right\rceil \leq k \leq \lfloor \tfrac{T}{\delta} \rfloor, k \in \mathcal{I}. \end{aligned} \tag{3.32}$$

We have the following lemma for the objective function of this problem.

<u>Lemma</u> 3.5. $C(k+1) - C(k) < C(k) - C(k-1)$, for $\left\lceil \frac{\sum_{n=1}^{N} z_n}{M} \right\rceil < k < \lfloor \frac{T}{\delta} \rfloor, k \in \mathcal{I}$.

Proof.

$$\begin{aligned} & [C(k+1) - C(k)] - [C(k) - C(k-1)] \\ = \ & \left[\left(1 - \frac{(k+1)\delta}{T}\right) U^*(k+1) - \left(1 - \frac{k\delta}{T}\right) U^*(k) \right] \\ & - \left[\left(1 - \frac{k\delta}{T}\right) U^*(k) - \left(1 - \frac{(k-1)\delta}{T}\right) U^*(k-1) \right] \\ = \ & \left(1 - \frac{k\delta}{T}\right) \left[\big(U^*(k+1) - U^*(k)\big) - \big(U^*(k) - U^*(k-1)\big) \right] \\ & + \frac{\delta}{T} \big(U^*(k-1) - U^*(k+1) \big). \end{aligned}$$

From Lemma 3.4, $U^*(k+1) - U^*(k) \leq U^*(k) - U^*(k-1)$, and $U^*(k-1) < U^*(k+1)$. Thus we have

$$[C(k+1) - C(k)] - [C(k) - C(k-1)] < 0.$$

This completes the proof. □

From Lemma 3.5, we see that as k grows, the increase of $C(k)$, denoted $D(k) = C(k) - C(k-1)$, becomes smaller. Then, the optimal point of k, denoted k^*, satisfies

$$D(k^*) \geq 0 \geq D(k^* + 1). \tag{3.33}$$

In the search for k^*, for a round (with a specific k value), the results in previous rounds are still useful, and can be used to reduce the computation complexity in the current round. And the total complexity is upper bounded by $O\left(\lfloor \frac{T}{\delta} \rfloor MN\right)$.

3.3 Optimal Sensing Time Setting in the Continuous-Time Sensing Mode

In this section, we consider the continuous-time sensing mode. The duration of the sensing phase in a slot is denoted τ. So the total sensing time among all the secondary users is $M\tau$. The total sensing time assigned to channel $n \in \mathcal{N}$ is τ_n. Note that τ and τ_n's are all continuous variables. Similar to problem 3.1, we have the following problem formulation.

<u>Problem</u> 3.4.

$$
\begin{aligned}
\max_{\tau,\{\tau_n\},\{\varepsilon_n\}} \quad & C_c = (1-\tfrac{\tau}{T}) \cdot \Big(\sum_{n=1}^{N} \big(\Pr(\mathcal{H}_n^0)\big(1-P_n^f(\tau_n,\varepsilon_n)\big)R_n^0 \\
& \qquad\qquad + \Pr(\mathcal{H}_n^1)\big(1-P_n^d(\tau_n,\varepsilon_n)\big)R_n^1\big)\Big) \\
\text{s.t.} \quad & P_n^d(\tau_n,\varepsilon_n) \geq P_{th}, n \in \mathcal{N} \\
& \sigma^2 \leq \varepsilon_n \leq \sigma^2(1+\gamma_n), n \in \mathcal{N} \\
& \sum_{n=1}^{N} \tau_n = \tau M \\
& 0 < \tau \leq T \\
& \tau_n > 0, n \in \mathcal{N}
\end{aligned}
$$

in which $P_n^f(\tau_n,\varepsilon_n)$ and $P_n^d(\tau_n,\varepsilon_n)$ are given in (3.5) and (3.4), respectively.

Similar to Lemma 3.1, it can be proved that problem 3.4 is not a convex problem. In order to solve it, we transform problem 3.4 to:

$$
\begin{aligned}
\max_{\tau} \quad & C_c(\tau) = (1-\tfrac{\tau}{T}) \cdot U_c^*(\tau) \\
\text{s.t.} \quad & 0 < \tau \leq T.
\end{aligned}
\tag{3.34}
$$

where $U_c^*(\tau)$ is the optimal objective value of the following problem with a specific τ value:

Problem 3.5. (with a specific τ value)

$$\begin{aligned}
\max_{\{\tau_n\},\{\varepsilon_n\}} \quad & U_c(\{\tau_n\},\{\varepsilon_n\}) = \sum_{n=1}^{N} \Big(\Pr(\mathcal{H}_n^0)\big(1 - P_n^f(\tau_n,\varepsilon_n)\big)R_n^0 \\
& \qquad\qquad\qquad\qquad + \Pr(\mathcal{H}_n^1)\big(1 - P_n^d(\tau_n,\varepsilon_n)\big)R_n^1 \Big) \\
\text{s.t.} \quad & P_n^d(\tau_n,\varepsilon_n) \geq P_{th}, n \in \mathcal{N} \\
& \sigma^2 \leq \varepsilon_n \leq \sigma^2(1+\gamma_n), n \in \mathcal{N} \\
& \sum_{n=1}^{N} \tau_n = \tau M \\
& \tau_n > 0, n \in \mathcal{N}.
\end{aligned}$$

Similar to the proof of Lemma 3.2, it can be proved that function $U_c(\{\tau_n\},\{\varepsilon_n\})$ achieves the maximal value when $P_n^d(\tau_n,\varepsilon_n) = P_{th}$.

Define

$$S_c(\{\tau_n\}) = \sum_{n=1}^{N} \Pr(\mathcal{H}_n^0) \left(1 - P_n^f(\tau_n, P_n^d = P_{th})\right) R_n^0 \tag{3.35}$$

where $P_n^f(\tau_n, P_n^d)$ is from (3.7), i.e.,

$$P_n^f\left(\tau_n, P_n^d\right) = Q\left((\gamma_n+1)Q^{-1}(P_n^d) + \gamma_n\sqrt{\mu\tau_n}\right). \tag{3.36}$$

We substitute $P_n^d(\tau_n,\varepsilon_n)$ with P_{th} in the objective function $U_c(\{\tau_n\},\{\varepsilon_n\})$ in problem 3.5, and we have

$$U_c(\{\tau_n\},\{\varepsilon_n\})|_{P_n^d(\tau_n,\varepsilon_n)=P_{th}} = S_c(\{\tau_n\}) + \sum_{n=1}^{N} \Pr(\mathcal{H}_n^1)(1-P_{th})R_n^1$$

in which the second term on the right-hand side of the equality is a fixed value.

Similar to the discussion in Chapter 3.2, the constraint $\varepsilon_n \leq \sigma^2(1+\gamma_n)$ in problem 3.5 is satisfied when $P_n^d = P_{th} > 0.5$, and the constraint $\varepsilon_n \geq \sigma^2$ is equivalent to the following inequality

$$(\gamma_n+1)Q^{-1}\left(P_{th}\right) + \gamma_n\sqrt{\mu\tau_n} \geq 0, n \in \mathcal{N} \tag{3.37}$$

which means

$$\tau_n \geq \left(\frac{-(\gamma_n+1)Q^{-1}(P_{th})}{\sqrt{\mu}\gamma_n}\right)^2, n \in \mathcal{N}. \tag{3.38}$$

Denote $z_n^c = \left(\frac{-(\gamma_n+1)Q^{-1}(P_{th})}{\sqrt{\mu}\gamma_n}\right)^2$. Then problem 3.5 is equivalent to the following problem.

Problem 3.6. (with a specific τ value)

$$\begin{aligned}
\max_{\{\tau_n\}} \quad & S_c(\{\tau_n\}) = \sum_{n=1}^{N} \Big(\Pr(\mathcal{H}_n^0)\big(1 - P_n^f(\tau_n, P_n^d = P_{th})\big) R_n^0 \Big) \\
\text{s.t.} \quad & \sum_{n=1}^{N} \tau_n = \tau M \\
& \tau_n \geq z_n^c, n \in \mathcal{N}.
\end{aligned}$$

It can be easily proved that in the feasible region of problem 3.6 (in which $P_n^f \leq 0.5$ is guaranteed), the objective function of problem 3.6 is a concave function, and therefore, problem 3.6 is a convex problem.

Let $S_c^*(\tau)$ denote the optimal objective value of problem 3.6 with a specific τ value.

Lemma 3.6. Function $S_c^*(\tau)$ is an increasing concave function for $\tau \geq \frac{\sum_{n=1}^{N} z_n^c}{M}$.

Proof. We first prove it is an increasing function.

Define a function

$$s_n^c(\tau_n) = \Pr(\mathcal{H}_n^0)\big(1 - P_n^f(\tau_n, P_n^d = P_{th})\big) R_n^0, \quad \tau_n \geq z_n^c. \tag{3.39}$$

It is easy to prove that function $s_n^c(\tau_n)$ is an increasing function.

For any two variables $\tau^\dagger \geq \frac{\sum_{n=1}^{N} z_n^c}{M}$ and $\tau^\ddagger \geq \frac{\sum_{n=1}^{N} z_n^c}{M}$, assume that $\tau^\dagger < \tau^\ddagger$. For $\tau = \tau^\dagger$, the optimal solution to problem 3.6 is denoted $\{\tau_n^\dagger\}$, with the optimal objective value being $S_c^*(\tau^\dagger)$. For $\tau = \tau^\ddagger$, the optimal solution to problem 3.6 is denoted $\{\tau_n^\ddagger\}$, with the optimal objective value being $S_c^*(\tau^\ddagger)$. Let $\tau_N' = \tau^\ddagger M - \sum_{i=1}^{N-1} \tau_n^\dagger$. Since $\sum_{i=1}^{N} \tau_n^\dagger = \tau^\dagger M$, we have

$$\tau_N' = \tau^\ddagger M - (\tau^\dagger M - \tau_N^\dagger) = (\tau^\ddagger - \tau^\dagger) M + \tau_N^\dagger > \tau_N^\dagger.$$

Since $s_N^c(\tau_N)$ is an increasing function, we have $s_N^c(\tau_N') > s_N^c(\tau_N^\dagger)$, and further we have

$$S_c^*(\tau^\dagger) = \sum_{i=1}^{N-1} s_i^c(\tau_i^\dagger) + s_N^c(\tau_N^\dagger) < \sum_{i=1}^{N-1} s_i^c(\tau_i^\dagger) + s_N^c(\tau_N'). \tag{3.40}$$

On the other hand, $\{\tau_1^\dagger, ..., \tau_{N-1}^\dagger, \tau_N'\}$ is also a feasible solution for problem 3.6 with $\tau = \tau^\ddagger$. Recall that $S_c^*(\tau^\ddagger)$ is the optimal objective value of problem 3.6 when $\tau = \tau^\ddagger$. So we have

$$S_c^*(\tau^\ddagger) \geq \sum_{i=1}^{N-1} s_i^c(\tau_i^\dagger) + s_N^c(\tau_N'). \tag{3.41}$$

From (3.40) and (3.41), we have $S_c^*(\tau^\dagger) < S_c^*(\tau^\ddagger)$, which means that $S_c^*(\tau)$ is an increasing function.

Next we prove that $S_c^*(\tau)$ is a concave function. Denote a vector $\boldsymbol{\nu}= (\tau_1, \tau_2, ..., \tau_N)$, and two functions

$$H_{tc}(\boldsymbol{\nu}) = \sum_{n=1}^{N} \Big(\Pr(\mathcal{H}_n^0)\big(1 - P_n^f(\tau_n, P_n^d = P_{th})\big)R_n^0 \Big) \tag{3.42}$$

and

$$H_c(x) = \max_{\boldsymbol{\nu}\in\mathcal{C}(x)} H_{tc}(\boldsymbol{\nu}) \tag{3.43}$$

where $\mathcal{C}(x) = \{(\tau_1, \tau_2, ..., \tau_N)| \sum_{n=1}^{N} \tau_n = x, \tau_n \geq z_n^c, n \in \mathcal{N}\}$. It is obvious that $H_{tc}(\boldsymbol{\nu})$ is a concave function with respect to $\boldsymbol{\nu}$ when $\tau_n \geq z_n^c, n \in \mathcal{N}$.

For any two values $x_1, x_2 > 0$, and any $\epsilon > 0$, there exist $\boldsymbol{\nu_1} \in \mathcal{C}(x_1)$ and $\boldsymbol{\nu_2} \in \mathcal{C}(x_2)$ such that $H_{tc}(\boldsymbol{\nu_1}) \geq H_c(x_1) - \epsilon$ and $H_{tc}(\boldsymbol{\nu_2}) \geq H_c(x_2) - \epsilon$. For any $\theta \in [0, 1]$, we have

$$\begin{aligned} H_c(\theta x_1 + (1-\theta)x_2) &= \max_{\boldsymbol{\nu}\in\mathcal{C}(\theta x_1+(1-\theta)x_2)} H_{tc}(\boldsymbol{\nu}) \\ &\geq H_{tc}(\theta\boldsymbol{\nu_1} + (1-\theta)\boldsymbol{\nu_2}) \\ &> \theta H_{tc}(\boldsymbol{\nu_1}) + (1-\theta)H_{tc}(\boldsymbol{\nu_2}) \\ &\geq \theta H_c(x_1) + (1-\theta)H_c(x_2) - \epsilon. \end{aligned} \tag{3.44}$$

This inequality holds for any small $\epsilon > 0$. So we have

$$H_c(\theta x_1 + (1-\theta)x_2) > \theta H_c(x_1) + (1-\theta)H_c(x_2) \tag{3.45}$$

which means $H_c(x)$ is a concave function. And therefore, $S_c^*(\tau) = H_c(\tau M)$ is also a concave function [80, page 84] .

This completes the proof. □

With the optimal objective value of problem 3.5, i.e., $U_c^*(\tau)$, problem 3.4 is equivalent to the problem in (3.34) with constraint $\frac{\sum_{n=1}^{N} z_n^c}{M} \leq \tau \leq T$.[4] And the objective function $C_c(\tau)$ of the problem in (3.34) has the following property.

Lemma 3.7. $C_c(\tau)$ is a concave function when $\tau \geq \frac{\sum_{n=1}^{N} z_n^c}{M}$.

Proof. The second-order derivative of $C_c(\tau)$ is given by

$$\frac{d^2 C_c(\tau)}{d\tau^2} = \frac{d^2 S_c^*(\tau)}{d\tau^2}\left(1 - \frac{\tau}{T}\right) + 2 \cdot \left(-\frac{1}{T}\right)\frac{dS_c^*(\tau)}{d\tau}. \tag{3.46}$$

Since the function $S_c^*(\tau)$ is increasing and concave (from Lemma 3.6), we have $\frac{dS_c^*(\tau)}{d\tau} > 0$ and $\frac{d^2 S_c^*(\tau)}{d\tau^2} < 0$. Therefore, $\frac{d^2 C_c(\tau)}{d\tau^2} < 0$, which means that function $C_c(\tau)$ is a concave function. □

From Lemma 3.7, it can be seen that, problem 3.4 is transformed to a *bilevel optimization problem* [78], where the upper level problem (i.e., the problem in (3.34), which optimizes τ) and lower level problem (i.e., problem 3.5 or 3.6, which optimizes $\{\tau_n\}$) are both convex problems. The bilevel problem can be solved by existing methods.

3.4 Performance Evaluation

In this section, numerical results are presented to verify the effectiveness of our proposed algorithms. Similar to [49], the system setup is as follows. The sampling rate is μ= 6 MHz; the slot duration is T= 100 ms; the threshold of detection probability is P_{th} = 0.9; and SNR_n^{ss} between secondary users are all assumed to have a mean 20 dB. Both SNR_n^{ps} (the SNR from the primary user to a secondary user on channel n) and SNR_n^{ss} are assumed to be exponentially distributed.

3.4.1 Verification of Our Algorithms

The slotted-time sensing mode is tested first. Consider $N = 2$ channels, with the available probability as $\Pr(\mathcal{H}_1^0) = 0.8$ and $\Pr(\mathcal{H}_2^0) = 0.6$, and mean SNR from primary user to both secondary users as $\gamma_1 = -15$ dB in channel 1 and $\gamma_2 = -20$ dB in channel 2. There are

[4]Note that when $0 < \tau < \frac{\sum_{n=1}^{N} z_n^c}{M}$, other constraints in the original problem, problem 3.4, cannot be satisfied simultaneously.

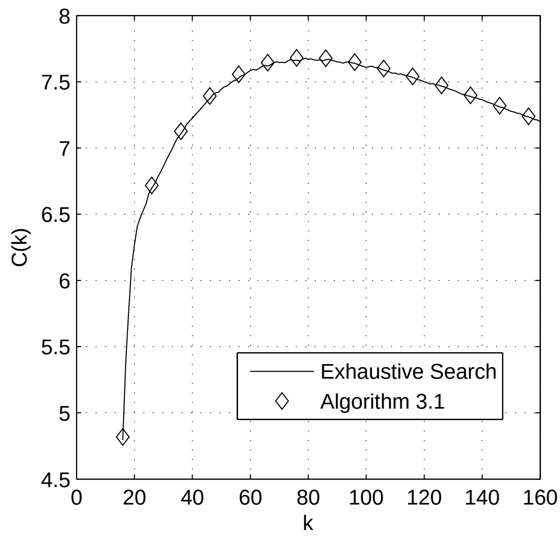

Fig. 3.3. $C(k)$ versus k.

$M = 2$ secondary users. The mini-slot duration is $\delta = 0.1$ ms. To get the value of $C(k)$ (given in (3.19)) at each k, we use two methods: through exhaustive search of the k_n's and ε_n's, and through our algorithm 3.1. The results are demonstrated in Fig. 3.3. It is obvious that the results match well with each other. And the concave-shaped $C(k)$ is also consistent with the conclusion in Lemma 3.5. All these verify the correctness of our algorithm.

For continuous-time sensing mode, similar observations are noticed, and the results are omitted here.

3.4.2 Complexity of Algorithm 3.1

Next, we illustrate the complexity of algorithm 3.1 for slotted-time sensing mode, when the mini-slot duration is 0.01 ms and the available probability of each channel is set as 0.8. The number of iterations to reach the optimal solution is taken as the measure of complexity. In the solving algorithm for slotted-time sensing mode, k is increased from $k_0 \triangleq \left\lceil \frac{\sum_{n=1}^{N} z_n}{M} \right\rceil$ until k^*. Then the number of iterations to reach the optimal solution is given as $(k^* - k_0)\,MN$.

Fig. 3.4 shows the complexity when the number of channels, N, is fixed as 10, while Fig. 3.5 shows the complexity when the number of secondary users, M, is fixed as 10. It can be seen that the number of iterations ($(k^* - k_0)\,MN$) is monotonically increasing

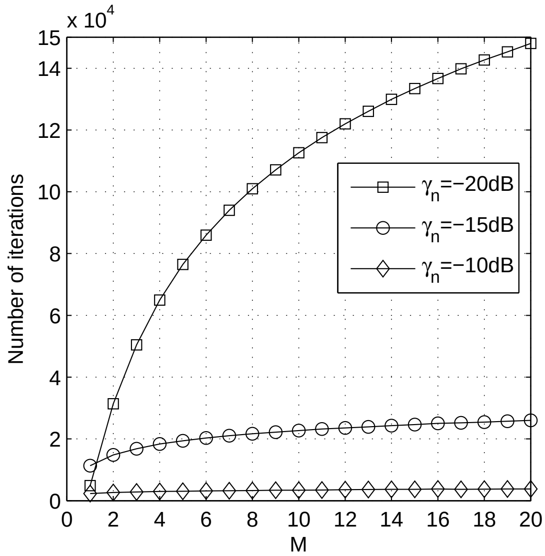

Fig. 3.4. Number of iterations versus M ($N = 10$, $\delta = 0.01$ ms).

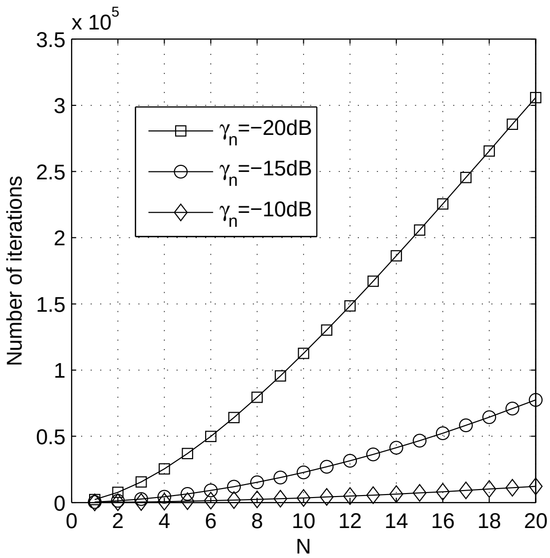

Fig. 3.5. Number of iterations versus N ($M = 10$, $\delta = 0.01$ ms).

with M sublinearly, and with N superlinearly in the demonstrated intervals. This means $(k^* - k_0)$ is monotonically decreasing with respect to M, and monotonically increasing with respect to N in the demonstrated intervals.

Note that Fig. 3.5 does not mean that the number of iterations always increases with N superlinearly. This is because the upper bound of the number of iterations, $\lfloor \frac{T}{\delta} \rfloor MN$, grows linearly with N. So the number of iterations will grow with N sublinearly or even decrease with N when N further increases. Due to space limit in the figure, the results for larger N are not shown here.

3.4.3 Comparison between the Slotted-time Sensing Mode and Continuous-time Sensing Mode

To compare the performance between slotted-time sensing mode and continuous-time sensing mode, consider $N = 5$ channels. These 5 channels have available probabilities as 0.8, 0.7, 0.6, 0.5, and 0.4, with γ_n being -19 dB, -18 dB, -17 dB, -16 dB, and -15 dB, respectively. By running our algorithms for the continuous-time sensing mode, and slotted-time sensing mode with different mini-slot duration δ =0.01 ms, 0.05 ms, 0.1 ms, 0.5 ms, or 1 ms, the obtained optimal network throughput is 17.4 in all cases.

In comparison of slotted-time sensing mode and continuous-time sensing mode, it is difficult to tell which one is better. Generally, it is easier to implement the slotted-time sensing mode in a real system. On the other hand, the complexity in solving the problem with continuous-time sensing mode is lower than that with the slotted-time sensing mode.

Since both sensing modes can reach almost the same optimal network throughput, in the following, we evaluate the performance of slotted-time sensing mode only.

3.4.4 Impact of the Number of Channels and the Number of Users

In the system setup, there are N =5, 9, or 19 channels. Channel n ($1 \leq n \leq N$) has a free probability $\Pr(\mathcal{H}_n^0) = 1 - 0.05 * n$, and the average channel gain from the primary user to a secondary user is $\gamma_n = (-20 + n - 1)$ dB. Fig. 3.6 shows the optimal network throughput (obtained from our algorithm) as a function of the number, M, of secondary users, with different N values. It can be seen that when the number of channels or the number of secondary users increases, the optimal network throughput also increases. Interestingly, with a fixed N value, when the value of M further increases beyond a certain

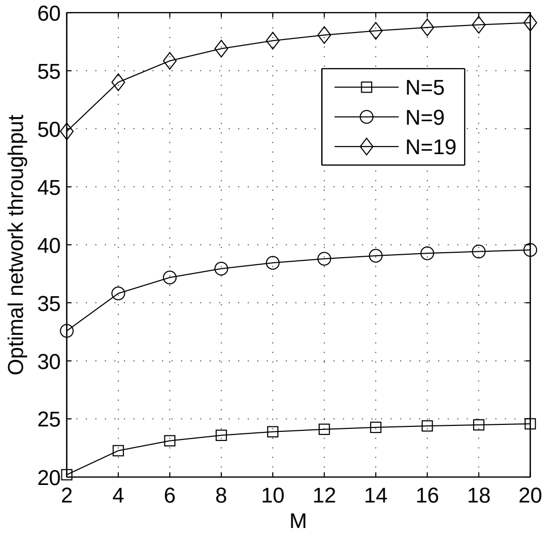

Fig. 3.6. Optimal network throughput versus M and N.

value, the optimal network throughput seems to keep almost constant. The reason is as follows. When M increases, the total sensing time for the channels also increases, which means that a smaller false alarm probability for each channel is expected. When M is large enough, the false alarm probability for each channel is almost zero. Recall that the optimal objective function of problem 3.1 is achieved when the detection probability is exactly the threshold P_{th} (from Lemma 3.2). Therefore, when the false alarm probability of each channel is almost zero, the optimal objective function in problem 3.1 can be approximated as $\sum_{n=1}^{N}\Big(\Pr(\mathcal{H}_n^0)R_n^0 + \Pr(\mathcal{H}_n^1)\,(1-P_{th})\,R_n^1\Big)$, which is a constant. Note that in the approximation, we omit the factor $\frac{T-k\delta}{T}$, since the sensing duration is only a small portion of the total duration T.

3.5 Conclusions

In this chapter, we have explored the optimal multichannel cooperative spectrum sensing strategies in cognitive radio networks. We have studied the problem of how to determine the total sensing time and how to distribute the total sensing time to different channels in cooperative soft-decision spectrum sensing. For the slotted-time sensing mode, we have transformed the initial nonconvex mixed-integer problem into convex mixed-integer subproblems, and provided a polynomial-complexity algorithm to achieve the optimal solution of the initial problem. For the continuous-time sensing mode, we have successfully transformed the initial nonconvex optimization problem into a convex bilevel optimization problem. This research should provide helpful insights into the sensing time configuration in cognitive radio networks.

Chapter 4

Channel Sensing Order Setting in Cognitive Radio Networks: A Two-User Case

This chapter investigates the sensing order problem in two-user multichannel cognitive medium access control. When adaptive modulation is not adopted, although brute force search can be used to find the optimal sensing order setting of the two users, it has huge computational complexity. Accordingly, we propose two sub-optimal algorithms, namely greedy search algorithm and incremental algorithm, which have comparable performance to brute force search and have much less computational complexity. It is shown that, with a high probability, either sub-optimal algorithm can reach an optimal point if a backoff mechanism is used for contention resolution. When adaptive modulation is adopted, it is observed that the traditional stopping rule does not lead to an optimal point in the two-user case. Furthermore, we demonstrate that the adoption of adaptive modulation affects the optimal sensing order setting of the two users, compared with the case without adaptive modulation. These findings imply that the stopping rule and the sensing order setting should be jointly designed in a systematic point of view. [1]

[1] *A version of this chapter has been published in IEEE Trans.Veh. Technol., 58: 4997-5008 (2009).*

4.1 Introduction

A brief review for optimal sensing order and optimal stopping rule in cognitive radio is given as follows. In [43], the optimal channel sensing strategy including the sensing order and the stopping rule is derived for a single-user case, with an assumption that *recall* (i.e., the secondary user can go back to access a previously sensed channel) and *guess* (i.e., the secondary use is permitted to access a channel that has not be sensed yet) are allowed. In [81], it is shown that prohibitive complexity is needed to obtain the optimal channel sensing strategy, and polynomial-complexity algorithms with parameter ϵ are given whose rewards are at most ϵ less than that of the optimal strategy. Recall and guess are permitted. In [56], the optimal channel sensing order problem is investigated for a single-user case, when neither recall nor guess is allowed. It is shown that, in some special scenarios, a simple sensing order does exist. It can be seen that references [43], [56], [81] all focus on the optimality of a single user inside a time frame, assuming that the state (e.g., free/busy, channel capacity, etc.) of a channel is independent of other channels' states and also independent of the states of the same channel in other time slots. In [82], a centralized coordinator is used to discover the spectrum opportunities for a cognitive radio network. The optimal channel sensing order is derived for channels with homogeneous capacities, and it is shown that the problem of the optimal sensing order for channels with heterogeneous capacities is NP-hard. In [54], all the channels have the same free probabilities. So the sensing order problem is not applicable. The optimal stopping rule is derived.

From the above discussion, it can be seen that the optimal sensing order problem and related stopping rule problem are investigated in the literature either for a single-user case or in a centralized system, when only one sensing order (for the single user or for the coordinator) is considered. The problem is still open for a *distributed multi-user case* when the channel sensing and estimation are performed by the users in a distributed manner (i.e., each user has a unique sensing order). As a follow-up of the research in [56] and also as our first step to solve the sensing order problem with multiple users, here we target at a cognitive radio network with two secondary users.

4.2 System Model

Consider a cognitive radio network with a coordinator and two secondary users[2], user 1 and user 2. For the network, there are a number, N, of potential channels, with channel indices 1, 2, ..., N, respectively. Each user can sense one channel at a time. Each user has a sensing order, $(a_1, a_2, ..., a_N)$ for user 1 and $(b_1, b_2, ..., b_N)$ for user 2, which are permutations of $(1, 2, ..., N)$. The *sensing position* of a channel with a user is defined as the position of the channel in the user's sensing order. So channel a_k's (or b_k's) sensing position with user 1 (or user 2) is k.

CAT slot structure is adopted here with fixed total slot duration T. In each time slot, each channel is either occupied by primary activities for the whole slot duration or free of primary activity for the whole slot duration. For channel $n \in \mathcal{N} = \{1, 2, ..., N\}$ in a slot, it is free of primary activity with probability $\theta_n \in (0, 1)$, which is referred to as *primary-free probability* of channel n in the sequel. For each channel, the primary activity state (idle or busy) in a slot is independent of the states in other slots, and also independent of the primary activity state of any other channel in any slot. A similar assumption is also taken in [43], [56], [81]. For simplicity of presentation, we assume $\theta_n \geq \theta_{n+1}, \forall n \in \mathcal{N} \backslash \{N\}$.

A coordinator exists in the cognitive radio network. The coordinator keeps scanning the N channels at all time, and estimates the primary-free probability of each channel. Based on estimated primary-free probabilities, the coordinator determines the sensing orders of the two users, and announces the sensing orders to the two users. Note that the primary-free probabilities may be time-varying. So new sensing orders will be determined and updated to the two users upon significant changes in the estimated primary-free probabilities. In this research, we are interested in how the coordinator determines the sensing orders of the two users, with any specific estimated $\{\theta_1, \theta_2, ..., \theta_N\}$.

In the sensing part of each slot, a user first senses the channels sequentially according to its sensing order until it finds a free channel and transmits in the channel in the remaining time of the slot. Here "a channel is (sensed) free" means that the channel is primary-free and is not occupied by the other user in its prior sensings in the slot. A user decides to access a channel (i.e., starts its transmission if a contention resolution strategy, which will be introduced later in detail, is not adopted, or starts its channel contention procedure if

[2] By "secondary users", we mean secondary transmitters in this chapter.

a contention resolution strategy is adopted), or say the user *stops* at that channel, if the channel is sensed to be free. This procedure is repeated in each of the following slots, i.e., each user repeats sensing channel(s) & choosing a channel to transmit in each slot. No recall or guess is allowed. The time (in a slot) required for sensing a channel is denoted by δ. If a user stops at the k^{th} channel in its sensing order, the length of the sensing phase and transmission phase are $k\delta$ and $T - k\delta$, respectively. The *effectiveness* of a slot is defined as the ratio of the transmission phase length to the slot length. So if a user stops at the k^{th} channel in its sensing order, the effectiveness is

$$c_k = 1 - \frac{k\delta}{T}. \tag{4.1}$$

In this research, each user may sense up to N channels sequentially in a slot. Thus, a short sensing duration for each channel is required. Energy detection is qualified to fulfill this requirement and is used in this research.

If the two users decide to stop at the same channel simultaneously (i.e., the two users sense the same channel in their k^{th} sensing, find the channel is primary-free, and decide to access the channel), one of the following three contention resolution strategies will apply.

- `FAIL_THEN_CONTINUE`: Each user uses a backoff mechanism to avoid possible collision (e.g., similar to the IEEE 802.11 and its variants [83]–[86], each user can pick up a random backoff time before its channel access[3]). One user wins in the contention and transmits in the channel until the end of the slot. The other user (that fails in the contention) continues to sense other channels in the slot according to its sensing order.

- `FAIL_THEN_QUIT`: Similar to `FAIL_THEN_CONTINUE`, the two users use a backoff mechanism to avoid possible collision, and one user wins in the contention and transmits in the channel until the end of the slot. However, the other user (that fails in the contention) quits its subsequent sensings in the slot (i.e., loses its chance to transmit and earns no reward in the slot). Note that compared with `FAIL_THEN_CONTINUE`, `FAIL_THEN_QUIT` is not practical. It is adopted here only as an extreme case.

[3]Similar to the IEEE 802.11, a backoff slot duration can be 20 μs. So if the contention window size is bounded (say ≤ 5), the backoff overhead is also bounded by a value much less than the sensing duration. This means that backoff does not cause significant overhead.

- `COLLIDE`: No backoff mechanism is applied. When the two users decide to stop at the same channel simultaneously, they both transmit in the channel until the end of the slot, and thus, a collision happens. In other words, no user can earn rewards in the slot due to the collision.

If one user decides to stop at a channel, it may select a transmission rate with or without adaptive modulation. Without adaptive modulation, the transmission rate of each user is fixed, denoted by R. With adaptive modulation, the transmitter can select a rate according to its SNR in the channel.

In the following, the case without adaptive modulation is discussed first, in Chapters 4.3-4.6, then the case with adaptive modulation is investigated in Chapter 4.7.

4.3 Some Examples without Adaptive Modulation

If adaptive modulation is not adopted, a user will decide to stop at a channel if the channel is sensed free. We define the reward of a user as the information bits it can transmit at a slot. So if a user stops at its k^{th} channel in its sensing order, the reward is

$$U_k = R \cdot \left(1 - \frac{k\delta}{T}\right). \tag{4.2}$$

The system throughput is the summation of the two users' rewards. And our objective is to find an optimal sensing order setting of the two users so as to achieve the maximal system throughput.

In this section, we use some examples to demonstrate the factors that affect the optimal sensing order setting of the two users. In the examples, a cognitive radio network with $N = 4$ potential channels is considered. The fraction of sensing time for each channel in a slot is $\frac{\delta}{T} = 0.1$.

Example 1: The primary-free probabilities of the four channels are $\theta_1 = 0.9$, $\theta_2 = 0.8$, $\theta_3 = 0.7$, and $\theta_4 = 0.6$, respectively. `FAIL_THEN_QUIT` is the contention resolution strategy.

By brute force search, we can obtain the optimal sensing order setting of the two users, which is (1, 4, 3, 2) for user 1 and (2, 3, 4, 1) for user 2. This setting is reasonable because the four primary-free probabilities are comparable to each other. So if a user picks up a channel to sense, the other user will pick up another.

Example 2: The primary-free probabilities of the four channels are $\theta_1 = 0.9$, $\theta_2 = 0.8$, $\theta_3 = 0.7$, and $\theta_4 = 0.1$, respectively. `FAIL_THEN_QUIT` is the contention resolution strategy.

By brute force search, the optimal sensing order setting is (1, 3, 4, 2) for user 1 and (2, 3, 4, 1) for user 2, which is different from that in Example 1. This is because θ_4 is much less than θ_3. Consider the moment when user 2 is selecting[4] its second channel to sense (among channels 1, 3, and 4, since channel 2 has been selected as the first channel in the sensing order). User 2 knows[5] that the first two channels in user 1's sensing order are 1 and 3, respectively. From user 2's point of view at its second sensing, we observe the following:

- Channel 1 is always busy, since channel 1 should be occupied by either primary users or user 1. Therefore, if channel 1 is selected as user 2's second sensing, no reward is earned;

- Consider the case that user 2 selects channel 3 as its second sensing: when both channel 1 and channel 3 are primary-free (with probability $\theta_1\theta_3$), user 2's reward at its second sensing is U_2; when channel 1 is primary-busy and channel 3 is primary-free (with probability $(1-\theta_1)\theta_3$), user 2's reward at its second sensing is $0.5U_2$ (where the term 0.5 means the probability that user 2 wins the contention with user 1 in channel 3); the reward is 0 otherwise. On the other hand, when channel 1 is primary-busy and channel 3 is primary-free, user 1 has a reward loss (compared with the case when user 2 does not select channel 3 in its second sensing), given by $(1-\theta_1)\theta_3 \cdot 0.5 \cdot U_2$. Therefore, if user 2 selects channel 3 in its second sensing, the *additional* reward to the system is given by $\theta_1\theta_3U_2 + (1-\theta_1)\theta_3 \cdot 0.5U_2 - (1-\theta_1)\theta_3 \cdot 0.5 \cdot U_2 = 0.63U_2$;

- Channel 4 is free with probability $\theta_4 = 0.1$. If user 2 selects channel 4 in its second sensing, the additional reward to the system is $0.1U_2$.

So user 2 selects the channel with the maximal additional reward to the system, i.e., channel 3. Both users select channel 3 as the second channel in their sensing orders. If the two users both proceed to sense channel 3, one will succeed in the contention and the other will quit

[4]For presentation simplicity, when we say "a user selects a channel", we mean that the coordinator selects the channel for the user.

[5]For presentation simplicity, here "user 2 knows" actually means that "the coordinator knows".

subsequent sensings in the slot (i.e., lose the chance to sense other channels in the slot). However, the probability of this event is quite small, i.e., $(1-\theta_1)(1-\theta_2) = 0.02$ in the example. Therefore, considering the probabilities and rewards/costs of all events, it is optimal that both users select channel 3 as their second channel to sense. In the next section, a detailed procedure is given, regarding how to model the cost when the two users contend in the same channel, and how to determine the sensing orders of the two users.

Example 3: The primary-free probabilities of the four channels are the same as those in Example 2, i.e., $\theta_1 = 0.9$, $\theta_2 = 0.8$, $\theta_3 = 0.7$, and $\theta_4 = 0.1$. However, `COLLIDE` is the contention resolution strategy.

By brute force search, the optimal sensing order setting is (1, 4, 3, 2) for user 1 and (2, 3, 4, 1) for user 2, which is different from that in Example 2. If the optimal sensing order setting in Example 2 is applied, although the probability that both users simultaneously proceed to sense channel 3 is small, the cost of this event (i.e., no user will gain any reward in the slot) is significant. A mathematical model for the cost is given in next section. So the optimal sensing order setting in Example 2 is not optimal anymore when `COLLIDE` is the contention resolution strategy.

From the examples, it can be seen that both the primary-free probabilities of the channels and the contention resolution strategy will affect the optimal sensing order setting.

Generally, brute force search can be used to find an optimal sensing order setting of the two users, but with significant complexity. If the complexity to calculate the average system throughput with a specific sensing order setting is $O(1)$, then the complexity of brute force search is $O\big((N!)^2\big)$. Our research targets at algorithms that have much less complexity and have comparable performance to brute force search, as discussed in the following two sections.

4.4 Greedy Search Algorithm

From Example 2 in the preceding section, we have the following observations. For a target user to determine which channel to sense in its k^{th} sensing, when a channel has already been in the other user's prior ($\leq k$) sensing positions, the target user should estimate the probability that the channel is accessed by the other user. From Examples 2 and 3, it can be seen that the cost should be considered when the two users both proceed to sense the same channel at the same time. Based on these observations, we propose a sub-optimal algorithm, namely *greedy search algorithm*, to select the sensing orders of the two users, i.e., $\boldsymbol{A}$=$(a_1, a_2, ..., a_N)$ for user 1 and $\boldsymbol{B}$=$(b_1, b_2, ..., b_N)$ for user 2.

The greedy search algorithm consists of N rounds. In the k^{th} round, we determine a_k and b_k. When user 1 (or 2) is selecting a_k (or b_k) among channels not in its prior sensing positions, the user estimates the probability that each channel is free of primary activities and has not been accessed by the other user. The user also assigns each channel a reward accordingly. Then the channel with the maximal reward is selected as a_k (or b_k). The detailed selecting procedure is as follows.

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ denote the channel selection vectors of users 1 and 2, respectively, which will be updated after each round. Let $\mathcal{A}$ and $\mathcal{B}$ denote the sets of channels in $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively. So after the k^{th} round, $l(\boldsymbol{A})$ = $l(\boldsymbol{B})$ = $|\mathcal{A}|$ = $|\mathcal{B}|$ = k. Note that $\mathcal{A}$ and $\mathcal{B}$ have the same elements as $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively. The difference is that $\mathcal{A}$ and $\mathcal{B}$ are sets, while $\boldsymbol{A}$ and $\boldsymbol{B}$ are vectors. $l(\cdot)$ means the length of a vector.

4.4.1 Round-1 Procedure

At the beginning, set $\mathcal{A} = \mathcal{B} = \emptyset$ (empty set). We first select a_1 for user 1. We need to select a channel from $\bar{\mathcal{A}} = \mathcal{N} \backslash \mathcal{A}$. For each channel $n \in \bar{\mathcal{A}}$, we denote $G_n^{(1)}(1)$ as the reward user 1 can add to the system through sensing channel n in its first sensing, which is given by $G_n^{(1)}(1) = \theta_n U_1$, where U_1 is given in (4.2). In the sequel, we use superscript $(\cdot)$ to represent user index, and subscript to represent channel index (except for c_i and U_i, where the subscript i stands for the i^{th} sensing in a slot).

Then the first channel to be sensed by user 1 is selected as

$$a_1 = \arg\max_{n \in \bar{\mathcal{A}}} G_n^{(1)}(1)$$

and after that we have $\boldsymbol{A} = (a_1)$, and $\mathcal{A} = \{a_1\}$.

Next we select b_1 for user 2. We need to select a channel from $\bar{\mathcal{B}} = \mathcal{N} \backslash \mathcal{B}$. If user 2 also selects a_1, it does not add any reward to the system (compared with the case when users 2 does not select a_1), because user 1 has already selected a_1. Therefore, we have

$$b_1 = \arg\max_{n \in \bar{\mathcal{B}} \backslash \{a_1\}} G_n^{(2)}(1)$$

and after that we have $\boldsymbol{B} = (b_1)$, and $\mathcal{B} = \{b_1\}$.

4.4.2 Round-k ($k \in \{2, 3, ..., N-1\}$) Procedure

At round $k \in \{2, 3, ..., N-1\}$, we should determine which user we should start with. To achieve a certain level of fairness[6], we start with the user that has less cumulative reward in the prior $k-1$ rounds, which is given by

$$\theta_{a_1} U_1 + (1-\theta_{a_1})\theta_{a_2} U_2 + \cdots + \left(\Pi_{i=1}^{k-2}(1-\theta_{a_i})\right)\theta_{a_{k-1}} U_{k-1}$$

for user 1, or

$$\theta_{b_1} U_1 + (1-\theta_{b_1})\theta_{b_2} U_2 + \cdots + \left(\Pi_{i=1}^{k-2}(1-\theta_{b_i})\right)\theta_{b_{k-1}} U_{k-1}$$

for user 2.

Without loss of generality, assume user 1 has a less cumulative reward, and thus we should first select a_k from $\bar{\mathcal{A}}$, then select b_k from $\bar{\mathcal{B}}$.

In the sequel, for each user, primary-free probability of a channel means the probability that the channel is free of primary activity, and free probability of a channel means the probability that the channel is free of either primary activity or the other secondary user's activity.

4.4.2.1 Selection of a_k

To select a_k, we need to know the free probability (from user 1's viewpoint) of each channel $n \in \bar{\mathcal{A}}$. If $n \notin \mathcal{B}$, then channel n's free probability is the same as its primary-free probability, i.e., θ_n. However, if $n \in \mathcal{B}$, channel n is free for user 1 at its k^{th} sensing only when there is no primary activity in channel n (with probability θ_n) and user 2 does not proceed to sense channel n. In order to know the probability that user 2 does not proceed to sense channel n, we first update the primary-free probability of each channel in $\mathcal{B}$. Note that for each channel n, θ_n is its initially-known primary-free probability. When user 1 proceeds to sense a_k, we have more information that user 1 does not stop at channel a_1,a_2,..., or a_{k-1}.

[6] It is hard to achieve strict fairness between the two users in a slot, since the throughput maximization and strict fairness are two conflicting objectives in a short duration [87], [88]. However, The fairness of the two users can be guaranteed among a number of slots. This can be achieved, for example, through switching the sensing orders of the two users after a short period.

The information can provide a more accurate estimation of the primary-free probabilities of the channels. We use the following example to explain.

As an example, assume we have $N = 6$ channels with initially known primary-free probability for channel n being θ_n. And we have determined that $(a_1, a_2, a_3) = (1, 3, 4)$ and $(b_1, b_2, b_3) = (2, 4, 3)$. Now we need to determine a_4. When user 1 proceeds to sense a_4, we know that it has already sensed channel 3 but has not stopped at channel 3. Further, if user 2 also senses channel 3, it should be later than user 1's sensing of channel 3, because channel 3 is the second in user 1's sensing order, and the third in user 2's sensing order. These two facts determine that channel 3 should be occupied by primary activities, i.e., $\theta_3^* = 0$ in this example. We use superscript * to denote the updated primary-free probability.

The detailed primary-free probability update procedure, as shown in the Appendix B, is repeated from channel b_1 to b_{k-1}. In the Appendix B, $g(n, \boldsymbol{B})$ is used to denote the position of n in $\boldsymbol{B}$, if channel n is in $\boldsymbol{B}$. In other words, channel n is at the $g(n, \boldsymbol{B})^{th}$ position in $\boldsymbol{B}$.

Then for a channel $n \in \bar{\mathcal{A}}$, we have the following.

- If $n \notin \mathcal{B}$, the free probability of channel n is $\theta_n^{\text{free}} = \theta_n$. So the reward if user 1 selects channel n in its k^{th} sensing is $G_n^{(1)}(k) = \theta_n^{\text{free}} U_k = \theta_n U_k$.
- If $n \in \mathcal{B}$, the free probability of channel n is $\theta_n^{\text{free}} = \theta_n^* \big(1 - \Pi_{j=1}^{g(n,\boldsymbol{B})-1}(1 - \theta_{b_j}^*)\big)$, where the second term on the right hand side means the probability that user 2 does not proceed to sense channel n. So the reward if user 1 selects channel n in its k^{th} sensing is

$$G_n^{(1)}(k) = \theta_n^{\text{free}} U_k = \Big[\theta_n^* \big(1 - \Pi_{j=1}^{g(n,\boldsymbol{B})-1}(1 - \theta_{b_j}^*)\big)\Big] U_k. \tag{4.3}$$

Then the selection of the k^{th} sensing of user 1 is given by

$$a_k = \arg\max_{n \in \bar{\mathcal{A}}} G_n^{(1)}(k). \tag{4.4}$$

4.4.2.2 Selection of b_k

To select b_k, the procedure is similar to that for a_k. The only difference lies in the reward of user 2 also selecting channel a_k in its k^{th} sensing, if $a_k \notin \mathcal{B}$. We have three possible situations, which are listed as follows.

- `FAIL_THEN_CONTINUE` is adopted as the contention resolution strategy: Similar to (4.3), we use the following equation to calculate the reward of user 2 selecting channel a_k as its k^{th} channel to sense

$$G^{(2)}_{a_k}(k) = \theta^*_{a_k}\big(1 - \Pi_{j=1}^{k-1}(1-\theta^*_{a_j})\big)U_k. \tag{4.5}$$

- `FAIL_THEN_QUIT` is adopted as the contention resolution strategy: Consider the scenario when user 2 selects channel a_k as its k^{th} channel to sense. If user 1 also proceeds to sense its k^{th} channel in its sensing order (i.e., a_k), and channel a_k is primary-free, then user 1 and 2 will contend for the channel access, and thus one will succeed and the other will quit subsequent channel sensings in the slot. So if user 2 also selects a_k as its k^{th} channel to sense, compared with the situation that user 2 selects another channel, user 2 will lose the chance to obtain reward via sensing the subsequent channels (i.e., channels $b_{k+1}, b_{k+2}, ..., b_N$) in its sensing order. This loss of reward is approximated with an assumption that the subsequent channels in the sensing order of user 2 are with descending order of their free probabilities, i.e., $\theta^{\text{free}}_{b_{k+1}} \geq \theta^{\text{free}}_{b_{k+2}} \geq ... \geq \theta^{\text{free}}_{b_N}$. Then the loss of reward is approximated by

$$L(a_k) = \sum_{i=k+1}^{N} \big(\Pi_{j=k+1}^{i-1}(1-\theta^{\text{free}}_{b_j})\big)\theta^{\text{free}}_{b_i} U_i. \tag{4.6}$$

So the reward in (4.5) should be rewritten for user 2 selecting channel a_k as its k^{th} channel to sense:

$$G^{(2)}_{a_k}(k) = \theta^*_{a_k}\Big[\big(1 - \Pi_{j=1}^{k-1}(1-\theta^*_{a_j})\big)U_k - \big(\Pi_{j=1}^{k-1}(1-\theta^*_{a_j})\big)L(a_k)\Big] \tag{4.7}$$

where the factor $\Pi_{j=1}^{k-1}(1-\theta^*_{a_j})$ for $L(a_k)$ means the probability that user 1 does not stop at channel a_1, a_2..., or a_{k-1}.

- `COLLIDE` is adopted as the contention resolution strategy: Consider the scenario when user 2 selects channel a_k as its k^{th} channel to sense. If user 1 also proceeds to sense its k^{th} channel in its sensing order (i.e., a_k), and channel a_k is primary-free, then a collision will happen. So if user 2 also selects a_k as its k^{th} channel to sense, compared with the situation that user 2 selects another channel, user 1 will lose its reward at channel a_k and user 2 will lose the chance to obtain reward via sensing the subsequent channels in its sensing order. So the loss of reward in (4.6) should be modified to

$$L(a_k) = U_k + \sum_{i=k+1}^{N} \left(\Pi_{j=k+1}^{i-1}(1-\theta_{b_j}^{\text{free}})\right)\theta_{b_i}^{\text{free}} U_i \tag{4.8}$$

and equation (4.7) is still used to calculate the reward of user 2 selecting a_k as its k^{th} channel to sense.

After a_k and b_k are selected, they are added into $\boldsymbol{A}$, $\boldsymbol{B}$, $\mathcal{A}$, and $\mathcal{B}$. Then $\boldsymbol{A} = (a_1, a_2, ..., a_k)$, $\boldsymbol{B} = (b_1, b_2, ..., b_k)$, $\mathcal{A} = \{a_1, a_2, ..., a_k\}$, and $\mathcal{B} = \{b_1, b_2, ..., b_k\}$.

4.4.3 Round-N Procedure

In the last round, the only element in $\mathcal{N}\backslash\mathcal{A}$ is selected as a_N, and the only element in $\mathcal{N}\backslash\mathcal{B}$ is selected as b_N.

4.4.4 Complexity of the Greedy Search Algorithm

In the k^{th} round of the preceding procedure, there are $(N-k+1)$ rewards to calculate. To calculate each reward, the primary-free probability need to be calculated with maximal complexity $O(k-1)$. So the complexity of the greedy search algorithm is upper-bounded by $O(\sum_{k=1}^{N}(N-k+1)(k-1))$, i.e., $O(N^3)$, which is much less than the complexity in brute force search.

4.5 The Incremental Algorithm

It is observed in many scenarios that, if we remove the channel with the smallest primary-free probability from an optimal sensing order setting of the two users and keep other channels' relative positions in each sensing order unchanged, the remaining sensing order setting is still optimal for the remaining channels. Accordingly, we propose a sub-optimal algorithm, namely *incremental algorithm*, as follows. Here “incremental" means we gradually increase the size of the set of potential channels until the set size denoted by k reaches N.

We start with $k = 2$. The set of potential channels to be examined is $\{1, 2\}$, since θ_1 and θ_2 are the two largest primary-free probabilities. By brute force search, we can find an optimal sensing order setting, $\boldsymbol{A}_2 = (a_1, a_2)$ for user 1 and $\boldsymbol{B}_2 = (b_1, b_2)$ for user 2. Here the subscript in $\boldsymbol{A}_2$ and $\boldsymbol{B}_2$ means the round index. The computation complexity of this round is $O(2^2)$.

In the round with $k = 3$, from the previous round we already know the sensing order setting with two channels that have the two largest primary-free probabilities (i.e., channels 1 and 2). As an example, assume $\boldsymbol{A}_2 = (1, 2)$ for user 1 and $\boldsymbol{B}_2 = (2, 1)$ for user 2. We need to add channel 3 into $\boldsymbol{A}_2$ and $\boldsymbol{B}_2$, while we keep the relative positions of channels 1 and 2 in $\boldsymbol{A}_2$ and $\boldsymbol{B}_2$ unchanged. After channel 3 is added, the new sensing orders are $\boldsymbol{A}'$ and $\boldsymbol{B}'$ for users 1 and 2, respectively. $\boldsymbol{A}'$ has three possible cases: (3,1,2), (1,3,2), and (1,2,3), while $\boldsymbol{B}'$ also has three possible cases: (3,2,1), (2,3,1), and (2,1,3). So we have $3^2 = 9$ possible combinations of $(\boldsymbol{A}', \boldsymbol{B}')$. Among them, the combination with the maximal average system throughput is selected as $(\boldsymbol{A}_3, \boldsymbol{B}_3)$. The computational complexity in this round is $O(3^2)$.

In the round with $k \in \{4, ..., N\}$, from previous rounds we already know the sensing order setting with $k - 1$ channels that have the $k - 1$ largest primary-free probabilities, denoted by $\boldsymbol{A}_{k-1} = (a_1, a_2, ..., a_{k-1})$ for user 1 and by $\boldsymbol{B}_{k-1} = (b_1, b_2, ..., b_{k-1})$ for user 2. Both $\boldsymbol{A}_{k-1}$ and $\boldsymbol{B}_{k-1}$ are permutations of $(1, 2, ..., k - 1)$. We need to add channel k into $\boldsymbol{A}_{k-1}$ and $\boldsymbol{B}_{k-1}$, while we keep the relative positions of existing elements in $\boldsymbol{A}_{k-1}$ and $\boldsymbol{B}_{k-1}$ unchanged. After channel k is added, the new sensing orders are $\boldsymbol{A}'$ and $\boldsymbol{B}'$ for users 1 and 2, respectively. Either $\boldsymbol{A}'$ or $\boldsymbol{B}'$ has k possible cases. Then we have k^2 possible combinations of $(\boldsymbol{A}', \boldsymbol{B}')$. Among them, the combination with the maximal average system throughput is selected as $(\boldsymbol{A}_k, \boldsymbol{B}_k)$. The computational complexity in this round is $O(k^2)$.

The above procedure is repeated until $k = N$. Then $\boldsymbol{A}_N$ and $\boldsymbol{B}_N$ are the sensing orders of the two users, with N potential channels.

The computational complexity of the incremental algorithm is thus $O(\sum_{k=2}^{N} k^2)$, i.e., $O(N^3)$, which is much less than the complexity of brute force search.

4.6 Numerical Results

To evaluate the performance of the proposed greedy search and incremental algorithms, we consider $N = 7$ potential channels with $\frac{\delta}{T} = 0.1$. The primary-free probability of channel n (i.e., θ_n) is given by $\theta_n = 0.05 \cdot \kappa_n$, where κ_n is an integer. We get the numerical results for the following three tests:

- Test I: all the scenarios with κ_n's being even integers, $2 \leq \kappa_1 \leq 18$, $2 \leq \kappa_2 \leq \kappa_1$, $2 \leq \kappa_3 \leq \kappa_2$, and $0 \leq \kappa_n \leq \kappa_{n-1}$ for $n = 4, 5, 6, 7$. Note that in this test, all θ_n's are within [0, 0.9];

- Test II: all the scenarios with $10 \leq \kappa_n \leq 18$, and $\kappa_1 \geq \kappa_2 \geq ... \geq \kappa_7$. Note that in this test, all θ_n's are within [0.5, 0.9];

- Test III: all the scenarios with $1 \leq \kappa_n \leq 10$, and $\kappa_1 \geq \kappa_2 \geq ... \geq \kappa_7$. Note that in this test, all θ_n's are within [0.05, 0.5].

In each scenario of each test, we obtain the maximal average system throughput through brute force search (for an optimal sensing order setting), denoted by T_{opt}, and the average system throughput of the greedy search and incremental algorithms, denoted by T_{greedy} and $T_{\text{incremental}}$, respectively. We also obtain the relative difference of the average system throughput of the greedy search and incremental algorithms from the maximal average system throughput, defined as $\frac{T_{\text{opt}}-T_{\text{greedy}}}{T_{\text{opt}}}$ and $\frac{T_{\text{opt}}-T_{\text{incremental}}}{T_{\text{opt}}}$, respectively. It is observed that all the relative difference values in all scenarios are below 5%. Table 4.1 shows the percentage of the relative difference values being in different intervals and the mean and standard deviation (Std) of the relative difference values when the contention resolution strategy is `FAIL_THEN_CONTINUE`, `FAIL_THEN_QUIT`, or `COLLIDE`. It can be seen that, with `FAIL_THEN_CONTINUE`, the greedy search algorithm is very likely (with probability 93.02% in Test I, 71.62% in Test II, and 63.91% in Test III) to achieve an optimal solution, while the likelihood is decreased with `FAIL_THEN_QUIT` or `COLLIDE`. This is because of the approximation used to estimate the loss of reward in a contention. On the other hand, the incremental algorithm with `FAIL_THEN_CONTINUE` can achieve an optimal solution with probability 99.64% in Test I, 97.31% in Test II, and 91.79% in Test III. The likelihood is decreased slightly when `FAIL_THEN_QUIT` is the contention resolution strategy. This is because in each round of the incremental algorithm, the relative positions of the channels determined in prior rounds do not change. So it is likely that a channel may have the same sensing position in the two users' sensing orders. And thus, a user will fail in the contention and lose the reward of subsequent sensings if both users proceed to sense the channel simultaneously and the channel is primary-free. The reward loss determines that the incremental algorithm may not always be optimal. When `COLLIDE` is adopted for the incremental algorithm, the loss of reward is much more severe, and thus the incremental algorithm only has

TABLE 4.1
PERCENTAGE OF THE RELATIVE DIFFERENCE VALUES AND THE MEAN AND STANDARD DEVIATION (STD).

Test	Relative difference		0	(0, 0.1%]	(0.1%, 1%]	(1%, 5%]	Mean	Std
I	FAIL_THEN_CONTINUE	Greedy Search	93.02%	6.73%	0.25%	0	0.0015%	0.0097%
		Incremental	99.64%	0.36%	0	0	0.000037%	0.0008%
	FAIL_THEN_QUIT	Greedy Search	64.47%	21.55%	13.47%	0.50%	0.0541%	0.1565%
		Incremental	88.86%	10.87%	0.26%	0	0.0022%	0.0112%
	COLLIDE	Greedy Search	51.33%	33.23%	13.63%	1.81%	0.0881%	0.2810%
		Incremental	16.50%	28.63%	54.59%	0.29%	0.1733%	0.1873%
II	FAIL_THEN_CONTINUE	Greedy Search	71.62%	28.38%	0	0	0.0005%	0.0016%
		Incremental	97.31%	2.69%	0	0	0.000062 %	0.00064%
	FAIL_THEN_QUIT	Greedy Search	53.55%	42.88%	3.57%	0	0.014%	0.056%
		Incremental	85.67%	14.33%	0	0	0.0007%	0.0027%
	COLLIDE	Greedy Search	47.94%	50.67 %	1.37%	0.02 %	0.0096%	0.051%
		Incremental	10.97%	87.85%	1.18%	0	0.022%	0.023%
III	FAIL_THEN_CONTINUE	Greedy Search	63.91%	32.93%	3.16%	0	0.015%	0.031%
		Incremental	91.79%	7.80%	0.41%	0	0.0024%	0.013%
	FAIL_THEN_QUIT	Greedy Search	54.37%	33.88%	11.09%	0.66%	0.045%	0.14%
		Incremental	84.49%	12.19%	3.32%	0	0.0088%	0.029%
	COLLIDE	Greedy Search	37.99%	40.78%	21.23%	0	0.057%	0.081%
		Incremental	0.48%	1.86%	94.95%	2.71%	0.42%	0.23%

a small probability to achieve an optimal solution. However, since almost all the relative difference values in all the scenarios are below 1%, the performance of the incremental algorithm is still promising when `COLLIDE` is the contention resolution strategy.

In summary, generally both algorithms work very well with all the three contention resolution strategies. As a comparison, both algorithms have the same order of computational complexity. The incremental algorithm can achieve slightly larger system throughput with `FAIL_THEN_CONTINUE` and `FAIL_THEN_QUIT`, while the greedy search algorithm leads to better throughput performance with `COLLIDE`.

4.7 Impact of Adaptive Modulation

If a secondary user knows its channel quality (i.e., SNR) to its receiver, adaptive modulation can be adopted so as to utilize the time-varying feature of the channels. The secondary user can adapt its transmission rate according to the channel quality. For channel n, if a user senses it free and the SNR of the user is SNR_n, then the achievable transmission rate of the user is $f(\text{SNR}_n)$. Here $f(\cdot)$ is a non-descending function mapping SNR to the transmission rate.

By adaptive modulation, a secondary user can opportunistically select a free channel with good channel quality to transmit, referred to as *multichannel diversity*. In multichannel diversity, additional overhead is needed to estimate the channel gain, e.g., small control packets may be sent between the sender and receiver. Despite the overhead, it is shown in [89] that multichannel diversity can result in an overall benefit. In the open literature, multichannel diversity has been well exploited in cognitive radio related research efforts [43], [56], [81], [90].

In Chapter 4.7.4, a scheme will be given for a secondary user to estimate its channel SNR to its receiver. With adaptive modulation, when a user senses a free channel with poor channel quality, it may skip this channel, with expectation that it may have better quality in other channels [56], [90]. To determine when a user should stop at a channel (i.e., the channel is free and has satisfactory SNR), similar to the research in [56], [90], an SNR threshold is assigned for the channel. When the user proceeds to sense the channel, if the channel is free and the channel SNR is above the threshold, the user will stop at the channel (i.e., start its contention in the channel); otherwise, the user will proceed to sense the next channel in its sensing order. Therefore, the configuration of each user includes: a sensing order and an SNR threshold for each channel in its sensing order. For user 1, its sensing order is $(a_1, a_2, ..., a_N)$ with SNR thresholds $(\Gamma^{(1)}_{a_1}, \Gamma^{(1)}_{a_2}, ..., \Gamma^{(1)}_{a_N})$. For user 2, its sensing order is $(b_1, b_2, ..., b_N)$ with SNR thresholds $(\Gamma^{(2)}_{b_1}, \Gamma^{(2)}_{b_2}, ..., \Gamma^{(2)}_{b_N})$.

In order to find an optimal system configuration, two questions need to be answered: 1) what is the optimal stopping rule? That is, how should a user determine the SNR threshold of a channel in its sensing order? 2) Does the adoption of adaptive modulation affect the optimal sensing order setting of the two users (compared to the case without adaptive modulation)? The two questions are investigated as follows.

4.7.1 Stopping Rule

Consider that a user proceeds to sense its k^{th} channel in its sensing order, say channel n, which is free. If the user has an SNR value denoted by SNR_n, its instantaneous reward at the channel is $c_k f(\text{SNR}_n)$ times the probability that it will win the contention (if any) in the channel. Here c_k is given in (4.1). The user also has an expected reward if it skips channel n and proceeds to sense the next channel in its sensing order. It may seem plausible and intuitive that the user should stop at channel n if the instantaneous reward of the channel is larger than the expected reward of skipping the channel. This stopping rule is referred to as the *traditional stopping rule* in the following.

Next we investigate whether the traditional stopping rule is optimal in a single-user case and a two-user case, respectively.

4.7.1.1 Optimality of the Traditional Stopping Rule in a Single-User Case

The traditional stopping rule has been used in single-user multichannel medium access control in the literature [56], [90]. Here, we prove the optimality of the traditional stopping rule in the single-user case.

Consider a (secondary) user with N potential channels. Channel n is primary-free with probability θ_n. The probability density function (PDF) of channel n's SNR is denoted by $h_n(\gamma)$ for SNR value γ. The SNR of each channel is independent of any other channel. The user's sensing order is $(a_1, a_2, ..., a_N)$, which is a permutation of $(1, 2, ..., N)$. For channel a_k, the SNR threshold is Γ_{a_k}. Particularly, for the last channel in the sensing order, the SNR threshold is 0, i.e., $\Gamma_{a_N} = 0$. If the user proceeds to sense the k^{th} channel in the sensing order, i.e., a_k, the user will transmit at the channel if $\text{SNR}_{a_k} \geq \Gamma_{a_k}$, where SNR_{a_k} is the instantaneous SNR in channel a_k. No recall is allowed. This means if the user skips a primary-free channel, it is not allowed to return to transmit at that channel.

If the user proceeds to sense channel a_N, its instantaneous reward if channel a_N is primary-free is given by $w_{a_N} = c_N f(\text{SNR}_{a_N})$. And the reward expectation of sensing channel a_N is $W_{a_N} = \theta_{a_N}\mathbb{E}[w_{a_N}]$.

When the user proceeds to sense its i^{th} ($i \leq N-1$) channel in its sensing order, i.e., channel a_i, its reward (if the channel is primary-free) is

$$w_{a_i} = \begin{cases} c_i f(\text{SNR}_{a_i}), \text{ if } \text{SNR}_{a_i} > \Gamma_{a_i} (\text{stop at channel } a_i) \\ W_{a_{i+1}}, \text{ otherwise (proceed to sense channel } a_{i+1}) \end{cases} \tag{4.9}$$

where $W_{a_{i+1}}$ ($i \leq N-1$) is the expected reward if the user proceeds to sense channel a_{i+1}, given by

$$W_{a_{i+1}} = \begin{cases} \theta_{a_{i+1}}\mathbb{E}[w_{a_{i+1}}] + (1-\theta_{a_{i+1}})W_{a_{i+2}}, \text{if } i < N-1 \\ \theta_{a_{i+1}}\mathbb{E}[w_{a_{i+1}}], \qquad\qquad\qquad \text{if } i = N-1. \end{cases} \tag{4.10}$$

From (4.9) and (4.10), we can obtain the values of all W_{a_i}'s recursively, starting from W_{a_N} until W_{a_1} [56], [90]. W_{a_1} is the average throughput of the user with the N potential channels.

Lemma 4.1. With a specific sensing order setting $(a_1, a_2, ..., a_N)$, the traditional stopping rule is optimal. This means the optimal threshold setting $(\Gamma_{a_1}, \Gamma_{a_2}, ..., \Gamma_{a_N})$ satisfies $c_i f(\Gamma_{a_i}) = W_{a_{i+1}}, 1 \le i \le N-1$ and $\Gamma_{a_N} = 0$.

Proof. If the user proceeds to sense its last channel in its sensing order, i.e., a_N, it should transmit if the channel is sensed free. So $\Gamma_{a_N} = 0$.

From (4.9) and (4.10), $\forall i \le N-1$, we have

$$W_{a_1} = \left\{ \sum_{l=1}^{i} \left[\Pi_{k=1}^{l-1} (1 - \theta_{a_k} \int_{\Gamma_{a_k}}^{\infty} h_{a_k}(\gamma) d\gamma) \right] \theta_{a_l} \cdot c_l \int_{\Gamma_{a_l}}^{\infty} f(\gamma) h_{a_l}(\gamma) d\gamma \right\} + \left[\Pi_{k=1}^{i} \left(1 - \theta_{a_k} \int_{\Gamma_{a_k}}^{\infty} h_{a_k}(\gamma) d\gamma \right) \right] \cdot W_{a_{i+1}}. \tag{4.11}$$

To achieve the optimal W_{a_1} value, we should have $\frac{\partial W_{a_1}}{\partial \Gamma_{a_i}} = 0$. We get

$$\left[\Pi_{k=1}^{i-1} \left(1 - \theta_{a_k} \int_{\Gamma_{a_k}}^{\infty} h_{a_k}(\gamma) d\gamma \right) \right] \cdot \left[\theta_{a_i} c_i f(\Gamma_{a_i}) h_{a_i}(\Gamma_{a_i}) - \theta_{a_i} h_{a_i}(\Gamma_{a_i}) W_{a_{i+1}} \right] = 0$$

and further we have $c_i f(\Gamma_{a_i}) = W_{a_{i+1}},\ 1 \le i \le N-1$. This completes the proof. $\square$

4.7.1.2 Stopping Rule for a Two-User Case

When two users, user 1 and user 2, are considered, the traditional stopping rule may not be optimal anymore. We use the following example to show the difference of the traditional stopping rule from an optimal stopping rule for a two-user case.

Consider the situation with two channels, channel 1 and channel 2, with primary-free probabilities $\theta_1 > 0$ and $\theta_2 > 0$. The PDF of user i's SNR in channel j is $h_j^{(i)}(\gamma) > 0$, for $i, j \in \{1, 2\}, \gamma \in (0, \infty)$. User 1 has a sensing order $(1, 2)$ and associated SNR thresholds $(\Gamma_1^{(1)}, \Gamma_2^{(1)})$. User 2 has a sensing order $(2, 1)$ and associated SNR thresholds $(\Gamma_2^{(2)}, \Gamma_1^{(2)})$. If user 1 proceeds to channel 2 and senses it free (which means channel 2 is primary-free and user 2 does not stop at channel 2 in its first sensing), user 1 should transmit in channel 2 with whatever SNR. Thus, $\Gamma_2^{(1)} = 0$. Similarly we have $\Gamma_1^{(2)} = 0$. Therefore, we only need to determine $\Gamma_1^{(1)}$ and $\Gamma_2^{(2)}$.

Traditional Stopping Rule:

If user 1 proceeds to sense channel 2, it can transmit only when channel 2 is primary-free and user 2 does not stop at channel 2 at its first sensing. Thus the reward expectation of user 1 proceeding to sense channel 2 is

$$W_2^{(1)} = \theta_2 \cdot \int_0^{\Gamma_2^{(2)}} h_2^{(2)}(\gamma)d\gamma \cdot c_2 \int_0^{\infty} f(\gamma)h_2^{(1)}(\gamma)d\gamma \tag{4.12}$$

where the second term on the right hand side is the probability that user 2 does not stop at channel 2 at its first sensing. So according to the traditional stopping rule, we have

$$c_1 f(\Gamma_1^{(1)}) = W_2^{(1)} = \theta_2 \cdot \int_0^{\Gamma_2^{(2)}} h_2^{(2)}(\gamma)d\gamma \cdot c_2 \int_0^{\infty} f(\gamma)h_2^{(1)}(\gamma)d\gamma. \tag{4.13}$$

Similarly, we have

$$c_1 f(\Gamma_2^{(2)}) = W_1^{(2)} = \theta_1 \cdot \int_0^{\Gamma_1^{(1)}} h_1^{(1)}(\gamma)d\gamma \cdot c_2 \int_0^{\infty} f(\gamma)h_1^{(2)}(\gamma)d\gamma. \tag{4.14}$$

Optimal Stopping Rule:

For user 1, its throughput is given by

$$\begin{aligned} W_1^{(1)} = \theta_1 c_1 \int_{\Gamma_1^{(1)}}^{\infty} f(\gamma)h_1^{(1)}(\gamma)d\gamma + \left(1 - \theta_1 \int_{\Gamma_1^{(1)}}^{\infty} h_1^{(1)}(\gamma)d\gamma\right) \\ \cdot\theta_2 \cdot \int_0^{\Gamma_2^{(2)}} h_2^{(2)}(\gamma)d\gamma \cdot c_2 \int_0^{\infty} f(\gamma)h_2^{(1)}(\gamma)d\gamma. \end{aligned}$$

For user 2, its throughput is given by

$$\begin{aligned} W_2^{(2)} = \theta_2 c_1 \int_{\Gamma_2^{(2)}}^{\infty} f(\gamma)h_2^{(2)}(\gamma)d\gamma + \left(1 - \theta_2 \int_{\Gamma_2^{(2)}}^{\infty} h_2^{(2)}(\gamma)d\gamma\right) \\ \cdot\theta_1 \cdot \int_0^{\Gamma_1^{(1)}} h_1^{(1)}(\gamma)d\gamma \cdot c_2 \int_0^{\infty} f(\gamma)h_1^{(2)}(\gamma)d\gamma. \end{aligned}$$

To achieve the optimal system throughput, we should have

$$\frac{\partial(W_1^{(1)} + W_2^{(2)})}{\partial\Gamma_1^{(1)}} = 0$$

and

$$\frac{\partial(W_1^{(1)} + W_2^{(2)})}{\partial\Gamma_2^{(2)}} = 0.$$

Then we get

$$
\begin{aligned}
c_1 f(\Gamma_1^{(1)}) &= \theta_2 \cdot \int_0^{\Gamma_2^{(2)}} h_2^{(2)}(\gamma)d\gamma \cdot c_2 \int_0^{\infty} f(\gamma)h_2^{(1)}(\gamma)d\gamma \\
&\quad + \left(1 - \theta_2 \int_{\Gamma_2^{(2)}}^{\infty} h_2^{(2)}(\gamma)d\gamma\right) c_2 \int_0^{\infty} f(\gamma)h_1^{(2)}(\gamma)d\gamma, \\
c_1 f(\Gamma_2^{(2)}) &= \theta_1 \cdot \int_0^{\Gamma_1^{(1)}} h_1^{(1)}(\gamma)d\gamma \cdot c_2 \int_0^{\infty} f(\gamma)h_1^{(2)}(\gamma)d\gamma \\
&\quad + \left(1 - \theta_1 \int_{\Gamma_1^{(1)}}^{\infty} h_1^{(1)}(\gamma)d\gamma\right) c_2 \int_0^{\infty} f(\gamma)h_2^{(1)}(\gamma)d\gamma
\end{aligned}
$$

which are apparently different from (4.13) and (4.14) for the Traditional Stopping Rule.

4.7.2 Impact of Adaptive Modulation on Optimal Sensing Order Setting

The answer to the question whether the adoption of adaptive modulation affects the optimal sensing order setting of the two users is essential. This is because if the answer is “no", then the derivation of an optimal system configuration (including the sensing order setting and SNR threshold setting) with adaptive modulation can be divided into two steps: one step to find an optimal sensing order setting without adaptive modulation, and the other step to select an optimal SNR threshold setting when adaptive modulation is adopted. However, unfortunately, the answer to that question is “yes". We use the following example to demonstrate that an optimal sensing order setting without adaptive modulation may not be optimal when adaptive modulation is adopted.

Consider two users, user 1 and 2, with two channels, channel 1 and 2. The primary-free probabilities of the two channels are $\theta_1 \gg 0$ and $\theta_2 = 0^+$, respectively. Here 0^+ means an infinitely small positive value. The PDF of the SNR of each user in each channel is common, denoted by $h(\gamma)$. `FAIL_THEN_QUIT` is the contention resolution strategy. So when adaptive modulation is not adopted, an optimal sensing order setting is (1,2) for user 1 and (2,1) for user 2, referred to as the DIFF_ORDER setting. When adaptive modulation is adopted with the sensing order setting being the DIFF_ORDER setting, we denote the optimal SNR threshold setting as $(\Gamma_1^{(1)}, \Gamma_2^{(1)})$ for user 1, and $(\Gamma_2^{(2)}, \Gamma_1^{(2)})$ for user 2. When user 2(1) proceeds to sense channel 1(2) in its second sensing, its SNR threshold should be 0 because if user 2(1) senses channel 1(2) free, channel 1(2) should be primary-free and user 1(2) should have skipped channel 1(2) in its first sensing due to unsatisfactory SNR.

TABLE 4.2
RATIO OF $c_1 \int_{\Gamma_1^{(1)}}^{\infty} f(\gamma)h(\gamma)d\gamma$ TO $c_2 \int_0^{\infty} f(\gamma)h(\gamma)d\gamma$.

$\frac{\delta}{T}$	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45
Ratio	0.79	0.89	1.01	1.16	1.36	1.64	2.09	2.95	5.48

Therefore, optimal SNR threshold setting should be $(\Gamma_1^{(1)}, 0)$ for user 1, and $(\Gamma_2^{(2)}, 0)$ for user 2. We have the following lemma.

Lemma 4.2. $c_1 f(\Gamma_1^{(1)}) = c_2 \int_0^{\infty} f(\gamma)h(\gamma)d\gamma$.

Proof. The system throughput is given by

$$T_{\text{DIFF}} = \theta_1 \cdot c_1 \int_{\Gamma_1^{(1)}}^{\infty} f(\gamma)h(\gamma)d\gamma + \theta_1 \cdot \int_0^{\Gamma_1^{(1)}} h(\gamma)d\gamma \cdot c_2 \int_0^{\infty} f(\gamma)h(\gamma)d\gamma. \quad (4.15)$$

On the right hand side of (4.15), the first term means user 1 transmits in channel 1 upon its first sensing if channel 1 is primary-free and user 1's SNR in channel 1 is above the threshold, and the second term means user 2 transmits in channel 1 upon its second sensing if channel 1 is primary-free and user 1 skips channel 1 in its first sensing.

For the optimal SNR threshold setting, we have

$$\frac{\mathrm{d}T_{\text{DIFF}}}{\mathrm{d}\Gamma_1^{(1)}} = 0$$

which can lead to

$$c_1 f(\Gamma_1^{(1)}) = c_2 \int_0^{\infty} f(\gamma)h(\gamma)d\gamma.$$

This completes the proof. □

Lemma 4.3. The DIFF_ORDER setting is not optimal if $c_1 \int_{\Gamma_1^{(1)}}^{\infty} f(\gamma)h(\gamma)d\gamma > c_2 \int_0^{\infty} f(\gamma)h(\gamma)d\gamma$.

Proof. We consider a new sensing order setting where both user 1 and 2 use (1,2) as the sensing order, referred to as the SAME_ORDER setting. With the SAME_ORDER setting, we let the SNR threshold setting be $(\Gamma_1^{(1)}, 0)$ for both users, where $\Gamma_1^{(1)}$ is user 1's optimal SNR threshold in channel 1 with the DIFF_ORDER setting (as given in Lemma 4.2).

The system throughput of the SAME_ORDER setting is

$$T_{\text{SAME}} = \theta_1 \bigg(c_1 \int_{\Gamma_1^{(1)}}^{\infty} h(\gamma) d\gamma \int_{\Gamma_1^{(1)}}^{\infty} f(\gamma) h(\gamma) d\gamma + 2 \cdot c_1 \int_0^{\Gamma_1^{(1)}} h(\gamma) d\gamma \int_{\Gamma_1^{(1)}}^{\infty} f(\gamma) h(\gamma) d\gamma \bigg). \tag{4.16}$$

On the right hand side of (4.16), the first term in the bracket means both users have SNR in channel 1 above the threshold (i.e., one user wins the contention and transmits, and the other user fails and quits), and the second term means one user has SNR in channel 1 above the threshold (i.e., stops at channel 1) while the other user has SNR in channel 1 below the threshold (i.e., skips channel 1 and proceeds to the next channel). After some mathematical manipulation, we have

$$T_{\text{SAME}} = \theta_1 \cdot c_1 \int_{\Gamma_1^{(1)}}^{\infty} f(\gamma) h(\gamma) d\gamma + \theta_1 \cdot c_1 \int_0^{\Gamma_1^{(1)}} h(\gamma) d\gamma \int_{\Gamma_1^{(1)}}^{\infty} f(\gamma) h(\gamma) d\gamma. \tag{4.17}$$

From (4.15) and (4.17) we have

$$T_{\text{SAME}} - T_{\text{DIFF}} = \theta_1 \int_0^{\Gamma_1^{(1)}} h(\gamma) d\gamma \cdot \left(c_1 \int_{\Gamma_1^{(1)}}^{\infty} f(\gamma) h(\gamma) d\gamma - c_2 \int_0^{\infty} f(\gamma) h(\gamma) d\gamma \right)$$

which means if $c_1 \int_{\Gamma_1^{(1)}}^{\infty} f(\gamma) h(\gamma) d\gamma > c_2 \int_0^{\infty} f(\gamma) h(\gamma) d\gamma$, we have $T_{\text{SAME}} > T_{\text{DIFF}}$ (i.e., the DIFF_ORDER setting is not optimal). This completes the proof. $\square$

Consider the Shannon's channel capacity as the achievable transmission rate, i.e., $f(\gamma) = \log(1+\gamma)$, and Rayleigh fading channels, i.e., $h(\gamma) = \frac{1}{\Gamma} e^{-\frac{\gamma}{\Gamma}}$ (where $\Gamma = 10$ is the average SNR in a channel as an example). Table 4.2 shows the ratio of $c_1 \int_{\Gamma_1^{(1)}}^{\infty} f(\gamma) h(\gamma) d\gamma$ to $c_2 \int_0^{\infty} f(\gamma) h(\gamma) d\gamma$ for different values of $\frac{\delta}{T}$. It can be seen that when $\frac{\delta}{T} \geq 0.15$, we have $c_1 \int_{\Gamma_1^{(1)}}^{\infty} f(\gamma) h(\gamma) d\gamma > c_2 \int_0^{\infty} f(\gamma) h(\gamma) d\gamma$, and thus the SAME_ORDER setting is the optimal sensing order setting.

4.7.3 Discussion

In the preceding subsections, we have two observations for a two-user cognitive radio network: 1) when the sensing order setting of the two users is fixed, the traditional stopping rule may not lead to an optimal SNR threshold setting and 2) the adoption of adaptive modulation may affect the optimal sensing order setting of the two users. These observations imply that the sensing order setting and the SNR threshold setting should be jointly determined together from a systematic viewpoint to achieve the maximal system throughput. Although in this research, a detailed algorithm to derive the optimal system configuration with adaptive modulation is not given, the two observations should be able to pave the way for further research efforts.

4.7.4 Implementation Issues

To achieve adaptive modulation in cognitive radio networks, two challenges are experienced: hidden terminal problem and channel SNR estimation problem.

- Hidden terminal problem: Consider a scenario with a primary transmitter and its receiver, and a secondary transmitter and its receiver. The primary transmitter is active, and its location is far away from the secondary transmitter but close to the secondary receiver. Then the secondary transmitter may not sense the signal from the primary transmitter and, thus, transmits its own signal. However, the secondary receiver cannot successfully receive this signal due to strong interference from the primary transmitter. This is the well known hidden terminal problem.

- Channel SNR estimation problem: In adaptive modulation, the secondary transmitter needs to know the channel SNR to its receiver before its information transmission.

To solve the preceding problems, we propose to let the secondary receiver instead of the secondary transmitter perform the sensing task. And in the CAT time slot structure a probing time with duration Δ is inserted after each sensing duration δ. It might be difficult to give a specific value of Δ, since the value should depend on the channel bandwidth, acquisition technique, modulation scheme, etc. In general, an appropriate Δ value should ensure that the transmission phase still occupies a significant portion in a time slot.

At the beginning of each time slot, the secondary receiver first senses its first channel (in its sensing order) for a duration of δ.

- If the channel is sensed free, the secondary receiver transmits a probe in the following probing duration Δ. The probe is transmitted using spread spectrum technology. Specifically, each secondary transmitter is assigned a unique spreading code. The probe sent from the secondary receiver is spread by the spreading code of its secondary transmitter. The secondary transmitter keeps scanning its own spreading code and, thus, can receive the probe successfully. Through the reception of the probe, the secondary transmitter can estimate the channel SNR from the secondary receiver to itself. This SNR is also the estimated channel SNR from the secondary transmitter to the secondary receiver, as channel reciprocity is assumed here.

- If the channel is sensed busy, the secondary receiver skips the following probing time Δ, and then starts to sense its second channel in its sensing order.

The above procedure is repeated in the sensing phase of each slot.

As the sensing task is performed by the secondary receiver, the preceding hidden terminal problem does not exist. And through the channel probing from the secondary receiver, the channel SNR can be estimated at the secondary transmitter side.

4.8 Conclusion and Further Discussions

In cognitive radio networks, one objective is to achieve the maximal spectrum efficiency. In this chapter, we have demonstrated our research efforts to achieve the maximal spectrum efficiency through selecting of sensing orders in a two-user cognitive radio network. For the case without adaptive modulation, low-complexity algorithms have been proposed that have comparable performance to an optimal solution, and particularly, have a high probability to reach an optimal point when a backoff mechanism is used. For the case with adaptive modulation, our observations imply that the sensing order setting and stopping rule should be jointly designed from a systematic point of view.

Sensing is assumed to be perfect in this research. In reality, sensing errors are inevitable. A channel sensing error is either a missed detection (i.e., the channel is occupied by primary users, but the secondary user senses it to be idle) or a false alarm (i.e., the channel is idle, but the secondary user senses it to be busy). Let P_m and P_f denote the missed detection probability and false alarm probability, respectively. Then the probability that a secondary user senses channel n to be idle is given by $\theta_n(1-P_f)+(1-\theta_n)P_m$. If a secondary user senses an idle channel and transmits in the channel, its probability of successful transmission (i.e., not colliding with possible primary activities) is $\frac{\theta_n(1-P_f)}{\theta_n(1-P_f)+(1-\theta_n)P_m}$. Generally speaking, for the greedy search algorithm, the sensing errors will affect the estimation of channel free probabilities and the reward calculation, while the basic principle in the algorithm is still to select the channel with the largest additional reward to the system. For the incremental algorithm, the average throughput calculation in each round will be affected. The detailed derivations are omitted here due to similarity.

Chapter 5

Joint Optimal Cooperative Sensing and Resource Allocation in Multichannel Cognitive Radio Networks

The reported work in Chapter 3 only considers optimal configuration of sensing time while supposing a fixed transmission rate. In other words, the gain of dynamic resource allocation is not exploited. In addition, it is assumed that each user can sense or access only one channel at a time. Recently wideband sampling technique has received some initial investigation [44], which makes wideband spectrum sensing and access possible. All these have motivated us to investigate the joint optimization of spectrum sensing duration configuration and resource allocation with wideband spectrum sensing and access. In this research, the problem of joint multichannel cooperative sensing and resource allocation with wideband spectrum sensing and access is investigated. A cognitive radio network with multiple potential channels and multiple secondary users is considered. Each secondary user carries out wideband spectrum sensing as introduced in Chapter 2.2.1 to get a test statistic for each channel, and transmits the test statistic to a coordinator. The coordinator performs the data fusion cooperative spectrum sensing to detect the presence of primary user on each channel. When a channel is estimated to be free, secondary users can get access to the channel with assigned bandwidth and power. An optimization problem is formulated, which maximizes the weighted sum of secondary users' throughputs while guaranteeing a certain level of protection for the activities of primary users. Although the problem is nonconvex, it is shown that the problem can be solved by bilevel optimization and monotonic programming. This research is also extended to cases with consideration of proportional fairness and max-min fairness. [1]

5.1 System Model

A cognitive radio network is considered with N frequency bands and M secondary communication pairs (with the m^{th} pair including secondary transmitter m and secondary receiver m). Each channel has a bandwidth of W, and is licensed to one primary user.

CAT slot structure is employed with fixed total slot length T. In each slot, the primary user in a channel is assumed to be either active or idle for the whole slot. The channels among primary users and secondary users are assumed to keep unchanged within each time slot. In other words, block-fading model is assumed, and the channel coherent time is assumed to be longer than the slot duration T. This assumption is reasonable when all the primary users and secondary users in the system are static or in low-speed mobility.

The duration of the sensing phase in each CAT slot is denoted τ, which is a parameter to be optimized. In the sensing phase of every CAT slot, each secondary user senses the signal spanning over the N channels with the wideband spectrum sensing method in frequency domain introduced in Chapter 2.2.1 at a sampling rate μ. Then the secondary users report to a coordinator their test statistics in the sensing phase and their channel gains. Based on these, the coordinator determines how the secondary users can get access to the channels and broadcasts the resource allocation decision to secondary users. The information exchange between secondary users and the coordinator can be finished on a common control channel, for example, on the industrial, scientific and medical (ISM) radio bands. By limiting the number of secondary users involved [46], the time overhead for signaling interchange between the secondary users and coordinator is assumed to be negligible.

[1] *A version of this chapter has been published in IEEE Trans.Veh. Technol., 60: 722-729 (2011).*

5.1.1 Spectrum Sensing

By wideband sampling technique as introduced in Chapter 2.2.1, for secondary transmitter m, its received primary signal at channel n in frequency domain, $Y_{n,m}$, can be given in the following binary test hypothesis

$$\begin{aligned} \mathcal{H}_n^0: &\quad Y_{n,m} = W_{n,m}, \\ \mathcal{H}_n^1: &\quad Y_{n,m} = H_{n,m}S_n + W_{n,m}, \end{aligned} \tag{5.1}$$

where $\mathcal{H}_n^0$ and $\mathcal{H}_n^1$ mean the primary user on channel n is idle and busy, respectively, and $W_{n,m}$, $H_{n,m}$, and S_n are N-point FFT of the additive complex white Gaussian noise (with mean being zero and variance being σ^2), the discrete-time channel impulse response, and the primary signal, respectively. For presentation simplicity, when primary signal is present, the power spectrum density of the primary signal on each channel at the transmitter side is normalized to be 1, i.e., $\mathbb{E}\{|S_n|^2\} = 1$.

Following the method introduced in Chapter 2.2.1, the test statistic for channel n of secondary user m is given as

$$T_{n,m} \triangleq \frac{N}{\mu\tau} \sum_{k=1}^{\mu\tau/N} |Y_{n,m}(k)|^2. \tag{5.2}$$

Data fusion cooperative spectrum sensing is used. The coordinator collects test statistics $T_{n,m}$'s from all secondary users, and makes estimation whether the primary signal in channel n exists by comparing $\sum_{m=1}^{M} T_{n,m}$ with a threshold ε_n. The detection probability and false alarm probability for channel n are given, respectively, by following the way in [44] as

$$Q_n^d(\tau, \varepsilon_n) = Q\left(\frac{\big(\varepsilon_n - (\sigma^2 M + ||\boldsymbol{H}_n||_1)\big)\sqrt{\mu\tau}}{\sigma\sqrt{2||\boldsymbol{\Sigma}_n||_1 N}}\right) \tag{5.3}$$

$$Q_n^f(\tau, \varepsilon_n) = Q\left(\frac{(\varepsilon_n - \sigma^2 M)\sqrt{\mu\tau}}{\sigma^2\sqrt{2MN}}\right) \tag{5.4}$$

where $\boldsymbol{H}_n = \big(|H_{n,1}|^2, |H_{n,2}|^2, ..., |H_{n,M}|^2\big)^T$ (here superscript T means transpose operation), $\boldsymbol{\Sigma}_n = \sigma^2\mathbf{I} + 2\text{Diag}(\boldsymbol{H}_n)$ (here $\mathbf{I}$ means an identity matrix, and $\text{Diag}(\boldsymbol{H}_n)$ is a diagonal matrix formed by elements in $\boldsymbol{H}_n$), and $||\cdot||_1$ of a matrix represents its entrywise 1-norm, which is the summation of absolute values of all the elements. Based on (5.3) and

(5.4), Q_n^f can be expressed in terms of Q_n^d as

$$Q_n^f(\tau, Q_n^d) = Q\left(\frac{Q^{-1}(Q_n^d)\sigma\sqrt{2||\mathbf{\Sigma}_n||_1 N} + ||\boldsymbol{H}_n||_1\sqrt{\mu\tau}}{\sigma^2\sqrt{2MN}}\right). \tag{5.5}$$

Similar with the case in Chapter 3, the false alarm probability should be no larger than 0.5. For a given value of Q_n^d, constraint $Q_n^f \leq 0.5$ is equivalent to the following constraint

$$Q^{-1}(Q_n^d)\sigma\sqrt{2||\mathbf{\Sigma}_n||_1 N} + ||\boldsymbol{H}_n||_1\sqrt{\mu\tau} \geq 0 \tag{5.6}$$

which is from (5.5).

5.1.2 Resource Allocation

Recall that the coherence time of the channels (from primary users to secondary users, and among secondary users) is assumed to be relative long. Within the duration of coherence time, the resource allocation strategy does not change. For the m^{th} secondary user pair in channel n, denote channel gain from the primary user to the secondary receiver and from the secondary transmitter to the secondary receiver as $|g_{n,m}^{ps}|^2$ and $|g_{n,m}^{ss}|^2$, respectively.

In the resource allocation, each secondary user is assigned a portion of each channel at a certain transmission power. In specific, secondary user m is assigned a portion $x_{n,m}(0 \leq x_{n,m} \leq 1)$ of channel n at a transmission power $p_{n,m}$. If channel n is estimated by the coordinator to be free, then secondary user m can transmit in the assigned portion of channel n with the power. It is assumed that the channel gain in channel n between the m^{th} secondary transceiver pair, i.e., $|g_{n,m}^{ss}|^2$, is a constant within the bandwidth of channel n. Therefore, it does not make any difference which portion in channel n is assigned to secondary user m.

When channel n is estimated (by the coordinator) to be free, one of two scenarios happens:

- Channel n is indeed free (i.e., the estimation is correct), with probability $\Pr(\mathcal{H}_n^0)\big(1 - Q_n^f(\tau, \varepsilon_n)\big)$, where $\Pr(\mathcal{H}_n^0)$ is the free probability of channel n. The achievable transmission rate of secondary user m is

$$r_{n,m}^0 = x_{n,m} \cdot \log\left(1 + \frac{|g_{n,m}^{ss}|^2 p_{n,m}}{W x_{n,m}\sigma^2}\right).$$

- Channel n is actually busy (i.e., the estimation is wrong), with probability $\Pr(\mathcal{H}_n^1)\big(1-Q_n^d(\tau,\varepsilon_n)\big)$, where $\Pr(\mathcal{H}_n^1) = 1 - \Pr(\mathcal{H}_n^0)$ is the busy probability of channel n. In this scenario, the primary signal serves as interference to secondary transmission. The achievable transmission rate of secondary user m is

$$r_{n,m}^1 = x_{n,m} \cdot \log\left(1 + \frac{|g_{n,m}^{ss}|^2 p_{n,m}}{x_{n,m} W(|g_{n,m}^{ps}|^2 + \sigma^2)}\right).$$

Then the average throughput of secondary user m is given as

$$R_m = \left(1 - \frac{\tau}{T}\right) \cdot \sum_{n=1}^{N} \Big[\Pr(\mathcal{H}_n^0)\big(1 - Q_n^f(\tau,\varepsilon_n)\big) \cdot r_{n,m}^0 + \Pr(\mathcal{H}_n^1)\big(1 - Q_n^d(\tau,\varepsilon_n)\big) \cdot r_{n,m}^1\Big]. \tag{5.7}$$

Suppose the average transmission power of secondary transmitter m is bounded by P_m^{avg}, and the maximal instantaneous transmission power of secondary transmitter m is bounded by P_m^{peak}. Then for secondary transmitter m, we have the following constraints for resource allocation:

$$\sum_{n=1}^{N} p_{n,m} \le P_m^{\text{peak}}, \quad \sum_{n=1}^{N} \Pr(\mathcal{H}_n^0) p_{n,m} \le P_m^{\text{avg}}. \tag{5.8}$$

5.2 Joint Optimization of Sensing Time Setting and Resource Allocation

In this section, a joint sensing duration configuration and resource allocation optimization problem is formulated and addressed, which maximizes the weighted sum of average throughputs of all the secondary users given as $R_w = \sum_{m=1}^{M} \alpha_m \cdot R_m$, where $\alpha_m \ge 0$ is the weight assigned to secondary user m.

To protect the activities of primary users, the detection probability in each channel should be larger than a threshold denoted P_{th} (which should be much larger than 0.5). [2] Together with the constraints discussed in the preceding section, an optimization problem can be formulated:

[2]When additional constraints $P_n^f \le P_{th}^f (P_{th}^f \le 0.5), \forall n \in \mathcal{N}$ are imposed on the proposed optimization problems, the methods presented in this chapter also work.

Problem 5.1.

$$
\begin{aligned}
\underset{\tau,\{\varepsilon_n\},\{p_{n,m}\},\{x_{n,m}\}}{\text{maximize}} \quad & R_w(\tau,\{\varepsilon_n\},\{p_{n,m}\},\{x_{n,m}\}) = \Big(1-\frac{\tau}{T}\Big)\cdot\sum_{n=1}^{N}\Big[\Pr(\mathcal{H}_n^0) \\
& \cdot\Big(1-Q_n^f(\tau,\varepsilon_n)\Big)\cdot\sum_{m=1}^{M}\alpha_m r_{n,m}^0 \qquad (5.9a)\\
& +\Pr(\mathcal{H}_n^1)\cdot\Big(1-Q_n^d(\tau,\varepsilon_n)\Big)\cdot\sum_{m=1}^{M}\alpha_m r_{n,m}^1\Big] \\
\text{s.t.} \quad & 0\le\tau\le T \qquad (5.9b)\\
& Q_n^d(\tau,\varepsilon_n)\ge P_{th},\ \ n\in\mathcal{N} \qquad (5.9c)\\
& Q^{-1}\big(P_n^d(\tau,\varepsilon_n)\big)\sigma\sqrt{2||\mathbf{\Sigma}_n||_1 N} \\
& \quad +||\boldsymbol{H}_n||_1\sqrt{\mu\tau}\ge 0,\ \ n\in\mathcal{N} \qquad (5.9d)\\
& \sum_{m=1}^{M}x_{n,m}=1,\ \ n\in\mathcal{N} \qquad (5.9e)\\
& \sum_{n=1}^{N}p_{n,m}\le P_m^{\text{peak}},\ \ m\in\mathcal{M} \qquad (5.9f)\\
& \sum_{n=1}^{N}\Pr(\mathcal{H}_n^0)p_{n,m}\le P_m^{\text{avg}},\ \ m\in\mathcal{M} \qquad (5.9g)\\
& 0\le x_{n,m}\le 1,\ \ n\in\mathcal{N},m\in\mathcal{M} \qquad (5.9h)\\
& 0\le p_{n,m}\le P_m^{\text{peak}},\ \ n\in\mathcal{N},m\in\mathcal{M} \qquad (5.9i)
\end{aligned}
$$

where $\mathcal{N}=\{1,2,...,N\}$ and $\mathcal{M}=\{1,2,...,M\}$.

Problem 5.1 is nonconvex. And it can be proved that problem 5.1 achieves the optimal solution only when $Q_n^d(\tau,\varepsilon_n)=P_{th}$. The proof is omitted, as it is similar to that in Lemma 3.2.

Then, in problem 5.1, the term $Q_n^d(\tau,\varepsilon_n)$ can be replaced with P_{th}, and $Q_n^f(\tau,\varepsilon_n)$ can be expressed in terms of τ and P_{th} according to equation (5.5). Note that by replacing $Q_n^d(\tau,\varepsilon_n)$ with P_{th}, constraint (5.9d) becomes

$$Q^{-1}(P_{th})\sigma\sqrt{2||\mathbf{\Sigma}_n||_1 N}+||\boldsymbol{H}_n||_1\sqrt{\mu\tau}\ge 0$$

which is equivalent to

$$\tau \geq \tau_{\min}^{n} \triangleq \frac{2\left(Q^{-1}(P_{th})\right)^{2}\sigma^{2}\|\mathbf{\Sigma}_{n}||_{1}N}{\mu\|\boldsymbol{H}_{n}||_{1}^{2}}.$$

Denote $\tau_{\min} = \max(\tau_{\min}^{1}, \tau_{\min}^{2}, ..., \tau_{\min}^{N})$. Then problem 5.1 is equivalent to the following problem:

<u>Problem</u> 5.2.

$$\begin{aligned}
\underset{\tau,\{p_{n,m}\},\{x_{n,m}\}}{\text{maximize}} \quad & R_{w}(\tau,\{p_{n,m}\},\{x_{n,m}\}) = \left(1-\frac{\tau}{T}\right)\cdot\sum_{n=1}^{N}\Big(\Pr(\mathcal{H}_{n}^{0}) \\
& \cdot\Big(1-Q_{n}^{f}(\tau,P_{th})\Big)\cdot\sum_{m=1}^{M}\alpha_{m}r_{n,m}^{0}+\Pr(\mathcal{H}_{n}^{1})\cdot(1-P_{th})\cdot\sum_{m=1}^{M}\alpha_{m}r_{n,m}^{1}\Big) \\
\text{s.t.} \quad & \tau_{\min} \leq \tau \leq T \qquad (5.10a) \\
& \text{Constraints (5.9e) – (5.9i).} \qquad (5.10b)
\end{aligned}$$

Problem 5.2 is still a nonconvex problem. To solve problem 5.2, we use bilevel optimization in which the lower level problem is to optimize $\{p_{n,m}\}$ and $\{x_{n,m}\}$ with τ fixed, while the upper level problem is to optimize τ. Specifically, the lower level problem is

<u>Problem</u> 5.3.

$$\begin{aligned}
U(\tau) = \quad & \max_{\{p_{n,m}\},\{x_{n,m}\}}\sum_{n=1}^{N}\Big[\Pr(\mathcal{H}_{n}^{0})\cdot\Big(1-Q_{n}^{f}(\tau,P_{th})\Big) \\
& \cdot\sum_{m=1}^{M}\alpha_{m}r_{n,m}^{0}+\Pr(\mathcal{H}_{n}^{1})\cdot(1-P_{th})\cdot\sum_{m=1}^{M}\alpha_{m}r_{n,m}^{1}\Big] \\
\text{s.t.} \quad & \text{Constraints (5.9e) – (5.9i).}
\end{aligned}$$

Denote $V(\tau) = (\tau/T)U(\tau)$. Then, the upper level problem is

<u>Problem</u> 5.4.

$$\begin{aligned}
\underset{\tau}{\text{maximize}} \quad & U(\tau) - V(\tau) \\
\text{s.t.} \quad & \tau_{\min} \leq \tau \leq T.
\end{aligned}$$

The lower level problem, problem 5.3, is convex (the proof method can be found in Appendix C, which is originally for the convexity proof of problem 6.1). Thus the optimal solution of $\{p_{n,m}\}$ and $\{x_{n,m}\}$ for a given τ can be obtained.

For the upper level problem, problem 5.4, we have the following lemma.

Lemma 5.1. Functions $U(\tau)$ and $V(\tau)$ are monotonically increasing functions with respect to τ within interval $[\tau_{\min}, T]$.

Proof. We first prove $U(\tau)$ is a monotonically increasing function with respect to τ.

Suppose $\tau^\dagger < \tau^\ddagger$, and the optimal solutions of problem 5.3 with $\tau = \tau^\dagger$ and $\tau = \tau^\ddagger$ are $\{p^\dagger_{n,m}\}$, $\{x^\dagger_{n,m}\}$ and $\{p^\ddagger_{n,m}\}$, $\{x^\ddagger_{n,m}\}$, respectively. Since $\Big(1 - P^f_n(\tau, P_{th})\Big)$ is a monotonically increasing function with respect to τ within the interval $[\tau_{\min}, T]$ according to equation (5.5), we have

$$\begin{aligned}
U(\tau^\dagger) \;&=\; \textstyle\sum_{n=1}^N \Big[\Pr(\mathcal{H}^0_n)\cdot\Big(1 - Q^f_n(\tau^\dagger, P_{th})\Big)\cdot \sum\limits_{m=1}^M \alpha_m x^\dagger_{n,m}\log(1+\frac{|g^{ss}_{n,m}|^2 p^\dagger_{n,m}}{W x^\dagger_{n,m}\sigma^2}) \\
&\qquad + \Pr(\mathcal{H}^1_n)\cdot(1-P_{th})\cdot\sum\limits_{m=1}^M \alpha_m x^\dagger_{n,m}\log\big(1+\frac{|g^{ss}_{n,m}|^2 p^\dagger_{n,m}}{W x^\dagger_{n,m}(|g^{ps}_{n,m}|^2+\sigma^2)}\big)\Big] \\
&<\; \textstyle\sum_{n=1}^N \Big[\Pr(\mathcal{H}^0_n)\cdot\Big(1 - Q^f_n(\tau^\ddagger, P_{th})\Big)\cdot \sum\limits_{m=1}^M \alpha_m x^\dagger_{n,m}\log(1+\frac{|g^{ss}_{n,m}|^2 p^\dagger_{n,m}}{W x^\dagger_{n,m}\sigma^2}) \\
&\qquad + \Pr(\mathcal{H}^1_n)\cdot(1-P_{th})\cdot\sum\limits_{m=1}^M \alpha_m x^\dagger_{n,m}\log\big(1+\frac{|g^{ss}_{n,m}|^2 p^\dagger_{n,m}}{W x^\dagger_{n,m}(|g^{ps}_{n,m}|^2+\sigma^2)}\big)\Big] \\
&\leq\; \textstyle\sum_{n=1}^N \Big[\Pr(\mathcal{H}^0_n)\cdot\Big(1 - Q^f_n(\tau^\ddagger, P_{th})\Big)\cdot \sum\limits_{m=1}^M \alpha_m x^\ddagger_{n,m}\log(1+\frac{|g^{ss}_{n,m}|^2 p^\ddagger_{n,m}}{W x^\ddagger_{n,m}\sigma^2}) \\
&\qquad + \Pr(\mathcal{H}^1_n)\cdot(1-P_{th})\cdot\sum\limits_{m=1}^M \alpha_m x^\ddagger_{n,m}\log\big(1+\frac{|g^{ss}_{n,m}|^2 p^\ddagger_{n,m}}{W x^\ddagger_{n,m}(|g^{ps}_{n,m}|^2+\sigma^2)}\big)\Big] \\
&=\; U(\tau^\ddagger)
\end{aligned}$$

where the last inequality comes from the fact that the optimal solution of problem 5.3 when $\tau = \tau^\ddagger$ is $\{p^\ddagger_{n,m}\}$ and $\{x^\ddagger_{n,m}\}$. Therefore, $U(\tau)$ is monotonically increasing. And $V(\tau)$ is also monotonically increasing, since it is the product of two positive monotonically increasing functions, (τ/T) and $U(\tau)$. $\square$

Denote $\tau' = \tau - \tau_{\min}$. By introducing a new variable ω, problem 5.4 is equivalent to

Problem 5.5.

$$\begin{aligned}
\underset{\tau',\omega}{\text{maximize}} \quad & U(\tau' + \tau_{\min}) + \omega \\
\text{s.t.} \quad & \omega + V(\tau' + \tau_{\min}) \leq V(T) \\
& 0 \leq \tau' \leq T - \tau_{\min} \\
& \omega \geq 0.
\end{aligned}$$

The reason for the equivalence of problems 5.4 and 5.5 is as follows. In problem 5.5, the objective function $\big(U(\tau' + \tau_{\min}) + \omega\big)$ is a monotonically increasing function. And the function $\big(\omega + V(\tau' + \tau_{\min})\big)$ is also monotonically increasing. Therefore, the maximal objective function happens only when $\omega + V(\tau' + \tau_{\min}) = V(T)$. This means problem 5.5 is to maximize

$$\begin{aligned} U(\tau' + \tau_{\min}) + \omega \quad &= U(\tau' + \tau_{\min}) - V(\tau' + \tau_{\min}) + V(T) \\ &= U(\tau) - V(\tau) + V(T) \end{aligned}$$

which is equivalent to maximizing $\big(U(\tau) - V(\tau)\big)$ since $V(T)$ is a constant. So problems 5.4 and 5.5 are equivalent.

Problem 5.5 is in the form of *monotonic optimization problem* [78], because it has the following features: i) the object function, $\big(U(\tau'+\tau_{\min})+\omega\big)$, is a monotonically increasing function; ii) all constraint functions are monotonic functions; and iii) both τ' and ω lie in the region of $[0, +\infty)$. A monotonic optimization problem can be solved by a monotonic programming method, named *polyblock algorithm* [78] with parameter ϵ, which is ϵ-optimal (i.e., the difference of achieved utility in the polyblock algorithm from the global optimal utility is bounded by ϵ). In specific, for problem 5.5, a polyblock algorithm is given as follows. Interested readers may refer to reference [78] for a detailed discussion of the algorithm.

Algorithm 5.1 Polyblock algorithm.

Define a point set $\mathcal{S} = \{s_1, s_2, ..., s_{|\mathcal{S}|}\}$, in which each point has two elements. Initialize set $\mathcal{S}$ by one point $s_1 = (T - \tau_{\min}, V(T) - V(\tau_{\min}))$. In other words, the two elements in point s_1 are $s_1(1) = T - \tau_{\min}$ and $s_1(2) = V(T) - V(\tau_{\min})$.
while Set $\mathcal{S} \neq \emptyset$ **do**
  for $i = 1, 2, ..., |\mathcal{S}|$ **do**
    Calculate λ_i that satisfies $\lambda_i s_i(2) + V(\lambda_i s_i(1) + \tau_{\min}) = V(T)$ by a bisection search, and set $\pi_i = \lambda_i s_i$.
  Find $i^* = \underset{1 \le i \le |\mathcal{S}|}{\arg\max} \, [U(\pi_i(1) + \tau_{\min}) + \pi_i(2)]$
  For any point (say point s) in set $\mathcal{S}$, if $[U(s(1) + \tau_{\min}) + s(2)] \le [U(\pi_{i^*}(1) + \tau_{\min}) + \pi_{i^*}(2)] + \epsilon$, then remove the point from $\mathcal{S}$.
  if Set $\mathcal{S} \neq \emptyset$ **then**
    Find $j^* = \underset{1 \le j \le |\mathcal{S}|}{\arg\max} \, [U(s_j(1) + \tau_{\min}) + s_j(2)]$
    Calculate λ_{j^*} that satisfies $\lambda_{j^*} s_{j^*}(2) + V(\lambda_{j^*} s_{j^*}(1) + \tau_{\min}) = V(T)$ by a bisection search, and set $\pi_{j^*} = \lambda_{j^*} s_{j^*}$.
    Generate two points $s^\dagger = s_{j^*} + (\pi_{j^*} - s_{j^*}) \circ (1, 0)$ and $s^\ddagger = s_{j^*} + (\pi_{j^*} - s_{j^*}) \circ (0, 1)$, where $\circ$ means Hadamard product.
    Add $s^\dagger$ and $s^\ddagger$ into $\mathcal{S}$, delete s_{j^*} from $\mathcal{S}$.
Output the last π_{i^*} before $\mathcal{S}$ becomes an empty set.

By employing algorithm 5.1, the first and second elements of π_{i^*} are the optimal τ' and ω, respectively, for problem 5.5. Then, the optimal τ in problem 5.4 is obtained by $\tau = \pi_{i^*}(1) + \tau_{\min}$.

5.3 Fairness Consideration

In the preceding section, the weighted throughput maximization problem is solved with the aid of bilevel optimization and monotonic programming. In this section, we focus on the cases when fairness among users is taken into account. In specific, proportional fairness and max-min fairness are discussed in the following two subsections, respectively.

5.3.1 Proportional Fairness Optimization Problem

In this case, the utility function of secondary user m can be written as $\log R_m$. Therefore, the optimization problem is as follows.

<u>Problem</u> 5.6.

$$\begin{aligned}
\underset{\tau,\{\varepsilon_n\},\{p_{n,m}\},\{x_{n,m}\}}{\text{maximize}} \quad & R_{pf}(\tau,\{\varepsilon_n\},\{p_{n,m}\},\{x_{n,m}\}) = \sum_{m=1}^{M} \log\left(\left(1-\frac{\tau}{T}\right)\sum_{n=1}^{N}\Big[\Pr(\mathcal{H}_n^0)\right.\\
& \left.\cdot\big(1-Q_n^f(\tau,\varepsilon_n)\big)\cdot r_{n,m}^0 + \Pr(\mathcal{H}_n^1)\big(1-Q_n^d(\tau,\varepsilon_n)\big)\cdot r_{n,m}^1\Big]\right)\\
\text{s.t.} \quad & \text{Constraints (5.9b) – (5.9i).}
\end{aligned}$$

Similar to the preceding section, problem 5.6 is nonconvex, and it achieves the optimal solution only when $Q_n^d(\tau,\varepsilon_n) = P_{th}$. After substituting $Q_n^d(\tau,\varepsilon_n)$ with P_{th}, problem 5.6 is equivalent to a bilevel problem. The lower level problem is to optimize $\{p_{n,m}\}$ and $\{x_{n,m}\}$ with a fixed τ, as

<u>Problem</u> 5.7.

$$\begin{aligned}
Z(\tau) = \quad & \max_{\{p_{n,m}\},\{x_{n,m}\}} \sum_{m=1}^{M} \log\left(\sum_{n=1}^{N}\Big[\Pr(\mathcal{H}_n^0)\cdot\Big(1-Q_n^f(\tau,P_{th})\Big)\cdot r_{n,m}^0\right.\\
& \left. + \Pr(\mathcal{H}_n^1)\cdot(1-P_{th})\cdot r_{n,m}^1\Big]\right)\\
\text{s.t.} \quad & \text{Constraints (5.9e) – (5.9i).}
\end{aligned}$$

According to [80, Page 84] , the objective function of problem 5.7 can be shown to be a concave function. Therefore, problem 5.7 is convex. In addition, the maximal objective function in problem 5.7, $Z(\tau)$, is a monotonically increasing function with respect to τ. The proof is similar to that in Lemma 5.1, and is omitted here.

The upper level problem is

<u>Problem</u> 5.8.

$$\begin{aligned}
\underset{\tau}{\text{maximize}} \quad & R_{pf}(\tau) = \log(1-\frac{\tau}{T})^M + Z(\tau)\\
\text{s.t.} \quad & \tau_{\min} \leq \tau \leq T.
\end{aligned}$$

For expression

$$\left(1-\frac{\tau}{T}\right)^M=\sum_{l=0}^{M}\binom{M}{l}\left(-\frac{\tau}{T}\right)^l \tag{5.11}$$

denote $Y_1(\tau)$ as the summation of all the positive items on the right-hand side, and $Y_2(\tau)$ as the summation of all the absolution values of negative items on the right-hand side. So we have

$$\left(1-\frac{\tau}{T}\right)^M=Y_1(\tau)-Y_2(\tau). \tag{5.12}$$

It is clear that both $Y_1(\tau)$ and $Y_2(\tau)$ are positive and monotonically increasing functions with respect to τ. Then the object function in problem 5.8 can be changed to

$$e^{R_{pf}(\tau)}=Y_1(\tau)\cdot e^{Z(\tau)}-Y_2(\tau)\cdot e^{Z(\tau)}. \tag{5.13}$$

In the new objective function, both $Y_1(\tau)e^{Z(\tau)}$ and $Y_2(\tau)e^{Z(\tau)}$ are monotonically increasing functions with respect to τ. Therefore, problem 5.8 has a form similar to problem 5.4, and can be solved by a polyblock algorithm.

5.3.2 Max-min Fairness Optimization Problem

With max-min fairness, the optimization problem can be formulated as

Problem 5.9.

$$\begin{aligned}
\underset{\tau,\{\varepsilon_n\},\{p_{n,m}\},\{x_{n,m}\}}{\text{maximize}} \quad & \min_m \Bigg((1-\frac{\tau}{T})\cdot\sum_{n=1}^{N}\Big[\Pr(\mathcal{H}_n^0)\big(1-Q_n^f(\tau,\varepsilon_n)\big)\cdot r_{n,m}^0 \\
& +\Pr(\mathcal{H}_n^1)\big(1-Q_n^d(\tau,\varepsilon_n)\big)\cdot r_{n,m}^1\Big]\Bigg) \\
\text{s.t.} \quad & \text{Constraints (5.9b) – (5.9i).}
\end{aligned}$$

For problem 5.9, it can also be concluded that the optimal utility is achieved only when $Q_n^d(\tau,\varepsilon_n)=P_{th}$.

For the objective function in problem 5.9, we have

$$\begin{aligned}
&\max_{\tau,\{\varepsilon_n\},\{p_{n,m}\},\{x_{n,m}\}} \min_m \Bigg((1-\tfrac{\tau}{T}) \cdot \sum_{n=1}^{N} \Big[\Pr(\mathcal{H}_n^0) \\
&\cdot \big(1 - Q_n^f(\tau,\varepsilon_n)\big) \cdot r_{n,m}^0 + \Pr(\mathcal{H}_n^1)\big(1 - Q_n^d(\tau,\varepsilon_n)\big) \cdot r_{n,m}^1 \Big] \Bigg) \\
&= \max_{\tau} \max_{\{\varepsilon_n\},\{p_{n,m}\},\{x_{n,m}\}} \min_m \Bigg((1-\tfrac{\tau}{T}) \cdot \sum_{n=1}^{N} \Big[\Pr(\mathcal{H}_n^0) \\
&\cdot \big(1 - Q_n^f(\tau,\varepsilon_n)\big) \cdot r_{n,m}^0 + \Pr(\mathcal{H}_n^1)\big(1 - Q_n^d(\tau,\varepsilon_n)\big) \cdot r_{n,m}^1 \Big] \Bigg) \\
&= \max_{\tau} \; (1-\tfrac{\tau}{T}) \cdot \max_{\{\varepsilon_n\},\{p_{n,m}\},\{x_{n,m}\}} \min_m \Bigg(\sum_{n=1}^{N} \Big[\Pr(\mathcal{H}_n^0) \\
&\cdot \big(1 - Q_n^f(\tau,\varepsilon_n)\big) \cdot r_{n,m}^0 + \Pr(\mathcal{H}_n^1)\big(1 - Q_n^d(\tau,\varepsilon_n)\big) \cdot r_{n,m}^1 \Big] \Bigg).
\end{aligned} \tag{5.14}$$

Setting $Q_n^d(\tau,\varepsilon_n) = P_{th}$ and applying the transformation shown in (5.14), problem 5.9 is equivalent to a bilevel optimization problem. The lower level problem is

Problem 5.10.

$$\begin{aligned}
&F(\tau) = \max_{\{p_{n,m}\},\{x_{n,m}\},\Omega} \quad \Omega \\
\text{s.t.} \quad &\sum_{n=1}^{N} \Big(\Pr(\mathcal{H}_n^0)\big(1 - Q_n^f(\tau,P_{th})\big) \cdot r_{n,m}^0 + \Pr(\mathcal{H}_n^1)(1 - P_{th}) \cdot r_{n,m}^1 \Big) \geq \Omega, \; m \in \mathcal{M} \\
&\text{Constraints (5.9e) –(5.9i).}
\end{aligned}$$

Problem 5.10 is convex, and its maximal objective function, $F(\tau)$, is a monotonically increasing function with respect to τ.

The upper level problem is to maximize $(1-\tau/T)\cdot F(\tau)$ subject to τ within the region defined by constraint $\tau_{\min} \leq \tau \leq T$. Note that the objective function is a difference between two positive monotonically increasing functions, $F(\tau)$ and $(\tau/T)F(\tau)$, which is similar to problem 5.4. Therefore, the upper level problem can be solved by a polyblock algorithm.

5.4 Numerical Results

In this section, numerical results are demonstrated to verify our proposed algorithms. The system is set up as follows. There are four primary channels and four secondary users in a cognitive radio network. Each of the channels spans a bandwidth of 1 MHz. The

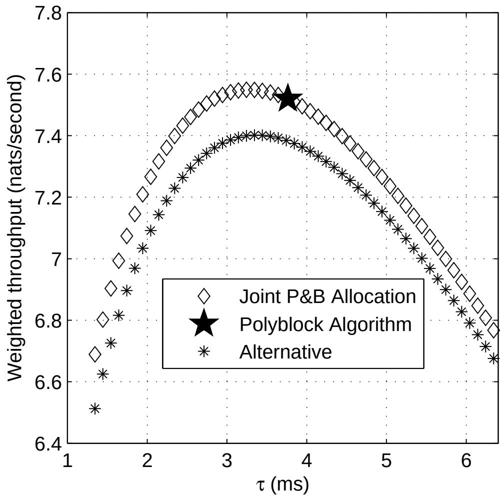

Fig. 5.1. Weighted sum of throughputs versus τ.

free probabilities of the four channels are 0.9, 0.8, 0.7 and 0.6, respectively. The weights assigned to secondary users, α_m's, are all set as 1. The sampling rate is $\mu = 8$ MHz; the slot duration is $T = 20$ ms; the threshold of detection probability is $P_{th} = 0.9$. Of the four secondary users, the average transmission power values P_m^{avg}'s are set as 0.5, 0.45, 0.4, and 0.35, respectively, and the peak transmission power values P_m^{peak}'s are set as 0.8, 0.9, 1, and 1.2, respectively. All links (from primary users to secondary users and between secondary users) experience Rayleigh fading. And in each channel, the channel SNR value between each secondary transceiver pair is with mean 15 dB, and the channel SNR value from the primary user to a secondary user (either the transmitter or the receiver) is with mean -15 dB. The value of ϵ is set as 0.05.

We first demonstrate the effect of sensing time duration configuration in our research. For different particular values of τ in each of the three optimization problems in Chapters 5.2 and 5.3, we obtain the optimal utilities under the method of joint power and bandwidth allocation (which is a convex optimization problem). The corresponding curves are marked as "Joint P&B Allocation" in Figs. 5.1-5.3. It can be seen that as τ grows, the optimal utilities first increase and then decrease. The first increase of the optimal utilities is due to the improvement in false alarm probability. When the sensing time duration reaches a certain value, increase of sensing time duration does not change much the false alarm probability. So the optimal utilities decrease, since less time is used in data transmission. From the

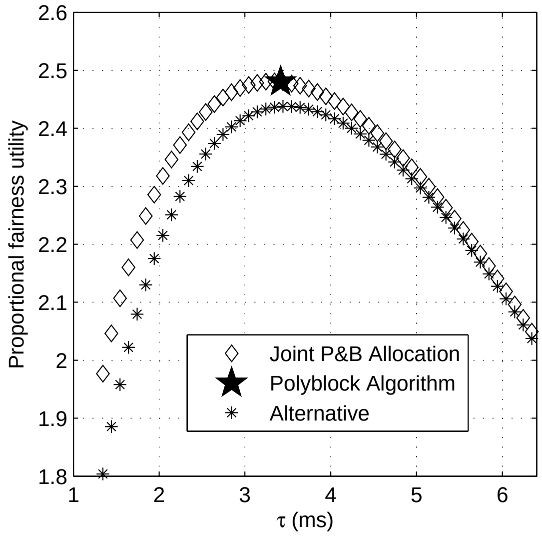

Fig. 5.2. Proportional fairness utility versus τ.

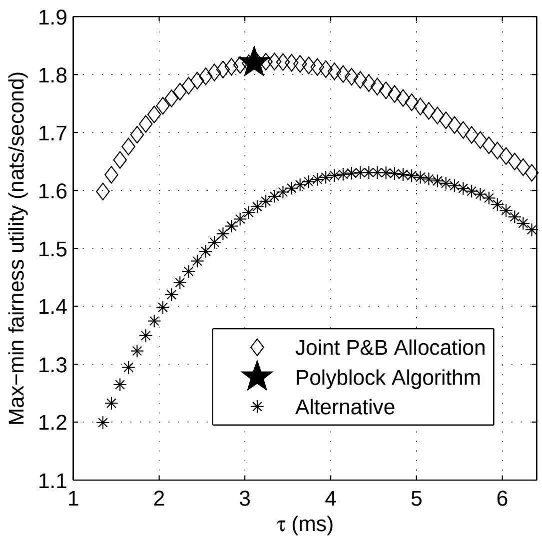

Fig. 5.3. Max-min fairness utility versus τ.

curves it can be seen that the sensing time duration affects largely the system performance, and thus, it is important to find the optimal sensing time setting. In Figs. 5.1-5.3, we also show the resulted sensing time duration in our proposed polyblock algorithms and the corresponding utilities, marked as "Polyblock Algorithm" in the figures. It is clear that utilities obtained in our polyblock algorithms are very close to the global optimal utilities in Figs. 5.1-5.3.

Next we demonstrate the effect of resource allocation in our research, in which each channel can be shared by multiple users. As a comparison, we consider an alternative case when each channel can be used by one secondary user only. The alternative case is equivalent to solving problems 5.1, 5.6 and 5.9 with additional constraints $x_{n,m} \in \{0,1\}, m \in \mathcal{M}, n \in \mathcal{N}$. In the newly formulated problems, it is possible that a secondary user is assigned more than one channel, and it is also possible that a secondary user is assigned no channel. It can be proved that the optimal utilities are achieved only when $P_n^d(\tau, \varepsilon_n) = P_{th}$. So for a given τ, the optimal ε_n can be obtained based on $P_n^d(\tau, \varepsilon_n) = P_{th}$. Therefore, to maximize the utility for a given τ, we need to determine optimal power allocation ($p_{n,m}$) and channel allocation ($x_{n,m}$) strategies. This is a mixed-integer problem which is usually NP-hard. To solve the problem, we exhaustively search all the combinations of assigning N channels to M secondary users. In each combination, the power allocation (over probably multiple channels) at each secondary user is determined by a convex optimization problem. Then for each τ, the optimal power and channel allocation strategy is the strategy associated with the combination whose optimal utility has the largest value. For the alternative case, the optimal utilities for different τ are also plotted in Figs. 5.1-5.3, marked as "Alternative". It can be seen that the alternative case has worse performance than our polyblock algorithms. This is because our algorithms take advantage of dynamic sharing of each channel. Another advantage of our algorithms over the algorithms in the alternative case is the computational complexity. Apparently our algorithms are with significantly less complexity, since our algorithms are for problems with continuously-valued variables, while the alternative case is associated with mixed-integer problems.

Last we compare the fairness among the average throughputs of secondary users achieved by our proposed polyblock algorithm, in the three optimization problems in Chapters 5.2 and 5.3. Table 5.1 lists the achieved average throughputs of the four secondary users in the three optimization problems. It can be seen that, the throughput summation maximization

TABLE 5.1
AVERAGE THROUGHPUTS OF THE FOUR SECONDARY USERS (SUS)

Objective	SU1	SU2	SU3	SU4	Throughput summation
Weighted throughput maximization	1.7722	2.4634	2.1211	1.163	7.5197
Proportional fairness	1.7455	1.8918	2.1667	1.6529	7.4569
Max-min fairness	1.8206	1.8206	1.8206	1.8206	7.2824

problem cares only throughput, and thus, has the largest throughput summation, but has the lowest fairness level. The max-min fairness problem achieves the best fairness among all the users, i.e., each user has the same throughput. But the throughput summation is the smallest. The proportional fairness problem achieves better fairness but smaller throughput summation than throughput summation maximization problem, and achieves worse fairness but more throughput summation than the max-min fairness problem. All these mean that there is a tradeoff between throughput summation and fairness among secondary users.

5.5 Conclusion

In this chapter, the problem of optimal multichannel cooperative sensing and resource allocation (of spectrum and power) in cognitive radio networks is explored. The issues of how to set spectrum sensing time, how to determine the spectrum sensing threshold, how to allocate spectrum resources, and how to set transmission power are investigated. A weighted throughput maximization problem is formulated. Although nonconvex, the problem is solved with the aid of bilevel optimization and monotonic programming methods. We also show that the bilevel optimization and monotonic programming methods are applicable when proportional fairness or max-min fairness is considered.

Chapter 6

Resource Allocation in a Cognitive Radio Network in Underlay Mode

In this chapter, the average rate maximization problem given constraints on transmission power and interference power in a cognitive radio network over slow fading channels is studied. For the problem, either long-term or short-term constraint is imposed on the transmission power and interference power. In each combination of the constraints on transmission power and interference power, the original optimization problem is decomposed into sub-problems, each corresponding to a specific channel gain realization. For each sub-problem, a fast algorithm with closed-form solutions is provided. The case with additional peak power spectrum density constraints is also investigated. [1]

6.1 Introduction

In a wireless communication system, the channel between a transceiver pair experiences fading. Therefore, *ergodic capacity*, the expectation of channel capacity over all the channel fading states, has attracted research attention recently in resource allocation problems. The ergodic capacity region together with resource allocation problems in multiple-access channels is studied in [91]. Ergodic capacity and resource allocation strategies in broadcast channels are given in [92]. The ergodic capacity and power allocation problems are studied in [67] with both multiple-access channel and broadcast channel in cognitive radio

[1] *The research method and proposed mathematical problem in this chapter are also applicable in the case of relay networks. The research results for relay networks were published in IEEE Trans.Veh. Technol., 60: 3865-3881 (2011).*

networks.

Note that ergodic capacity is defined for fast fading channels only. When slow fading channel is considered, outage capacity is a commonly used performance metric [93]. However, in this chapter, we are interested in a setup that is different from the traditional slow fading model. In our model, in addition to the slow fading, we assume that the transmitter has the channel state information (CSI) (the fact that the channel is in slow fading makes this assumption reasonable). As pointed out in many recent works (e.g. [94]), the presence of CSI at transmitter makes the average rate a reasonable alternative performance metric for the slow fading channel. The transmit rate can be matched to the fading (e.g., by adaptive modulation) when timely CSI is available at the transmitter in slow fading. In this case the fading-induced outages can be essentially eliminated and an average rate can be achieved, which is defined as the expectation of rate over a sufficiently long interval and coincides with the ergodic capacity (i.e., expectation over fading distribution) when the fading process is ergodic. In this chapter, we target at maximizing the average rate of secondary users in a cognitive radio network in underlay mode over slow fading channels.

6.2 System Model and Problem Formulation

Consider a cognitive radio network with N frequency bands (called channels) with equal bandwidth. In each channel, a primary user exists. The secondary base station (BS) needs to transmit to N secondary users, each over one of the N channels, in underlay mode. For simplicity of presentation, the secondary BS transmits to secondary user S_n over channel n, and the bandwidth of each channel is normalized to 1. Since there might be multiple secondary BSs in the area, the target secondary BS only occupies a portion of the bandwidth in each channel, and the total occupied bandwidth by the target secondary BS in all the N channels should be no more than B. For the secondary BS, denote the transmit power and bandwidth in channel n as p_n and x_n, respectively, which are to be optimized.

Denote the channel gain between the secondary BS and S_n is g_n and the channel gain between the secondary BS and the primary user (primary receiver) on the n^{th} channel as h_n. Assume that g_n's and h_n's ($n \in \mathcal{N}$, where $\mathcal{N} = \{1, 2, ..., N\}$) are flat fading. All the channel gains are supposed to be block-fading, which means the channel gains keep stable within each transmission block. At the beginning of each fading block, the secondary BS can measure channel gains g_n's and h_n's ($n \in \mathcal{N}$).

By normalizing the power spectrum density of background noise as 1, the channel capacity from the secondary BS to the secondary user S_n can be expressed as

$$C_n = x_n \ln\left(1 + \frac{p_n g_n}{x_n}\right). \tag{6.1}$$

Then the total downlink channel capacity from the secondary BS to all the secondary users is

$$C = \sum_{n=1}^{N} C_n. \tag{6.2}$$

Channel gains g_n and h_n are assumed to be ergodic processes over fading blocks. Denote the joint PDF of the channel gains as $f(\boldsymbol{g}, \boldsymbol{h})$, where $\boldsymbol{g} = (g_1, ..., g_N)$ and $\boldsymbol{h} = (h_1, ..., h_N)$. It is reasonable to further assume that g_n's and h_n's are independent from each other. In this case, power and bandwidth allocation should be based on the instantaneous channel states. In other words, the power allocation strategy p_n's and bandwidth allocation strategy x_n's are all dependent on the fading state $\boldsymbol{g}$ and $\boldsymbol{h}$. For the sake of simplicity, the fading state $\boldsymbol{g}$ and $\boldsymbol{h}$ are omitted in the notations of p_n's and x_n's. Then the total expected rate from the secondary BS to all the secondary users, referred to as the *average rate*, is

$$C_e = \mathbb{E}\left[\sum_{n=1}^{N} x_n \ln\left(1 + \frac{p_n g_n}{x_n}\right)\right] \tag{6.3}$$

where $\mathbb{E}$ means expectation for $\boldsymbol{g}$ and $\boldsymbol{h}$. And we have

$$p_n \geq 0, \ \forall n \in \mathcal{N} \tag{6.4}$$

$$0 \leq x_n \leq 1, \ \forall n \in \mathcal{N}; \ \ \sum_{n=1}^{N} x_n \leq B. \tag{6.5}$$

Throughout this chapter, B is assumed to be no larger than 1. The research result for the case when $B > 1$ will be studied in the future. When $B \leq 1$, the constraint in (6.5) can be simplified as

$$x_n \geq 0, \ \forall n \in \mathcal{N}; \ \ \sum_{n=1}^{N} x_n \leq B. \tag{6.6}$$

In addition, the transmission power at the secondary BS is limited. The instantaneous transmission power may be bounded by a pre-specified value P^{ST}, referred to as *short-term* *transmission-power constraint*, given as

$$\sum_{n=1}^{N} p_n \leq P^{ST}, \tag{6.7}$$

or the average transmission power may be bounded by a pre-specified value P^{LT}, referred to as *long-term transmission-power constraint*, given as

$$\mathbb{E}\left[\sum_{n=1}^{N} p_n\right] \leq P^{LT}. \tag{6.8}$$

The interference to the primary user on channel n should also be limited. The short-term interference-power constraint bounds the instantaneous interference power $p_n h_n$ by I_n^{ST}, given as

$$p_n h_n \leq I_n^{ST}, \ \forall n \in \mathcal{N} \tag{6.9}$$

while the long-term interference-power constraint bounds the average interference power $\mathbb{E}[p_n h_n]$ by I_n^{LT}, given as

$$\mathbb{E}[p_n h_n] \leq I_n^{LT}, \ \forall n \in \mathcal{N}. \tag{6.10}$$

A group of optimization problems can be formulated in the following form.

Problem 6.1.

$$\begin{array}{ll}
\max\limits_{\{p_n\},\{x_n\}} & \mathbb{E}\left[\sum\limits_{n=1}^{N} x_n \ln\left(1+\frac{p_n g_n}{x_n}\right)\right] \\
\text{s.t.} & p_n \geq 0, x_n \geq 0, \forall n \in \mathcal{N}; \quad \sum\limits_{n=1}^{N} x_n \leq B \\
& \sum\limits_{n=1}^{N} p_n \leq P^{ST} \text{ (short-term)} \quad \text{or} \quad \mathbb{E}\left[\sum\limits_{n=1}^{N} p_n\right] \leq P^{LT} \text{ (long-term)} \\
& p_n h_n \leq I_n^{ST} \text{ (short-term)} \quad \text{or} \quad \mathbb{E}[p_n h_n] \leq I_n^{LT} \text{ (long-term)}, \ \forall n \in \mathcal{N}.
\end{array} \tag{6.11}$$

In problem 6.1, we need to derive the power and bandwidth allocation p_n's and x_n's with respect to every specific channel realization $\boldsymbol{g}$ and $\boldsymbol{h}$, which is proved to be a convex problem in Appendix C. The resource allocation is conducted at the secondary BS. In the following sections, we will focus on how to solve the resource allocation problem and how to implement the resource allocation, respectively.

6.3 Average Rate Maximization

In this section, problem 6.1 with four scenarios of power constraints is studied sequentially. In each scenario, optimal bandwidth and power allocation strategy are derived.

6.3.1 Scenario with Long-term Transmission-power and Long-term Interference-power Constraints

In this scenario, the Lagrangian of problem 6.1 can be written as

$$\begin{aligned} & \mathcal{L}(\{p_n\},\{x_n\},\{\mu_n\},\lambda) \\ = \; & \mathbb{E}\left[\sum_{n=1}^{N} x_n \ln\left(1+\frac{p_n g_n}{x_n}\right)\right] - \lambda\left(\mathbb{E}\left[\sum_{n=1}^{N} p_n\right] - P^{LT}\right) \\ & - \sum_{n=1}^{N} \mu_n \left(\mathbb{E}\left[p_n h_n\right] - I_n^{LT}\right) \end{aligned} \tag{6.12}$$

where μ_n and λ are non-negative Lagrange multipliers corresponding to the constrains $\mathbb{E}\left[p_n h_n\right] \leq I_n^{LT}$ and $\mathbb{E}\left[\sum_{n=1}^{N} p_n\right] \leq P^{LT}$, respectively. Then the Lagrange dual function is given as

$$\begin{aligned} & \mathcal{D}(\lambda,\{\mu_n\}) = \max_{\{p_n\},\,\{x_n\}} \mathcal{L}(\{p_n\},\{x_n\},\{\mu_n\},\lambda) \\ & \text{s.t.} \quad p_n \geq 0, x_n \geq 0,\ \forall n \in \mathcal{N};\ \ \sum_{n=1}^{N} x_n \leq B. \end{aligned} \tag{6.13}$$

Since problem 6.1 in this scenario is convex, and strictly feasible points are available in this problem, which satisfies the Slater's condition [80], the minimum of the duality function

$$\min_{\lambda\geq 0,\ \mu_n \geq 0,\ n\in\mathcal{N}} \mathcal{D}(\lambda,\{\mu_n\}) \tag{6.14}$$

is guaranteed to achieve the optimal utility of problem 6.1. Therefore, optimizing problem 6.1 is equivalent to minimizing the duality function $\mathcal{D}(\lambda,\{\mu_n\})$. To converge to the minimum of $\mathcal{D}(\lambda,\{\mu_n\})$, the Lagrange multipliers λ and μ_n's can be updated by resorting to the sub-gradient method [95] as

$$\lambda(t+1) = \left(\lambda(t) - a(t)\left(P^{LT} - \mathbb{E}\left[\sum_{n=1}^{N} p_n(t)\right]\right)\right)^{+} \tag{6.15}$$

and

$$\mu_n(t+1) = \left(\mu_n(t) - a(t)\left(I_n^{LT} - \mathbb{E}\left[p_n(t)h_n\right]\right)\right)^+, \forall n \in \mathcal{N} \tag{6.16}$$

where $p_n(t)$ is the optimal power allocation solution of the dual function in (6.13) at the t^{th} iteration, $a(t)$ is the positive step size at the t^{th} iteration, and $(x)^+ = \max(x, 0)$.

In other words, problem 6.1 is decomposed into two levels. In the higher level, λ and μ_n's are updated iteratively as in (6.15) and (6.16). In the lower level, the Lagrange dual function in (6.13) is obtained for each iteration.

The update of λ and μ_n's in (6.15) and (6.16) will be discussed in Chapter 6.5. So next we focus on how to obtain the optimal solution of the dual function in (6.13),[2] i.e., how to obtain $\mathcal{D}(\lambda, \{\mu_n\})$ for given λ and μ_n's.

It can be seen that the optimization problem in (6.13) can be solved by investigating the sub-problem as follows for every realization of channel gains $\boldsymbol{g}$ and $\boldsymbol{h}$.

$$\begin{array}{ll} \max\limits_{\{p_n\},\,\{x_n\}} & \sum\limits_{n=1}^{N} x_n \ln\left(1 + \frac{p_n g_n}{x_n}\right) - \lambda \sum\limits_{n=1}^{N} p_n - \sum\limits_{n=1}^{N} \mu_n p_n h_n \\ \text{s.t.} & p_n \geq 0, x_n \geq 0,\ \forall n \in \mathcal{N};\ \ \sum_{n=1}^{N} x_n \leq B. \end{array} \tag{6.17}$$

Note that for the ease of presentation, the constant terms λP^{LT} and $\sum\limits_{n=1}^{N} \mu_n I_n^{LT}$ in the objective function are discarded.

The sub-problem in (6.17) is a convex problem, and traditionally, can be solved by numerical optimization methods, such as sub-gradient method and interior-point algorithm [96]. However, the traditional methods require iterative calculations and can only numerically achieve the optimal solution (i.e., no closed-form solution is achieved). In this paper, we will not use the numerical methods to solve the sub-problem in (6.17). Rather, we will use Karush-Kuhn-Tucker (KKT) conditions to analyze special properties of the convex sub-problem in (6.17). Based on the special properties, we give a fast algorithm with a closed-form optimal solution for the sub-problem, which avoids high computational complexity.

The sub-problem in (6.17) satisfies the Slater's condition. Then, the KKT condition which serves as a sufficient and necessary condition for the optimal solution can be listed

[2]Note that the method in Chapter 6.5 also applies for the scenarios discussed in Chapter 6.3.2, 6.3.3, and 6.3.4. So in those three subsections, we also focus only on the lower-level problems.

as follows [80].

$$\ln\left(1+\frac{p_ng_n}{x_n}\right)-\frac{p_ng_n}{x_n+p_ng_n}-\Gamma^*+\delta_n^*=0,\ \forall n\in\mathcal{N} \tag{6.18a}$$

$$\frac{x_ng_n}{x_n+g_np_n}-\lambda-\mu_nh_n+\Delta_n^*=0,\ \forall n\in\mathcal{N} \tag{6.18b}$$

$$\Delta_n^*p_n=0,\delta_n^*x_n=0,\ \forall n\in\mathcal{N};\ \Gamma^*\left(\sum_{n=1}^{N}x_n-B\right)=0 \tag{6.18c}$$

$$p_n\geq 0,x_n\geq 0,\ \forall n\in\mathcal{N};\ \sum_{n=1}^{N}x_n\leq B. \tag{6.18d}$$

in which Δ_n^*, δ_n^* and Γ^* are non-negative Lagrange multipliers associated with constraints $p_n \geq 0$, $x_n \geq 0$, and $\sum_{n=1}^{N} x_n \leq B$, respectively. In this chapter, when superscript $*$ is used for a Lagrange multiplier, it means the Lagrange multiplier is associated with a sub-problem for a realization of $(\boldsymbol{g},\boldsymbol{h})$ (to distinguish from the Lagrange multipliers associated with the original optimization problem 6.1). Equations (6.18a) and (6.18b) are obtained by setting the derivative of the Lagrangian of the sub-problem in (6.17) with respect to x_n and p_n as 0, respectively. The following lemma is in order for the KKT condition (6.18a)-(6.18d).

Lemma 6.1. If $\mathcal{A}=\{j|p_j>0,x_j>0\}\neq\emptyset$ (null set), then $|\mathcal{A}|\leq 1$.

Proof. Define $\text{SNR}_n=\frac{p_ng_n}{x_n}$. Equations (6.18a) and (6.18b) can be rewritten as

$$\ln(1+\text{SNR}_n)-\frac{\text{SNR}_n}{1+\text{SNR}_n}=\Gamma^*-\delta_n^*,\ \ \forall n\in\mathcal{N} \tag{6.19a}$$

$$\frac{1}{1+\text{SNR}_n}=\frac{\lambda}{g_n}+\frac{\mu_nh_n}{g_n}-\frac{\Delta_n^*}{g_n},\ \ \forall n\in\mathcal{N}. \tag{6.19b}$$

Suppose $n^\dagger\in\mathcal{A}$. From (6.18c) it can be seen that $\delta_{n^\dagger}^*$ and $\Delta_{n^\dagger}^*$ should be zero. Define $S(x)\triangleq\ln(1+x)-\frac{x}{1+x}$, which is a monotonic increasing non-negative function of non-negative x. Since $\text{SNR}_{n^\dagger}=\frac{p_{n^\dagger}g_{n^\dagger}}{x_{n^\dagger}}\in(0,+\infty)$ and $\delta_{n^\dagger}^*=0$, from (6.19a) we have

$$\Gamma^*=S(\text{SNR}_{n^\dagger})+\delta_{n^\dagger}^*=S(\text{SNR}_{n^\dagger})\in(0,+\infty).$$

When λ and $\mu_{n^\dagger}$ are nonzero (this case is general in the updating procedure of λ and $\mu_{n^\dagger}$), after dividing the left- and right-hand sides of (6.19a) by the left- and right-hand sides of

(6.19b), respectively, for $n^\dagger$, we get

$$(1+\text{SNR}_{n^\dagger})\ln(1+\text{SNR}_{n^\dagger}) - \text{SNR}_{n^\dagger} = \frac{\Gamma^*}{\frac{\lambda}{g_{n^\dagger}} + \frac{\mu_{n^\dagger}h_{n^\dagger}}{g_{n^\dagger}}}. \tag{6.20}$$

Define $T(x) \triangleq (1+x)\ln(1+x)-x$, which is a monotonic increasing non-negative function of non-negative x. From (6.19a) and (6.20), we have

$$S\left(T^{-1}\left(\frac{\Gamma^*}{\frac{\lambda}{g_{n^\dagger}} + \frac{\mu_{n^\dagger}h_{n^\dagger}}{g_{n^\dagger}}}\right)\right) = \Gamma^* \tag{6.21}$$

and $S(T^{-1}(\cdot))$ is a monotonic increasing function. So if there exists $n^\ddagger(\neq n^\dagger) \in \mathcal{A}$, it should also satisfy equation (6.21), which leads to

$$\frac{\Gamma^*}{\frac{\lambda}{g_{n^\dagger}} + \frac{\mu_{n^\dagger}h_{n^\dagger}}{g_{n^\dagger}}} = \frac{\Gamma^*}{\frac{\lambda}{g_{n^\ddagger}} + \frac{\mu_{n^\ddagger}h_{n^\ddagger}}{g_{n^\ddagger}}}. \tag{6.22}$$

Note that $g_{n^\dagger}$, $h_{n^\dagger}$, $g_{n^\ddagger}$ and $h_{n^\ddagger}$ are independent channel gains, while λ, $\mu_{n^\dagger}$ and $\mu_{n^\ddagger}$ are fixed for the problem in (6.17). Therefore, the probability that equation (6.22) holds is zero. So there is at most one element in set $\mathcal{A}$ almost surely, i.e., $|\mathcal{A}| \leq 1$. This completes the proof.[3] $\square$

Note that $\mathcal{A}$ is actually the set of selected channels, i.e., with non-zero power and bandwidth assignment. Therefore, Lemma 6.1 indicates that either of the two following cases happens: i) no channel is assigned bandwidth and power (e.g., when the channels are poor, it may be better not to transmit, since the power constraints are in a long-term scale); ii) only one channel is selected.

From Lemma 6.1, it is clear that, if $n \in \mathcal{A}$, then n is unique and $x_n = B$ since the objective function of the sub-problem in (6.17) is an increasing function with respect to x_n. Substituting $x_n = B$ in equation (6.18b), we have

$$p_n = B\left(\frac{1}{\lambda + \mu_n h_n} - \frac{1}{g_n}\right) \tag{6.23}$$

[3]Note that a similar proof method is adopted in [97].

and the corresponding achieved utility of the sub-problem in (6.17) is

$$B\left(\frac{\lambda+\mu_n h_n}{g_n}-\ln\left(\frac{\lambda+\mu_n h_n}{g_n}\right)-1\right). \tag{6.24}$$

Since $p_n > 0$, we have $\frac{\lambda+\mu_n h_n}{g_n} < 1$, based on (6.23). The above achieved utility can be proved to be always positive when $\frac{\lambda+\mu_n h_n}{g_n} > 0$. On the other hand, the utility in equation (6.24) decreases if $\frac{\lambda+\mu_n h_n}{g_n}$ increases within range $(0,1)$. Therefore, if there is no channel that satisfies $\frac{\lambda+\mu_n h_n}{g_n} < 1$, then no channel is selected; otherwise, the channel with the minimum value of $\frac{\lambda+\mu_n h_n}{g_n}$ is selected.

Then the optimal solution for the sub-problem in (6.17) can be obtained in the procedure as follows.

Algorithm 6.1 Searching procedure for the optimal solution of the sub-problem in (6.17).

1: Define set $\mathcal{I} \triangleq \{n | \frac{\lambda+\mu_n h_n}{g_n} < 1\}$.
2: **if** $\mathcal{I} = \emptyset$ **then**
3: The maximal utility is 0, with $p_n = 0$ and $x_n = 0$, $\forall n \in \mathcal{N}$.
4: **else**
5: Find $n^* = \arg\min_{n\in\mathcal{I}} \left(\frac{\lambda+\mu_n h_n}{g_n}\right)$.
Output the optimal index n^*, with optimal bandwidth and power allocation strategy $x_{n^*} = B$ and $p_{n^*} = B\left(\frac{1}{\lambda+\mu_{n^*} h_{n^*}} - \frac{1}{g_{n^*}}\right)$, respectively, and the maximal utility $B\left(\frac{\lambda+\mu_{n^*} h_{n^*}}{g_{n^*}} - \ln\left(\frac{\lambda+\mu_{n^*} h_{n^*}}{g_{n^*}}\right) - 1\right)$.

It can be seen that the complexity of algorithm 6.1 is $O(N)$.

Remark: When the transmission- and interference-power constraints are both long-term, if there is no good channel to transmit at one moment, then it is reasonable not to select any channel at this moment, to save power for moments when there are good channels. And if there are good channels to transmit, then it is good to select the best channel. The good channel is the one on which the secondary BS can transmit more information to the secondary user while generating less interference to the primary user. Algorithm 6.1 indicates a metric to measure the overall quality of a channel, i.e., $\frac{\lambda+\mu_n h_n}{g_n}$. This metric is reasonable since it includes the information transmission channel g_n and the interference channel gain h_n, and reflects the opposite monotonicity of g_n and h_n with the metric. Interestingly, when the measure is below a fixed value 1, the secondary BS is considered to be with overall good channels.

6.3.2 Scenario with Short-term Transmission-power and Long-term Interference-power Constraints

Similar to Chapter 6.3.1, problem 6.1 in this scenario (with short-term transmission-power and long-term interference-power constraints) can be solved by investigating the following sub-problem for each realization of $(\boldsymbol{g}, \boldsymbol{h})$

$$\begin{aligned} \max_{\{p_n\},\,\{x_n\}} \quad & \sum_{n=1}^{N} x_n \ln\left(1+\frac{p_n g_n}{x_n}\right) - \sum_{n=1}^{N} \mu_n p_n h_n \\ \text{s.t.} \quad & \sum_{n=1}^{N} p_n \le P^{ST} && (6.25a) \\ & p_n \ge 0, x_n \ge 0,\ \forall n \in \mathcal{N};\ \sum_{n=1}^{N} x_n \le B && (6.25b) \end{aligned}$$

which is a convex problem satisfying Slater's condition. To analyze this sub-problem, the KKT condition is listed as follows.

$$\ln\left(1+\frac{p_n g_n}{x_n}\right) - \frac{p_n g_n}{x_n + p_n g_n} - \Gamma^* + \delta_n^* = 0,\ \forall n \in \mathcal{N} \tag{6.26a}$$

$$\frac{x_n g_n}{x_n + g_n p_n} - \lambda^* - \mu_n h_n + \Delta_n^* = 0,\ \forall n \in \mathcal{N} \tag{6.26b}$$

$$\lambda^* \left(\sum_{n=1}^{N} p_n - P^{ST}\right) = 0 \tag{6.26c}$$

$$\Delta_n^* p_n = 0, \delta_n^* x_n = 0,\ \forall n \in \mathcal{N};\ \Gamma^* \left(\sum_{n=1}^{N} x_n - B\right) = 0 \tag{6.26d}$$

$$\text{Constraints (6.25a) – (6.25b)} \tag{6.26e}$$

where λ^*, Δ_n^*, δ_n^*, and Γ^* are non-negative Lagrange multipliers associated with constraints $\sum_{n=1}^{N} p_n \le P^{ST}$, $p_n \ge 0$, $x_n \ge 0$, and $\sum_{n=1}^{N} x_n \le B$, respectively.

For the KKT condition, the following lemmas are in order.

<u>Lemma</u> 6.2. If $\mathcal{A} \triangleq \{j | p_j > 0, x_j > 0\} \neq \emptyset$,

- when the constraint $\sum_{n=1}^{N} p_n \le P^{ST}$ in (6.25a) is inactive[4], $|\mathcal{A}| = 1$.

[4] At the optimal point of the problem (say in (6.25)), if the equality in constraint $\sum_{n=1}^{N} p_n \le P^{ST}$ holds (i.e., $\sum_{n=1}^{N} p_n = P^{ST}$), we say the constraint is active; otherwise, we say it is inactive.

- when the constraint $\sum\limits_{n=1}^{N} p_n \leq P^{ST}$ in (6.25a) is active, $|\mathcal{A}| \leq 2$.

Proof. Please refer to Appendix D. $\square$

Lemma 6.3. For $\forall i, j \in \mathcal{A}$, $\text{SNR}_i = \text{SNR}_j$.

Proof. Please refer to Appendix E. $\square$

Lemma 6.2 and Lemma 6.3 indicate that, in the scenario, at most two channels are selected. And the SNR for all selected channels should be the same.

Define a new set $\mathcal{B} \overset{\triangle}{=} \{j | p_j = 0, x_j = 0\}$. Note that set $\mathcal{B}$ includes channels that are not selected. Therefore, $\mathcal{A} \bigcup \mathcal{B} = \{1, ..., N\}$.[5] Considering the conclusion of Lemma 6.3 and constraint $\sum_{n=1}^{N} x_n \leq B$, the common SNR indicated by Lemma 6.3 is shown to be $\text{SNR} = \frac{\sum\limits_{n \in \mathcal{A}} g_n p_n}{B}$. Then the utility of the sub-problem in (6.25) is

$$\begin{aligned}
& \sum_{n=1}^{N} x_n \ln\left(1 + \tfrac{p_n g_n}{x_n}\right) - \sum_{n=1}^{N} \mu_n p_n h_n \\
= \; & \sum_{n \in \mathcal{A}} x_n \ln\left(1 + \text{SNR}\right) - \sum_{n=1}^{N} \mu_n p_n h_n \\
= \; & B \ln\left(1 + \frac{\sum\limits_{n \in \mathcal{A}} g_n p_n}{B}\right) - \sum_{n=1}^{N} \mu_n p_n h_n \\
= \; & B \ln\left(1 + \frac{\sum\limits_{n=1}^{N} g_n p_n}{B}\right) - \sum_{n=1}^{N} \mu_n p_n h_n
\end{aligned} \tag{6.27}$$

where the second equality is because $\sum\limits_{n \in \mathcal{A}} x_n = B$ and the third equality is because $p_n = 0$ for $\forall n \in \mathcal{B}$ and $\mathcal{A} \bigcup \mathcal{B} = \mathcal{N}$. Then the sub-problem in (6.25) is equivalent to

$$\max_{\{p_n\},\, \{x_n\}} \quad B \ln\left(1 + \frac{\sum\limits_{n=1}^{N} p_n g_n}{B}\right) - \sum_{n=1}^{N} \mu_n p_n h_n$$

$$\text{s.t.} \quad \sum_{n=1}^{N} p_n \leq P^{ST} \tag{6.28a}$$

$$p_n \geq 0, \forall n \in \mathcal{N}. \tag{6.28b}$$

Recall that at most two channels are selected.

[5] Note that cases of "$p_j > 0, x_j = 0$" and "$p_j = 0, x_j > 0$" are omitted because neither of them achieves a utility more than that in the case "$p_j = 0, x_j = 0$" for the sub-problem in (6.25).

- When there are two selected channels, there are totally $\binom{N}{2} = \frac{N(N-1)}{2}$ possible combinations of the two channels. In a specific combination, denote the two selected channels as i and j. The problem in (6.28) is re-written as

$$\begin{aligned} \max_{p_i, p_j} \quad & B \ln\left(1 + \tfrac{p_i g_i + p_j g_j}{B}\right) - (\mu_i p_i h_i + \mu_j p_j h_j) \\ \text{s.t.} \quad & p_i \geq 0, p_j \geq 0, \\ & p_i + p_j = P^{ST} \end{aligned}$$

  the closed-form solution of which can be obtained straightforwardly based on corresponding KKT condition, and is omitted here.

- When there is only one selected channel, there are totally N possible cases for the selected channel. For each specific case, denote the selected channel as channel n. Set the derivative of the utility function of the problem in (6.28) as zero. Then, we can get the power allocation for channel n, given as

$$p_n = \left[\frac{B}{g_n}\left(\frac{g_n}{\mu_n h_n} - 1\right)\right]_0^{P^{ST}}$$

  where $[x]_a^b$ is defined as $\max(\min(b, x), a)$. Note that when $p_n = 0$, it means that no channel is selected.

Comparing the $\frac{N(N-1)}{2} + N$ optimal utilities for the above cases, we can obtain the maximal utility of problem in (6.28) (the largest among the $N + \frac{N(N-1)}{2}$ optimal utilities) and the associated power allocation strategy. Note that when all the $N + \frac{N(N-1)}{2}$ optimal utilities are non-positive, then no channel is selected. The computational complexity is $O(N^2)$.

For bandwidth allocation strategy, when there is only one selected channel (say channel n) in the above optimal power allocation, it can be seen that $x_n = B$, as the utility function in (6.25) is an increasing function of x_n. When there are two selected channels (say channel i and channel j), we have $x_i + x_j = B$ (since the utility function in (6.25) is an increasing function of x_i and x_j), which leads to $\text{SNR}_i = \text{SNR}_j = \frac{p_i g_i + p_j g_j}{B}$ (from Lemma 6.3), and further we have $x_i = \frac{p_i g_i}{\text{SNR}_i} = \frac{B p_i g_i}{p_i g_i + p_j g_j}$, and $x_j = \frac{B p_j g_j}{\text{SNR}_j} = \frac{B p_j g_j}{p_i g_i + p_j g_j}$.

Remark: Comparing with Chapter 6.3.1, the scenario in Chapter 6.3.2 has short-term transmission-power constraint, which reduces the flexibility in power and bandwidth allocation. The reduced flexibility is the main reason that at most one channel is selected in the scenario in Chapter 6.3.1, while two channels may be selected in the scenario in Chapter 6.3.2. The reduced flexibility is also the reason that, with short-term transmission-power constraint, we do not have that simple metric as in algorithm 6.1 to determine which channel(s) to be selected.

6.3.3 Scenario with Long-term Transmission-power and Short-term Interference-power Constraints

In this scenario, the sub-problem for every realization of $(\boldsymbol{g}, \boldsymbol{h})$ is

$$\max_{\{p_n\},\,\{x_n\}} \quad \sum_{n=1}^{N} x_n \ln\left(1 + \frac{p_n g_n}{x_n}\right) - \lambda \sum_{n=1}^{N} p_n$$

$$\text{s.t.} \quad p_n h_n \leq I_n^{ST},\ \forall n \in \mathcal{N} \tag{6.29a}$$

$$p_n \geq 0, x_n \geq 0,\ \forall n \in \mathcal{N};\ \sum_{n=1}^{N} x_n \leq B. \tag{6.29b}$$

The sub-problem in (6.29) is a convex problem satisfying Slater's condition. The KKT condition is

$$\ln\left(1 + \frac{p_n g_n}{x_n}\right) - \frac{p_n g_n}{x_n + p_n g_n} - \Gamma^* + \delta_n^* = 0,\ \forall n \in \mathcal{N} \tag{6.30a}$$

$$\frac{x_n g_n}{x_n + g_n p_n} - \lambda - \mu_n^* h_n + \Delta_n^* = 0,\ \forall n \in \mathcal{N} \tag{6.30b}$$

$$\mu_n^* \left(p_n h_n - I_n^{ST}\right) = 0,\ \forall n \in \mathcal{N} \tag{6.30c}$$

$$\Delta_n^* p_n = 0, \delta_n^* x_n = 0,\ \forall n \in \mathcal{N};\ \Gamma^* \left(\sum_{n=1}^{N} x_n - B\right) = 0 \tag{6.30d}$$

$$\text{Constraints (6.29a) – (6.29b)} \tag{6.30e}$$

where μ_n^*, Δ_n^*, δ_n^*, and Γ^* are non-negative Lagrange multipliers associated with constraints $p_n h_n \leq I_n^{ST}$, $p_n \geq 0$, $x_n \geq 0$, and $\sum_{n=1}^{N} x_n \leq B$, respectively.

Define $\mathcal{A} \triangleq \{j | p_j > 0,\ x_j > 0\} = \mathcal{A}_1 \cup \mathcal{A}_2$, where $\mathcal{A}_1 \triangleq \left\{ j | p_j = \frac{I_j^{ST}}{h_j}, x_j > 0 \right\}$ and $\mathcal{A}_2 \triangleq \left\{ j | 0 < p_j < \frac{I_j^{ST}}{h_j}, x_j > 0 \right\}$. Note that $\mathcal{A}$ includes the selected channels. From (6.9) it can be seen that $\mathcal{A}_1$ includes the selected channels with the maximal allowed interference power, while $\mathcal{A}_2$ includes the selected channels with the interference power less than the maximal allowed value. By analyzing the KKT condition in (6.30), a series of lemmas can be expected as follows.

Lemma 6.4. When $\mathcal{A} \neq \emptyset$, $|\mathcal{A}_2| \leq 1$.

The proof is similar to that of Lemma 6.1, and is omitted here. Note that in the proof, we have $\mu_n^* = 0, n \in \mathcal{A}_2$ (based on (6.30c)).

Lemma 6.5. For $\forall i, j \in \mathcal{A}$, $\text{SNR}_i = \text{SNR}_j$.

Proof. Please refer to the proof of Lemma 6.3. □

Lemma 6.6. For $\forall i \in \mathcal{A}_1$ and $\forall j \in \mathcal{A}_2$, $g_i \geq g_j$.

Proof. Please refer to Appendix F. □

Lemmas 6.4–6.6 indicate that, among all the selected channels, at most one channel is with interference power less than the maximal allowed value, while all other selected channels are with the maximal allowed interference power; and the channel with the interference power less than the maximal allowed value has less channel gain of g than other selected channels. All selected channels have the same SNR.

With the aid of Lemma 6.5 and by following the same procedure in Chapter 6.3.2, the sub-problem in (6.29) is equivalent to the following optimization problem

$$\begin{aligned} \max_{\{p_n\}} \quad & B \ln\left(1 + \frac{\sum_{n=1}^{N} p_n g_n}{B}\right) - \lambda \sum_{n=1}^{N} p_n \\ \text{s.t.} \quad & 0 \leq p_n \leq \frac{I_n^{ST}}{h_n}, \forall n \in \mathcal{N}. \end{aligned} \tag{6.31}$$

For the solution of the problem in (6.31), the following lemma is in order.

Lemma 6.7. For $\forall k \in \mathcal{B}$ and $\forall j \in \mathcal{A}$, $g_k \leq g_j$.

Proof. Please refer to Appendix G. □

From Lemmas 6.4, 6.6, and 6.7, it is clear that, in the optimal solution of the problem in (6.31), the set of all the channels (i.e., $\{1, ..., N\}$) can be partitioned into three sub-sets, $\mathcal{A}_1$ (in each channel of which the interference power reaches its maximal allowed value), $\mathcal{A}_2$ (where $|\mathcal{A}_2| \leq 1$, in the channel of which, if exists, the interference power does not reach the maximal allowed value) and $\mathcal{B}$ (which includes the channels that are not selected); and for $\forall i \in \mathcal{A}_1, j \in \mathcal{A}_2, k \in \mathcal{B}$, we have $g_i \geq g_j \geq g_k$. If we sort the N channels in descending order of the channel gains g_n's, i.e., $g_{s_1} \geq g_{s_2} \geq .... \geq g_{s_N}$, where $(s_1, ..., s_N)$ is a permutation of $(1, ..., N)$, then there are $N+1$ possible cases for the optimal solution of the problem in (6.31): in Case j, there are $j-1$ channels in set $\mathcal{A}_1$, detailed as follows.

- Case 1: $\mathcal{A}_1 = \emptyset$, $s_1 \in \mathcal{A}_2 \cup \mathcal{B}$, and $s_2, s_3, ..., s_N \in \mathcal{B}$, which means $p_{s_1} \in [0, \frac{I_{s_1}^{ST}}{h_{s_1}})$, and $p_{s_2} = p_{s_3} = ... = p_{s_N} = 0$.
- Case 2: $\mathcal{A}_1 = \{s_1\}$, $s_2 \in \mathcal{A}_2 \cup \mathcal{B}$, and $s_3, s_4, ..., s_N \in \mathcal{B}$, which means $p_{s_1} = \frac{I_{s_1}^{ST}}{h_{s_1}}$, $p_{s_2} \in [0, \frac{I_{s_2}^{ST}}{h_{s_2}})$, and $p_{s_3} = p_{s_4} = ... = p_{s_N} = 0$.
  $\vdots$
- Case j: $\mathcal{A}_1 = \{s_1, s_2, ..., s_{j-1}\}$, $s_j \in \mathcal{A}_2 \cup \mathcal{B}$, and $s_{j+1}, s_{j+2}, ..., s_N \in \mathcal{B}$, which means $p_{s_n} = \frac{I_{s_n}^{ST}}{h_{s_n}}$ for $i = 1, 2, ..., j-1$, $p_{s_j} \in [0, \frac{I_{s_j}^{ST}}{h_{s_j}})$, and $p_{s_{j+1}} = p_{s_{j+2}} = ... = p_{s_N} = 0$.
  $\vdots$
- Case $N+1$: $\mathcal{A}_1 = \{s_1, s_2, ..., s_N\}$, which means $p_{s_n} = \frac{I_{s_n}^{ST}}{h_{s_n}}$ for $n = 1, ..., N$.

Therefore, to solve the problem in (6.31), we need to search the $N+1$ cases only. We can first find the optimal[6] solutions in Cases 1, 2, ..., and $N+1$, respectively. Then the optimal solution with the largest objective function among the $N+1$ cases will be the global optimal solution for problem in (6.31).

Next we show how to find the optimal solutions for the $N+1$ cases.

In Case $j \leq N$, the power allocation values for all channels (except channel s_j) are known (for $i < j$, the interference power on channel s_i reaches its maximal allowed value; for $i > j$, channel s_j is not assigned power). Denote the power allocation value for channel s_j in Case j as φ_j, the optimal value of which is to be determined as follows.

For the problem in (6.31), the achieved utility in Case j is given as

$$B \ln \left(1 + \frac{\sum_{n=1}^{j-1} \frac{I_{s_n}^{ST} g_{s_n}}{h_{s_n}} + \varphi_j g_{s_j}}{B} \right) - \lambda \left(\sum_{n=1}^{j-1} \frac{I_{s_n}^{ST}}{h_{s_n}} + \varphi_j \right). \tag{6.32}$$

[6]Note that this optimal solution is limited to a particular case, not global optimal solution.

The optimal φ_j can be obtained by setting the derivative of (6.32) to zero, given as

$$\frac{Bg_{s_j}}{B + \sum\limits_{n=1}^{j-1} \frac{I_{s_n}^{ST} g_{s_n}}{h_{s_n}} + \varphi_j g_{s_j}} - \lambda = 0. \tag{6.33}$$

Generally after optimal utility values in the $N+1$ cases are obtained, the largest one is the optimal utility of the problem in (6.31).

Actually it may not be necessary to search all the $N+1$ cases. To demonstrate this, we take a look at (6.33) first, for which we have the following three observations for the optimal φ_j, denoted φ_j^*.

- When $B + \sum\limits_{n=1}^{j-1} \frac{I_{s_n}^{ST} g_{s_n}}{h_{s_n}} < \frac{Bg_{s_j}}{\lambda} < B + \sum\limits_{n=1}^{j} \frac{I_{s_n}^{ST} g_{s_n}}{h_{s_n}}$, we have

$$\varphi_j^* = \frac{1}{g_{s_j}} \left(\frac{Bg_{s_j}}{\lambda} - B - \sum_{n=1}^{j-1} \frac{I_{s_n}^{ST} g_{s_n}}{h_{s_n}} \right) \in \left(0, \frac{I_{s_j}^{ST}}{h_{s_j}} \right).$$

- When $\frac{Bg_{s_j}}{\lambda} \leq B + \sum\limits_{n=1}^{j-1} \frac{I_{s_n}^{ST} g_{s_n}}{h_{s_n}}$, the solution for (6.33) is $\frac{1}{g_{s_j}} \left(\frac{Bg_{s_j}}{\lambda} - B - \sum\limits_{n=1}^{j-1} \frac{I_{s_n}^{ST} g_{s_n}}{h_{s_n}} \right)$ ≤ 0. Since the feasible region of φ_j is $[0, \frac{I_{s_j}^{ST}}{h_{s_j}})$, the optimal solution for Case j is $\varphi_j^* = 0$.

- When $\frac{Bg_{s_j}}{\lambda} \geq B + \sum\limits_{n=1}^{j} \frac{I_{s_n}^{ST} g_{s_n}}{h_{s_n}}$, the solution for (6.33) is $\frac{1}{g_{s_j}} \left(\frac{Bg_{s_j}}{\lambda} - B - \sum\limits_{n=1}^{j-1} \frac{I_{s_n}^{ST} g_{s_n}}{h_{s_n}} \right)$ $\geq \frac{I_{s_j}^{ST}}{h_{s_j}}$. However, the feasible region of φ_j is $[0, \frac{I_{s_j}^{ST}}{h_{s_j}})$. It can be seen that the maximal utility in the feasible region in Case j is less than the utility when $\varphi_j = \frac{I_{s_j}^{ST}}{h_{s_j}}$. The utility in the latter case is not greater than the maximal utility in Case $j+1$. Therefore, we can virtually set $\varphi_j^* = \frac{I_{s_j}^{ST}}{h_{s_j}}$.

Define $j^\dagger = \arg\min_j \left(\varphi_j^* \in [0, \frac{I_{s_j}^{ST}}{h_{s_j}}) \right)$. The following lemma is in order.

Lemma 6.8. The maximal utility in Case $j^\dagger$ is the maximal utility of problem in (6.31).

Proof. Please refer to Appendix H. □

Lemma 6.8 tells that, to obtain the optimal solution for problem (6.31), we need only search until the first case (say Case j) such that φ_j^* is not equal to $\frac{I_{s_j}^{ST}}{h_{s_j}}$. Accordingly, the optimal solution of the problem in (6.31) can be searched as follows.

Algorithm 6.2 Searching procedure for the optimal solution of the problem in (6.31).

1: Sort the N channels in descending order of the channel gains g_n's, i.e., $g_{s_1} \geq g_{s_2} \geq .... \geq g_{s_N}$.

2: Find $j^\dagger = \arg\min_n \left(\varphi_n^* < \frac{I_{s_n}^{ST}}{h_{s_n}}\right)$

3: Output the optimal power configuration: $p_{s_j} = \frac{I_{s_j}^{ST}}{h_{s_j}}$ for $0 < j < j^\dagger$, $p_{s_j} = \frac{1}{g_{s_j}}\left(\frac{Bg_{s_j}}{\lambda} - B - \sum\limits_{n=1}^{j-1}\frac{I_{s_n}^{ST}g_{s_n}}{h_{s_n}}\right)$ for $j = j^\dagger$, and $p_{s_j} = 0$ for $j > j^\dagger$. The maximal utility is $B\ln\left(1 + \frac{\sum\limits_{n=1}^{j^\dagger} p_{s_n}g_{s_n}}{B}\right) - \lambda\sum\limits_{n=1}^{j^\dagger} p_{s_n}$. The bandwidth configuration is $x_{s_j} = \frac{Bp_{s_j}g_{s_j}}{\sum\limits_{n=1}^{j^\dagger} p_{s_n}g_{s_n}}$, $j \in \mathcal{N}$.

In Step 3 of the algorithm, the bandwidth configuration is from the fact that, for $\forall j \in \mathcal{A}$, $\text{SNR}_j = \frac{\sum\limits_{n\in\mathcal{A}} g_n p_n}{B}$.

It can be seen that the complexity of the algorithm is $O(N)$.

Remark: Compared to the scenario in Chapter 6.3.1, the short-term transmission-power constraint in Chapter 6.3.2 changes the upper bound of the number of selected channels from one to two, while the short-term interference-power constraint in Chapter 6.3.3 makes all selected channels except one to be with their interference powers at maximal allowed value. The difference of the impact of short-term transmission-power constraint and short-term interference-power constraint lies in the fact that the transmission power is used to transmit over all selected channels while one interference power constraint only imposes a limit on one channel. Interestingly, with short-term interference-power constraints, channels with larger channel gains g_n's are selected. Note that it does not mean the channel selection is independent of h_n's, since h_n's affect how many channels are selected.

6.3.4 Scenario with Short-term Transmission-power and Short-term Interference-power Constraints

Similar to previous subsections, the following sub-problem is expected to be solved for every realization of $(\boldsymbol{g}, \boldsymbol{h})$ in this scenario.

$$\begin{aligned}
\max_{\{p_n\},\{x_n\}} \quad & \sum_{n=1}^{N} x_n \ln\left(1+\frac{p_n g_n}{x_n}\right) \\
\text{s.t.} \quad & p_n h_n \le I^{ST}, \forall n \in \mathcal{N} && \text{(6.34a)} \\
& \sum_{n=1}^{N} p_n \le P^{ST}, && \text{(6.34b)} \\
& p_n \ge 0, x_n \ge 0, \ \forall n \in \mathcal{N}; \ \sum_{n=1}^{N} x_n \le B. && \text{(6.34c)}
\end{aligned}$$

The KKT condition is written as follows.

$$\ln(1+\frac{p_n g_n}{x_n}) - \frac{p_n g_n}{x_n + p_n g_n} - \Gamma^* + \delta_n^* = 0, \forall n \in \mathcal{N} \quad \text{(6.35a)}$$

$$\frac{x_n g_n}{x_n + g_n p_n} - \lambda^* - \mu_n^* h_n + \Delta_n^* = 0, \forall n \in \mathcal{N} \quad \text{(6.35b)}$$

$$\lambda^* \left(\sum_{n=1}^{N} p_n - P^{ST}\right) = 0 \quad \text{(6.35c)}$$

$$\mu_n^* \left(p_n h_n - I_n^{ST}\right) = 0, \forall n \in \mathcal{N} \quad \text{(6.35d)}$$

$$\Delta_n^* p_n = 0, \delta_n^* x_n = 0, \forall n \in \mathcal{N}; \Gamma^* \left(\sum_{n=1}^{N} x_n - B\right) = 0 \quad \text{(6.35e)}$$

$$\text{Constraints (6.34a) – (6.34c)} \quad \text{(6.35f)}$$

where μ_n^*, λ^*, Δ_n^*, δ_n^*, and Γ^* are non-negative Lagrange multipliers associated with constraints $p_n h_n \le I^{ST}$, $\sum_{n=1}^{N} p_n \le P^{ST}$, $p_n \ge 0$, $x_n \ge 0$, and $\sum_{n=1}^{N} x_n \le B$, respectively.

Define $\mathcal{A} = \{j | p_j > 0, x_j > 0\} = \mathcal{A}_1 \bigcup \mathcal{A}_2$ where $\mathcal{A}_1 = \{j | p_j = \frac{I_j^{ST}}{h_j}, x_j > 0\}$ and $\mathcal{A}_2 = \{j | 0 < p_j < \frac{I_j^{ST}}{h_j}, x_j > 0\}$. Define $\mathcal{B} = \{j | p_j = 0, x_j = 0\}$. So $\mathcal{A}$, $\mathcal{A}_1$, $\mathcal{A}_2$, and $\mathcal{B}$ have the same physical meanings as in Chapter 6.3.3.

To solve the sub-problem in (6.34), the following lemmas are in order.

Lemma 6.9.

- When the constraint in (6.34b) is inactive, $|\mathcal{A}_2| = 0$.

- When the constraint in (6.34b) is active, the associated Lagrange multiplier $\lambda^* > 0$, and $|\mathcal{A}_2| \leq 1$.

Proof. Please refer to Appendix I. $\square$

Similar to Lemmas 6.5-6.7, we have the following lemmas (with proofs omitted).

Lemma 6.10. For $\forall i, j \in \mathcal{A}, \text{SNR}_i = \text{SNR}_j$.

Lemma 6.11. For $\forall i \in \mathcal{A}_1$ and $\forall j \in \mathcal{A}_2$, $g_i \geq g_j$.

Lemma 6.12. For $\forall j \in \mathcal{A}$ and $\forall k \in \mathcal{B}$, $g_j \geq g_k$.

Similar to Chapter 6.3.3, the sub-problem in (6.34) can be reduced (according to Lemma 6.10) to the following problem

$$\max_{\{p_n\}} \quad B \ln \left(1 + \frac{\sum\limits_{n=1}^{N} p_n g_n}{B} \right)$$

$$\text{s.t.} \quad 0 \leq p_n \leq \frac{I_n^{ST}}{h_n}, \forall n \in \mathcal{N} \tag{6.36a}$$

$$\sum_{n=1}^{N} p_n \leq P^{ST}. \tag{6.36b}$$

It can be easily proved that the problem in (6.36) can be solved as follows (the proof is omitted). First sort all the channels, $\{1, ..., N\}$, based on the descending order of channel gains g_n's such that $g_{s_1} \geq g_{s_2} \geq ... \geq g_{s_N}$. Denote $i^* = \arg\min_i \sum\limits_{n=1}^{i} \frac{I_{s_n}^{ST}}{h_{s_n}} > P^{ST}$.[7] Then the solution for the problem in (6.36) is: $p_{s_n} = \frac{I_{s_n}^{ST}}{h_{s_n}}, n = 1, ..., i^* - 1$; $p_{s_{i^*}} = P^{ST} - \sum\limits_{n=1}^{i^*-1} \frac{I_{s_n}^{ST}}{h_{s_n}}$; and $p_{s_n} = 0, n = i^* + 1, i^* + 2, ..., N$; $x_{s_n} = \frac{B p_{s_n} g_{s_n}}{\sum\limits_{n=1}^{i^*} p_{s_n} g_{s_n}}, n \in \mathcal{N}$. The computational complexity is $O(N)$.

Remark: Similar to Chapter 6.3.3, channels with larger g_n's are selected, and all selected channels except one are with the maximal allowed interference power. The difference is that the short-term transmission-power constraint in Chapter 6.3.4 makes the power allocation simpler, i.e., the short-term transmission power is assigned to channels according to the descending order of g_n's, until all the short-term transmission power is used up. The

[7]Note that if $\sum\limits_{n=1}^{N} \frac{I_{s_n}^{ST}}{h_{s_n}} < P^{ST}$, then $p_{s_n} = \frac{I_{s_n}^{ST}}{h_{s_n}}, n = 1, ..., N$.

simplicity is because, with short-term transmission-power constraint, when the channels are not good at a moment, we do not need to save transmission power for other moments with good channels.

6.4 The Case with Peak Power Spectrum Density Constraints

In this section, we consider the case when additional constraints of peak power spectrum density are imposed on the transmission power at the secondary BS and the interference power received at the primary user. Denote the peak power spectrum density constraints for the transmission power and interference power on channel n as P^{PEAK} and I_n^{PEAK}, respectively. Then, we have

$$\frac{p_n}{x_n} \leq P^{\text{PEAK}}, \forall n \in \mathcal{N} \tag{6.37}$$

and

$$\frac{p_n h_n}{x_n} \leq I_n^{\text{PEAK}}, \forall n \in \mathcal{N}. \tag{6.38}$$

The constraints (6.37) and (6.38) can be combined to be

$$\frac{p_n}{x_n} \leq U_n^{\text{PEAK}}, \forall n \in \mathcal{N} \tag{6.39}$$

where $U_n^{\text{PEAK}} = \min\left(P^{\text{PEAK}}, \frac{I_n^{\text{PEAK}}}{h_n}\right)$ is the minimal peak value of power spectrum density on channel n. Then the average rate maximization problem in this chapter with additional peak power spectrum density constraints can be formulated as problem 6.1 with additional constraints (6.39). Again, the problem can be decomposed into two levels, and next we focus on the sub-problems in the lower level. We still consider the four scenarios as in Chapter 6.3. Fast algorithms with closed-form solutions may exist in scenarios with long-term transmission-power and long-term interference-power constraints and with short-term transmission-power and long-term interference-power constraints, as shown in the following two subsections, respectively. However, for the other two scenarios, fast algorithms with closed-form solutions cannot be found, and thus, traditional methods (such as sub-gradient method and interior-point algorithm) may have to be resorted to.

6.4.1 Scenario with Long-term Source-power and Long-term Interference-power Constraints

In this scenario, a sub-problem as in (6.17) with additional constraints in (6.39) is supposed to be solved. Similar to Chapter 6.3, we define $\mathcal{A} \stackrel{\triangle}{=} \{j|p_j > 0, x_j > 0\}$ and $\mathcal{B} \stackrel{\triangle}{=} \{j|p_j = 0, x_j = 0\}$. It can be seen that $\mathcal{A} \cup \mathcal{B} = \{1, ..., N\}$. Denote $\mathcal{A} = \mathcal{C} \cup \mathcal{D}$, where $\mathcal{C} \stackrel{\triangle}{=} \left\{j|0 < \frac{p_j}{x_j} < U_j^{\text{PEAK}}, x_j > 0\right\}$ (i.e., the set of selected channels with transmission power spectrum density less than the minimal peak value) and $\mathcal{D} \stackrel{\triangle}{=} \left\{j|\frac{p_j}{x_j} = U_j^{\text{PEAK}}, x_j > 0\right\}$ (i.e., the set of selected channels with transmission power spectrum density equal to the minimal peak value). It can be proved using the similar method to that in Lemma 6.1 that

Lemma 6.13. If $\mathcal{C} \neq \emptyset$, then $|\mathcal{C}| \leq 1$.

Further, we have the following two lemmas.

Lemma 6.14. If $\mathcal{D} \neq \emptyset$, then $|\mathcal{D}| \leq 1$.

Proof. For $j \in \mathcal{D}$, $p_j = x_j U_j^{\text{PEAK}}$. The sub-problem can be rewritten as

$$\begin{array}{ll} \max\limits_{\{p_i\},\,\{x_i\}} & \sum\limits_{i\in\mathcal{C}} x_i \ln\left(1 + \frac{p_i g_i}{x_i}\right) - \lambda \sum\limits_{i\in\mathcal{C}} p_i - \sum\limits_{i\in\mathcal{C}} \mu_i p_i h_i \\ & + \sum\limits_{j\in\mathcal{D}} x_j \left(\ln\left(1 + U_j^{\text{PEAK}} g_j\right) - \lambda U_j^{\text{PEAK}} - \mu_j U_j^{\text{PEAK}} h_j\right) \\ \text{s.t.} & \sum\limits_{i\in\mathcal{C}} x_i + \sum\limits_{j\in\mathcal{D}} x_j \leq B. \end{array} \tag{6.40}$$

When $\mathcal{D} \neq \emptyset$,

$$\max_{j\in\mathcal{D}} \left(\ln\left(1 + U_j^{\text{PEAK}} g_j\right) - \lambda U_j^{\text{PEAK}} - \mu_j U_j^{\text{PEAK}} h_j\right) > 0.$$

Define

$$j^* = \arg\max_j \left(\ln\left(1 + U_j^{\text{PEAK}} g_j\right) - \lambda U_j^{\text{PEAK}} - \mu_j U_j^{\text{PEAK}} h_j\right).$$

If $|\mathcal{D}| > 1$, then it will achieve a higher objective function of (6.40) if the bandwidth assignments for channels in $|\mathcal{D}|$ are instead all assigned to channel j^* only. Therefore, in the optimal solution, we have $|\mathcal{D}| = 1$. □

Similar to proof of Lemma 6.14, we can prove the following lemma.

Lemma 6.15. $|\mathcal{A}| \leq 1$.

According to Lemma 6.15, at most one channel is selected. Suppose the only selected channel, if exists, is channel n. To maximize the objective function in (6.17), x_n should be set as B. By setting the derivative of the objective function in (6.17) with respect to p_n as 0, p_n can be obtained as $B\left(\frac{1}{\lambda+\mu_n h_n}-\frac{1}{g_n}\right)$. Considering the constraint $p_n \geq 0$ and $\frac{p_n}{x_n} \leq U_n^{\text{PEAK}}$ further, the optimal power configuration should be $p_n = B\left[\frac{1}{\lambda+\mu_n h_n}-\frac{1}{g_n}\right]_0^{U_n^{\text{PEAK}}}$.

In summary, the optimal solution of the problem in (6.17) with additional constraints in (6.39) can be obtained as follows.

Algorithm 6.3 Searching procedure for the optimal solution of the problem in (6.17) with additional constraints in (6.39).

1: Define $V_n = B\ln\left(1+\left[\frac{1}{\lambda+\mu_n h_n}-\frac{1}{g_n}\right]_0^{U_n^{\text{PEAK}}} g_n\right) - \lambda B\left[\frac{1}{\lambda+\mu_n h_n}-\frac{1}{g_n}\right]_0^{U_n^{\text{PEAK}}} - \mu_n B\left[\frac{1}{\lambda+\mu_n h_n}-\frac{1}{g_n}\right]_0^{U_n^{\text{PEAK}}} h_n$.

2: **if** $\max_n V_n \leq 0$ **then**

3: The maximal utility is 0, with $p_n = 0$ and $x_n = 0$, $\forall n \in \mathcal{N}$.

4: **else**

5: Find $i^* = \arg\max_n V_n$.

Output the optimal index i^*, with optimal bandwidth and power allocation strategy $x_{i^*} = B$ and $p_{i^*} = B\left[\frac{1}{\lambda+\mu_{i^*} h_{i^*}}-\frac{1}{g_{i^*}}\right]_0^{U_{i^*}^{\text{PEAK}}}$, respectively, and the maximal utility V_{i^*}.

It can be seen that the complexity of the algorithm is $O(N)$.

6.4.2 Scenario with Short-term Transmission-power and Long-term Interference-power Constraints

In this scenario, a sub-problem as in (6.25) with additional constraints (6.39) is going to be solved. For the sub-problem, still define $\mathcal{A} \triangleq \{j|p_j > 0, x_j > 0\}$ and $\mathcal{A} = \mathcal{C} \cup \mathcal{D}$, where $\mathcal{C} \triangleq \left\{j|0 < \frac{p_j}{x_j} < U_j^{\text{PEAK}}, x_j > 0\right\}$ and $\mathcal{D} \triangleq \left\{j|\frac{p_j}{x_j} = U_j^{\text{PEAK}}, x_j > 0\right\}$. The following lemmas can be expected.

Lemma 6.16. If $\mathcal{C} \neq \emptyset$,

- when the constraint $\sum_{n=1}^{N} p_n \leq P^{ST}$ in (6.25a) is inactive, $|\mathcal{C}| = 1$.
- when the constraint $\sum_{n=1}^{N} p_n \leq P^{ST}$ in (6.25a) is active, $|\mathcal{C}| \leq 2$.

Lemma 6.17. If $\mathcal{D} \neq \emptyset$, then $|\mathcal{D}| \leq 1$.

Lemma 6.18. When the constraint $\sum\limits_{n=1}^{N} p_n \leq P^{ST}$ in (6.25a) is inactive, $|\mathcal{A}| \leq 1$.

The proofs of the three lemmas are similar to those of Lemmas 6.2, 6.14, and 6.15, respectively, and thus, are omitted.

According to Lemma 6.18, when the constraint $\sum\limits_{n=1}^{N} p_n \leq P^{ST}$ is inactive, at most one channel is selected. Suppose the only selected channel, if exists, is channel n. Similar to Chapter 6.4.1, the optimal bandwidth and power configuration is $x_n = B$ and $p_n = \left[\frac{B}{g_n}\left(\frac{g_n}{\mu_n h_n} - 1\right)\right]_0^{U_n^{\text{PEAK}}}$.

In summary, when $\sum\limits_{n=1}^{N} p_n \leq P^{ST}$ is inactive (e.g., a sufficient condition for this is $\max_n(BU_n^{\text{peak}}) < P^{ST}$), the optimal solution of the sub-problem in (6.25) with additional constraints in (6.39) can be obtained as follows.

Algorithm 6.4 Searching procedure for the optimal solution of the sub-problem in (6.25) with additional constraints in (6.39) (when $\sum_{n=1}^{N} p_n \leq P^{ST}$ is inactive).

1: Define $Z_n = B \ln\left(1 + \left[\left(\frac{g_n}{\mu_n h_n} - 1\right)\right]_0^{U_n^{\text{PEAK}}}\right) - \mu_n \left[\frac{B}{g_n}\left(\frac{g_n}{\mu_n h_n} - 1\right)\right]_0^{U_n^{\text{PEAK}}} h_n$.

2: **if** $\max_n Z_n \leq 0$ **then**

3: The maximal utility is 0, with $p_n = 0$ and $x_n = 0$, $\forall n \in \mathcal{N}$.

4: **else**

5: Find $i^* = \arg\max\limits_n Z_n$.

Output the optimal index i^*, with optimal bandwidth and power allocation strategy $x_{i^*} = B$ and $p_{i^*} = \left[\frac{B}{g_{i^*}}\left(\frac{g_{i^*}}{\mu_{i^*} h_{i^*}} - 1\right)\right]_0^{U_{i^*}^{\text{PEAK}}}$, respectively, and the maximal utility Z_{i^*}.

It can be seen that the complexity of the algorithm is $O(N)$.

When $\sum_{n=1}^{N} p_n \leq P^{ST}$ is active, fast algorithms with closed-form solutions cannot be found, and therefore, traditional methods (such as sub-gradient method and interior-point algorithm) may have to be resorted to.

Table 6.1 summarizes the number of selected channels and algorithm complexity for the six scenarios discussed in Chapter 6.3 and Chapter 6.4, denoted as LT-LT, ST-LT, LT-ST, ST-ST, LT-LT (PK) and ST-LT (PK), respectively. Here the LT/ST before the hyphen means long-term/short-term transmission-power constraint, while LT/ST after the hyphen means long-term/short-term interference-power constraint, and PK means that peak power

TABLE 6.1
NUMBER OF SELECTED CHANNELS AND ALGORITHM COMPLEXITY

Scenario	Number of selected channels	Algorithm complexity
LT-LT	0 or 1	$O(N)$
ST-LT	$0, 1,$ or 2	$O(N^2)$
LT-ST	$0, 1, ..., N-1,$ or N	$O(N)$
ST-ST	$1, ..., N-1,$ or N	$O(N)$
LT-LT (PK)	0 or 1	$O(N)$
ST-LT (PK) (with $\sum_{n=1}^{N} p_n \leq P^{ST}$ inactive)	0 or 1	$O(N)$

spectrum density constraints are imposed.

6.5 Implementation Issue - Online Calculation of Resource Allocation

Take the scenario in Chapter 6.3.1 as an example. To update λ and μ_n's as in (6.15) and (6.16), $\left(P^{LT} - \mathbb{E}\left[\sum_{n=1}^{N} p_n(t)\right]\right)$ and $\left(I_n^{LT} - \mathbb{E}\left[p_n(t)h_n\right]\right)$ should be calculated, in which the two expectation terms require the resource allocation solution of the sub-problem in (6.17) for any possible realization of channel gains g_n's and h_n's. Thus, the computational complexity is high. In addition, PDF of the channel gains $f(\boldsymbol{g}, \boldsymbol{h})$ is required to be known in advance. To address these problems, an online (fading block-by-fading block) calculation method of resource allocation is adopted, similar to that in [98].

Step 1: Initialize $\lambda(0)$ and $\mu_n(0)$ $(n \in \mathcal{N})$.

Step 2: At the beginning of the t^{th} fading block, the secondary BS measures channel gains $g_n(t)$'s and $h_n(t)$'s $(n \in \mathcal{N})$, solves sub-problem (6.17) for that particular channel gain realization and gets solutions $p_n(t)$ and updates Lagrange multipliers

$$\lambda(t+1) = \left(\lambda(t) - a(t)\left(P^{LT} - \sum_{n=1}^{N} p_n(t)\right)\right)^{+} \tag{6.41}$$

$$\mu_n(t+1) = \left(\mu_n(t) - a(t)\left(I_n^{LT} - p_n(t)h_n(t)\right)\right)^{+}, \forall n \in \mathcal{N}. \tag{6.42}$$

It is proved in [98] that $\lambda(t)$ and $\mu_n(t)$'s will converge to the optimal λ and μ_n's values with probability 1, under the conditions 1) $a(t) \geq 0$, $\sum_{t=0}^{+\infty} a(t) = +\infty$, and $\sum_{t=0}^{+\infty} a(t)^2 < +\infty$, and 2) $(P^{LT} - \sum_{n=1}^{N} p_n(t))$ and $(I_n^{LT} - p_n(t)h_n(t))$ are bounded.

Compared with (6.15) and (6.16), expectation terms do not exist in (6.41) and (6.42), thus leading to much less complexity. In addition, the new updating method also works when the channel statistics information $f(\boldsymbol{g}, \boldsymbol{h})$ is unknown, because 1) the impact of $f(\boldsymbol{g}, \boldsymbol{h})$ is represented by the realizations of the channel gains g_n's and h_n's ($n \in \mathcal{N}$) in continuous fading blocks, and 2) the realizations of channel gains can be measured at the beginning of the fading blocks.

Note that the updating method in (6.41) and (6.42) is called *stochastic sub-gradient method*, broadly used for solving stochastic optimization problems. More details for this method can be found in references [98], [99] and [100].

6.6 Numerical Results

In this section, numerical results are given to show the performance of the algorithms for the four scenarios, LT-LT, ST-LT, LT-ST and ST-ST, discussed in Chapter 6.3. Peak power spectrum density constraints are not considered. The channel gains g_n's, h_n's ($n \in \mathcal{N}$) are independent and exponentially distributed (i.e., Rayleigh fading channels) with mean being 1. The channel gain distributions are supposed to be unknown by any node. So the stochastic sub-gradient method introduced in Chapter 6.5 is used to update Lagrange multipliers λ and μ_n's iteratively over fading blocks. All the initial values of Lagrange multipliers are set as 1. The step size $a(t)$ in (6.41) and (6.42) is selected as $a(t) = \frac{1}{t+200}$. Throughout this section B is set as 1.

6.6.1 Convergence of Lagrange Multipliers

In this subsection, the convergence of Lagrange multipliers λ and μ_i's is illustrated (note that for the original optimization problem, short-term constraints do not have Lagrange multipliers). In this numerical example, the long-term and short-term constraint of transmission power, P^{LT} and P^{ST}, are set to have a value equal to the number of channels, N. For example, when there are 4 channels, both P^{LT} and P^{ST} are set as 4. The long-term and short-term constraint of interference power, I_i^{LT} and I_i^{ST}, are set as 1.

Fig. 6.1, Fig. 6.2 and Fig. 6.3 show the updating iterations of the Lagrange multipliers as in (6.41) and (6.42) in different scenarios with 4 channels, 8 channels and 16 channels, respectively. Note that in Fig. 6.2 and Fig. 6.3, legends are not illustrated due to the

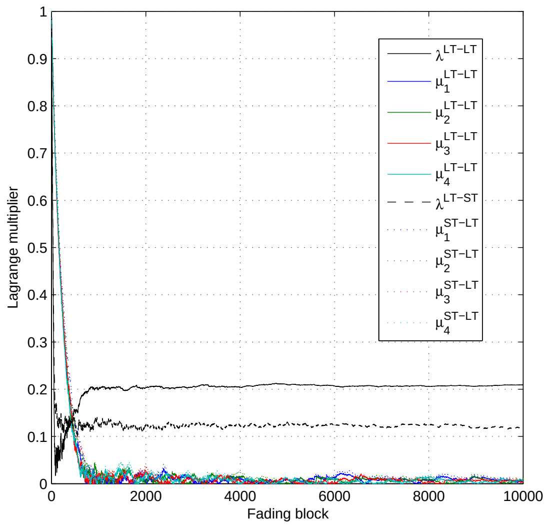

Fig. 6.1. Convergence of Lagrange multipliers for the average rate maximization problem with 4 relays.

large number of Lagrange multipliers. It can be seen from the figures that all the Lagrange multipliers converge within a few thousand fading blocks. In particular, all the Lagrange multipliers in Fig. 6.1, Fig. 6.2 and Fig. 6.3 converge to nonzero values, which means that the associated long-term constraints are all active. It can be also seen from the three figures that the number of channels has negligible impact on the convergence speed.

6.6.2 Average Rates

Consider 4 channels. When the interference-power constraint (either long-term or short-term) is fixed as 0.25, Fig. 6.4 shows the average rate, where the horizontal axis means the (long-term of short-term) transmission-power constraint. It can be seen that, with the same amount of power constraints, the LT-LT scenario always has the largest average rate, while the ST-ST scenario has the smallest average rate. This is because a long-term power constraint allows more flexibility than a short-term power constraint does, and therefore, leads to more efficient resource allocation.

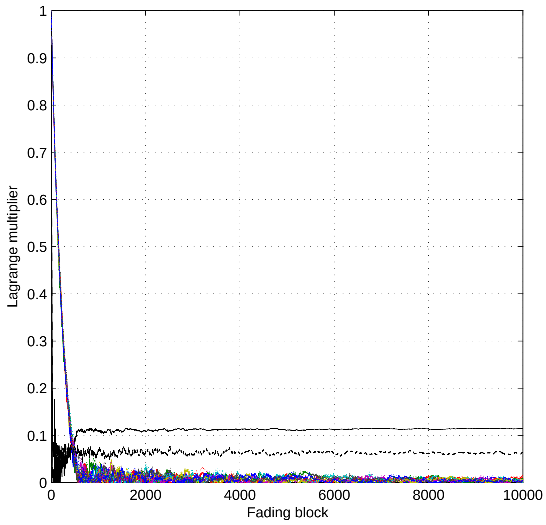

Fig. 6.2. Convergence of Lagrange multipliers for the average rate maximization problem with 8 relays.

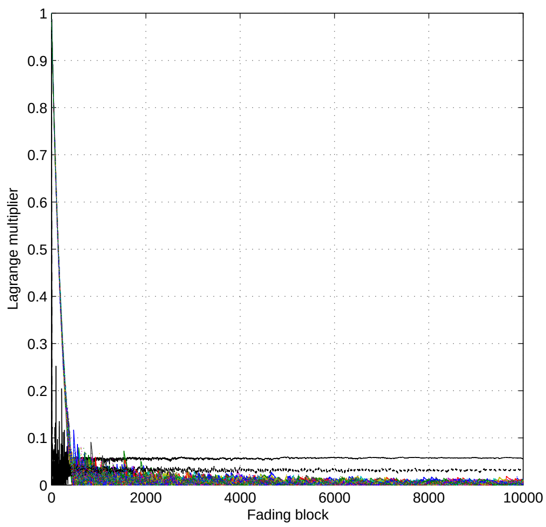

Fig. 6.3. Convergence of Lagrange multipliers for the average rate maximization problem with 16 relays.

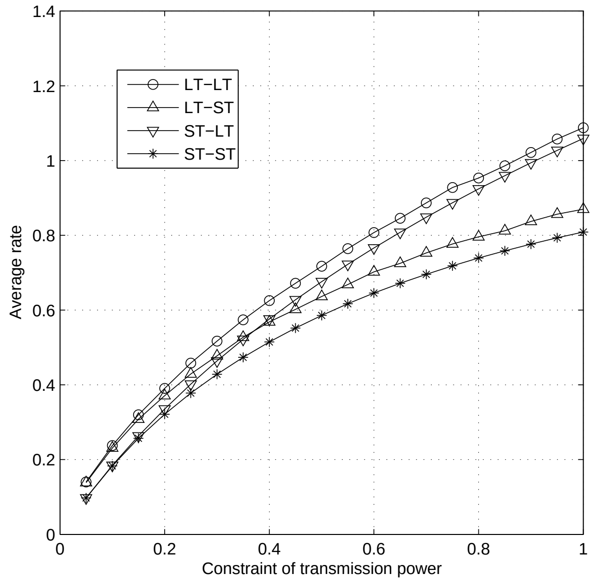

Fig. 6.4. Average rate vs. transmission-power constraint (interference-power constraint is 0.25).

6.7 Conclusion

In this chapter, the average rate maximization problem given constraints on transmission power and interference power in a cognitive radio network in underlay mode has been studied. With a consideration of either long-term or short-term constraint imposed on the transmission power and interference power, the average rate maximization under four scenarios of constraints on transmission power and interference power is investigated. In each scenario, the average rate maximization problem is decomposed to sub-problems, each corresponding to a realization of channel gains. For each sub-problem, special properties are found and a fast algorithm with closed-form resource allocation solution is given. This research should provide helpful insight to the design of cognitive radio networks in underlay mode over slow fading channels.

Chapter 7

Conclusions and Future Work

In this chapter the contributions of this thesis are summarized and some future research topics are indicated.

7.1 Conclusions

In this thesis, the problems of spectrum sensing and resource allocation in cognitive radio are investigated, to maximize the resource efficiency while protecting the activities of primary users. A number of optimization problems are formulated and then solved optimally, or sub-optimally but with comparable performance to optimal solutions.

In Chapter 3, the optimal multichannel cooperative spectrum sensing strategies in cognitive radio networks are studied to maximize the average throughput of all the secondary users. The strategies tell how to determine the total sensing time and how to distribute the total sensing time to different channels in cooperative data fusion spectrum sensing. By proving some interesting properties of the research problems, the original research problems, which are non-convex, are transformed to bi-level convex optimization problems. Polynomial-complexity algorithms are proposed to solve them optimally.

In Chapter 4, research efforts to achieve the maximal spectrum efficiency through selection of sensing orders are presented. Two cases are investigated: the case without adaptive modulation and the case with adaptive modulation. For the case without adaptive modulation, low-complexity algorithms are proposed and numerically verified to have comparable performance to an optimal solution. For the case with adaptive modulation, our observations imply that the sensing order setting and stopping rule should be jointly designed from a systematic point of view.

In Chapter 5, the joint problem of optimal multichannel cooperative sensing and resource allocation (of spectrum and power) in cognitive radio networks is explored. The research focus is on how to set spectrum sensing time, how to determine the spectrum sensing threshold, how to allocate spectrum resources, and how to set transmission power. A weighted throughput maximization problem, a proportional fairness problem, and a max-min fairness problem are formulated. Although the problems are nonconvex, a combination of bilevel optimization and monotonic programming is proposed to solve the problems optimally.

In Chapter 6, the average rate maximization problem in a cognitive radio network in underlay mode is studied. With a consideration of either long-term or short-term constraint imposed on the transmission power of secondary users and interference power to primary users, the average rate maximization under four scenarios of constraint combinations is investigated. In each scenario, the average rate maximization problem is decomposed to sub-problems, each corresponding to a realization of channel gains. For each sub-problem, special properties are found and a fast algorithm, with closed-form resource allocation solution, is given.

7.2 Future Work

In the optimal multichannel cooperative spectrum sensing discussed in Chapter 3, it is assumed that a secondary user can send its test statistic to the coordinator. An interesting research topic is to investigate the case when a secondary user sends a quantized version of its test statistic to the coordinator or sends its detection decision on the presence or absence of primary activities (i.e., 1-bit information is sent to the coordinator). It is also interesting to study the problem from a game theoretical point of view, in which each secondary user is assumed to be selfish but rational. Moreover, in the problem formulation, the constraint $P_n^f \leq 0.5$ may lead to some performance loss. For instance, for a specific channel (say channel n), consider the situation that the channel gains from the primary user to secondary users and from the secondary transmitter that is assigned channel n to its receiver are significantly lower than the channel gains in other channels. Then channel n will dominate the sensing time, but the secondary throughout in this channel is very low. A possible solution is to sense only a subset of the N channels. However, the selection of the subset depends on the channel SNR values and the channel available probabilities. How to find an optimal subset of channels is an interesting research topic for further investigation.

In the sensing order problem discussed in Chapter 4, for each channel, the idle/busy state of the channel in a slot is independent of the idle/busy states in other slots. On the other hand, when the channel idle/busy states are correlated (e.g., modeled by a two-state Markov chain), the sensing of a channel in a slot has two gains [74], [75]: short-term gain that is the transmission opportunity if the channel is sensed free, and long-term gain that is the updated state information of the channel (the updated state information can benefit the secondary user in later sensing selection). The proposed algorithm in this chapter can maximize the short-term gain in a slot. However, an optimal solution needs to strike a balance between the short-term gain and long-term gain. This means the user may still need to sense other channels so as to obtain updated state information (to benefit later sensing decisions). This is an interesting research topic deserving further significant investigation.

In the average rate maximization problem with underlay mode discussed in Chapter 6, parameter B, the total bandwidth that the cognitive BS is assigned, is assumed to be no larger than one. In real application, it is also possible that B is larger than 1. In this case, a constraint $x_n \leq 1, \forall n \in \mathcal{N}$ should be added to the formulated optimization problems. However, it is not clear whether special properties exist, and if exist, how to derive fast resource allocation algorithms. Those questions deserve further investigation.

Throughout this thesis, the utilities to be maximized are always related to data rates, such as average throughput in Chapter 3 and average rate in Chapter 6. Due to the limited power resources in a wireless communication system, energy efficiency is also an important utility to be maximized. In the future, optimal spectrum sensing and resource allocation aiming at maximizing the energy efficiency in the cognitive radio system will be studied.

References

[1] T. S. Rappaort, *Wireless Communications: Principles and Practice*, Prentice Hall, second edition, 2002.

[2] G. L. Stüber, *Principles of Mobile Communication*, Springer Press, 2011.

[3] A. Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.

[4] International Telecommunication Union, "Cellular standards for the third generation," Dec. 2005.

[5] International Telecommunication Union, "Recommendation M.1645: Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000," June 2003.

[6] J. Vilches, "Everything you need to know about 4G wireless technology," *TechSpot*, April 2010.

[7] A. F. Molisch, *Wireless Communications*, John Wiley and Sons, 2011.

[8] W. Song and W. Zhuang, "Interworking of 3G cellular networks and wireless LANs," *Int. J. Wireless and Mobile Computing*, vol. 2, no. 4, pp. 237-247, 2007.

[9] B. O'Hara, A. Petrick, *IEEE 802.11 Handbook: A Designer's Companion*, IEEE Press, second edition, 2005.

[10] http://www.ntia.doc.gov/page/2011/united-states-frequency-allocation-chart.

[11] Federal Communications Commission, "Spectrum policy task force report," Rep. ET Docket no. 02-155, Nov. 2002.

[12] I. J. Mitola, “Cognitive radio architecture evolution,” *Proc. IEEE*, vol. 97, no. 4, pp. 626-641, April 2009.

[13] S. Haykin, “Cognitive radio: Brain-empowered wireless communications,” *IEEE J. Select. Areas Commun.*, vol. 23, no. 2, pp. 201-220, Feb. 2005.

[14] I. J. Mitola, the MITRE Corporation and the Royal Insitiue of Technology, “Cognitive radio for flexible mobile multimedia communications,” *Proc. 1999 IEEE Int. Workshop on Mobile Multimedia Commun. (MoMuC'99)*, San Diego, CA, USA, Nov. 1999.

[15] Z. Qing and B. M. Sadler, “A survey of dynamic spectrum access," *IEEE Signal Processing Mag.*, vol. 24, no. 3, pp. 79-89, May 2007.

[16] Federal Communications Commission, “Spectrum policy task force,” Rep. ET Docket no. 02-135, Nov. 2002.

[17] http://ieee802.org/22/.

[18] http://www.ecma-international.org/publications/standards/Ecma-392.htm.

[19] http://www.dyspan-sc.org/.

[20] http://www.radio-electronics.com/info/wireless/wi-fi/ieee-802-11af-white-fi-tv-space.php.

[21] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran and S. Mohanty, “NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey,” *Computer Networks*, vol. 50, no. 13, pp. 2127-2159, Sept. 2006.

[22] Z. Zhang, and H. Jiang, “Cognitive radio with imperfect spectrum sensing: The optimal set of channels to sense,” *IEEE Wireless Commun. Letters*, vol. 1, no. 2, pp. 133-136, April 2012.

[23] Z. Zhang and H. Jiang, “Channel exploration and exploitation with imperfect spectrum sensing in cognitive radio networks,” [Online]. Available: http://arxiv.org/pdf/1201.4214.pdf.

[24] A. Rao and M. Alouini, "Performance of cooperative spectrum sensing over non-identical fading environments," *IEEE Trans. Wireless Commun.*, vol. 59, no. 12, pp. 3249-3253, Dec. 2011.

[25] A. S. Cacciapuoti, M. Caleffi, D. Izzo, and L. Paura, "Cooperative spectrum sensing techniques with temporal dispersive reporting channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3392-3402, Oct. 2011.

[26] C. Chen, H. Cheng and Y. Yao, "Cooperative spectrum sensing in cognitive radio networks in the presence of the primary user emulation attack," *IEEE Trans. Wireless Commun.*, vol. 10, no. 7, pp. 2135-2141, July 2011.

[27] J. Kim, C. Chae and J. G. Andrews, "Cooperative spectral covariance sensing under correlated shadowing," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3589-3593, Nov. 2011.

[28] K. W. Choi, E. Hossain and D. I. Kim, "Cooperative spectrum sensing under a random geometric primary user network model," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1932-1944, June 2011.

[29] E. C. Y. Peh, Y.-C. Liang, Y. L. Guan, and Y. Zeng, "Power control in cognitive radios under cooperative and non-cooperative spectrum sensing," *IEEE Trans. Wireless Commun.*, vol. 10, no. 12, pp. 4238-4248, Dec. 2011.

[30] Y. Zou, Y.-D. Yao and B. Zheng, "A selective-relay based cooperative spectrum sensing scheme without dedicated reporting channels in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1188-1198, April 2011.

[31] Z. Zhang, Z. Han, H. Li, D. Yang, and C. Pei, "Belief propagation based cooperative compressed spectrum sensing in wideband cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3020-3031, Sept. 2011.

[32] Z. Han, R. Fan, and H. Jiang, "Replacement of spectrum sensing in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2819-2826, June 2009.

[33] V. K. Tumuluru, P. Wang, and D. Niyato, "A novel spectrum scheduling scheme for multi-channel cognitive radio network and performance analysis," *IEEE Trans. Veh. Technol.*, vol. 60, no. 4, pp. 1849-1858, May 2011.

[34] D. T. Ngo and T. Le-Ngoc, "Distributed resource allocation for cognitive radio networks with spectrum-sharing constraints," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3436-3449, Sept. 2011.

[35] Q. Chen, Y.-C. Liang, M. Motani, and W.-C. Wong, "A two-level MAC protocol strategy for opportunistic spectrum access in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 5, pp. 2164-2180, June 2011.

[36] Y.-C. Liang, K.-C. Chen, G. Y. Li, and P. Mahonen, "Cognitive radio networking and communications: An overview," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3386-3407, Sept. 2011.

[37] A. A. El-Sherif and K. J. R. Liu, "Joint design of spectrum sensing and channel access in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1743-1753, June 2011.

[38] S. Gunawardena and W. Zhuang, "Capacity analysis and call admission control in distributed cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3110-3120, Sept. 2011.

[39] S. Stotas and A. Nallanathan, "Optimal sensing time and power allocation in multiband cognitive radio networks," *IEEE Trans. Commun.*, vol. 59, no. 1, pp. 226-235, Jan. 2011.

[40] M. H. Cheung, V. W. S. Wong and R. Schober, "SINR-based random access for cognitive radio: Distributed algorithm and coalitional game," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3887-3897, Nov. 2011.

[41] S. Stotas and A. Nallanathan, "On the throughput and spectrum sensing enhancement of opportunistic spectrum access cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 97-107, Jan 2012.

[42] E. Hossain and V. Bhargava, *Cognitive Wireless Communication Networks*, Springer Press, 2007.

[43] N. B. Chang and M. Liu, "Optimal channel probing and transmission scheduling for opportunistic spectrum access," *Proc. 13th ACM Annual Int. Conf. Mobile Computing and Networking (MobiCom'07)*, pp. 27-38, Montreal, QC, Canada, Sept. 2007.

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

[45] P. D. Sutton, K. E. Nolan and L. E. Doyle, "Cyclostationary signature in practical cognitive radio applications," *IEEE J. Select. Areas Commun.*, vol. 26, no. 1, pp. 13-24, Jan. 2008.

[46] K. B. Letaief and W. Zhang, "Cooperative communications for cognitive radio networks," *Proc. IEEE*, vol. 97, no. 5, pp. 878-893, May. 2009.

[47] Y. Hur, J. Park, W. Woo, J. S. Lee, K. Lim, C.-H. Lee, H. S. Kim, and J. Laskar, "A cognitive radio (CR) system employing a dual-stage spectrum sensing technique: A multi-resolution spectrum sensing (MRSS) and a temporal signature detection (TSD) technique," in *Proc. IEEE Global Telecommun. Conf.*, pp. 4090-4093, San Francisco, CA, Nov.-Dec., 2006.

[48] Y. Zeng and Y.-C. Liang, "Covariance based signal detections for cognitive radio," in *Proc. IEEE Int. Symp. New Frontiers Dyn. Spectrum Access Netw. (DySPAN 2007)*, pp. 202-207, Dublin, Ireland, Apr. 2007.

[49] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326-1337, Apr. 2008.

[50] S. Atapattu, C. Tellambura, and H. Jiang, "Energy detection based cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1232-1241, Apr. 2011.

[51] W.-Y. Lee and I. F. Akyildiz, "Optimal spectrum sensing framework for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3845-3857, Oct. 2008.

[52] T. Ferguson, Optimal Stopping and Applications 2006 [Online]. Available: http://www.math.ucla.edu/~tom/Stopping/Contents.html.

[53] D. Zheng, W. Ge, and J. Zhang, "Distributed opportunistic scheduling for ad-hoc networks with random access: An optimal stopping approach," *IEEE Trans. Inform. Theory*, vol. 55, no. 1, pp. 205-222, Jan. 2009.

[54] J. Jia, Q. Zhang, and X. Shen, "HC-MAC: A hardware-constrained cognitive MAC for efficient spectrum management," *IEEE J. Select. Areas Commun.*, vol. 26, no. 1, pp. 106-117, Jan. 2008.

[55] T. Shu and M. Krunz, "Throughput-efficient sequential channel sensing and probing in cognitive radio networks under sensing errors," *Proc. 13th ACM Annual Int. Conf. Mobile Computing and Networking (MobiCom'09)*, pp. 37-48, Beijing, China, Sept. 2009.

[56] H. Jiang, L. Lai, R. Fan, and H. V. Poor, "Optimal selection of channel sensing order in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 297-307, Jan. 2009.

[57] J. C. F. Li, S. Dey and J. Evans, "Maximal lifetime power and rate allocation for wireless sensor systems with data distortion constraint," *IEEE Trans. Signal Processing*, vol. 56, no. 5, pp. 2076-2090, May 2008.

[58] L. Hoo, B. Halder, J. Tellado, and J. Cioffi, "Multiuser transmit optimization for multicarrier broadcast channels: Asymptotic FDMA capacity region and algorithms," *IEEE Trans. Commun.*, vol. 52, no. 6, pp. 922-930, June 2004.

[59] J. Jang and K. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Select. Areas Commun.*, vol. 21, no. 2, pp. 171-178, Feb. 2003.

[60] D. Kivanc, G. Li, and H. Liu, "Computationally efficient bandwidth allocation and power control for OFDMA," *IEEE Trans. Wireless Commun.*, vol. 2, no. 6, pp. 1150-1158, Nov. 2003.

[61] C. Wong, R. Cheng, K. Lataief, and R. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Select. Areas Commun.*, vol. 17, no. 10, pp. 1747-1758, Oct. 1999.

[62] X. Wang, G. B. Giannakis, "Resource allocation for wireless multiuser OFDM networks," *IEEE Trans. Inform. Theory*, vol. 57, no. 7, pp. 4359-4372, July 2011.

[63] H. Kushner and P. Whiting, "Convergence of proportional-fair sharing algorithms under general conditions," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1250-1259, Jul. 2004.

[64] X. Wang, G. Giannakis, and A. Marques, "A unified approach to QoS-guaranteed scheduling for channel-adaptive wireless networks," *Proc. IEEE*, vol. 95, no. 12, pp. 2410-2431, Dec. 2007.

[65] Y. Xing, C. N. Mathur, M. A. Haleem etc, "Dynamic spectrum access with QoS and interference temperature constraints," *IEEE Trans. Mobile Computing*, vol. 6, no. 4, pp. 423-433, April. 2007.

[66] J. Huang, R. Berry, and M.L. Honig, "Auction-based spectrum sharing," *ACM/Springer Mobile Networks and Applications J. (MONET)*, vol. 11, no. 3, pp. 405-418, June 2006.

[67] R. Zhang, S. Cui, and Y.-C. Liang, "On ergodic sum capacity of fading cognitive multiple-access and broadcast channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 11, pp. 5161-5178, Nov. 2009.

[68] M. Gastpar, "On capacity under receive and spatial spectrum-sharing constraints," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 471-487, Feb. 2007.

[69] G. Caire, G. Taricco, and E. Biglieri, "Optimal power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468-1489, Jul. 1999.

[70] L. B. Le and E. Hossain, "Resource allocation for spectrum underlay in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5306-5315, Dec. 2008.

[71] X. Kang, H. K. Garg, Y.-C. Liang and R. Zhang, "Optimal power allocation for OFDM-based cognitive radio with new primary transmission protection criteria", *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, pp. 2066-2075, Dec. 2008.

[72] Y. T. Hou, Y. Shi, and H. D. Sherali, "Spectrum sharing for multi-hop networking with cognitive radios," *IEEE J. Select. Areas Commun.*, vol. 26, no. 1, pp. 146-155, Jan. 2008.

[73] S. M. Almalfouh and G. L. Stüber, "Interference-aware radio resource allocation in OFDMA-based cognitive radio network," *IEEE Trans. Veh. Technol.*, vol. 60, no. 4, pp. 1699-1713, Feb. 2011.

[74] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE J. Select. Areas Commun.*, vol. 25, no. 3, pp. 589-600, Apr. 2007.

[75] L. Lai, H. El Gamal, H. Jiang, and H. V. Poor, "Cognitive medium access: Exploration, exploitation and competition," *IEEE Trans. Mobile Computing*, vol. 10, no. 2, pp. 239-253, Feb. 2011.

[76] X. Kang, Y.-C. Liang, H. K. Garg, and L. Zhang, "Sensing-based spectrum sharing in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4649-4654, Oct. 2009.

[77] M. R. Bussieck and A. Pruessner, "Mixed-Integer Nonlinear Programming," [Online]. Available: http://www.gamsworld.org/minlp/siagopt.pdf.

[78] C. Floudas and P. M. Pardalos, *Encyclopedia of Optimization.* Springer Press, second edition, 2009.

[79] T. Ibaraki and N. Katoh, *Resource Allocation Problems: Algorithmic Approaches.* The MIT Press, 1988.

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

[81] S. Guha, K. Munagala, and S. Sarkar, "Approximation schemes for information acquisition and exploitation in multichannel wireless networks," in *Proc. 44th Annual Allerton Conf. Commun., Control and Computing*, Allerton, Monticello, IL, USA, Sept. 2006.

[82] H. Kim and K. G. Shin, "Fast discovery of spectrum opportunities in cognitive radio networks," in *Proc. IEEE Symp. New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Chicago, IL, USA, Oct. 2008.

[83] L. X. Cai, X. Shen, J. W. Mark, L. Cai, and Y. Xiao, "Voice capacity analysis of WLAN with unbalanced traffic," *IEEE Trans. Veh. Technol.*, vol. 55, no. 3, pp. 752-761, May 2006.

[84] P. Wang, H. Jiang, and W. Zhuang, "A new MAC scheme supporting voice/data traffic in wireless ad hoc networks," *IEEE Trans. Mobile Computing*, vol. 7, no. 12, pp. 1491-1503, Dec. 2008.

[85] H. Jiang, P. Wang, W. Zhuang, and H. V. Poor, "Service time analysis of a distributed medium access control scheme," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3988-3998, Oct. 2008.

[86] H. Jiang, P. Wang, and W. Zhuang, "A distributed channel access scheme with guaranteed priority and enhanced fairness," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2114-2125, June 2007.

[87] H. Luo, J. Cheng, and S. Lu, "Self-coordinating localized fair queueing in wireless ad hoc networks," *IEEE Trans. Mobile Computing*, vol. 3, no. 1, pp.86-98, Jan.-Mar. 2004.

[88] H. Jiang, W. Zhuang, X. Shen, and Q. Bi, "Quality-of-service provisioning and efficient resource utilization in CDMA cellular communications," *IEEE J. Select. Areas Commun.*, vol. 24, no. 1, pp. 4-15, Jan. 2006.

[89] V. Kanodia, A. Sabharwal, and E. Knightly, "MOAR: A multi-channel opportunistic auto-rate media access protocol for ad hoc networks," in *Proc. First Annual Int'l Conf. Broadband Networks (BROADNETS)*, San Jose, CA, USA, Oct. 2004.

[90] A. Sabharwal, A. Khoshnevis, and E. Knightly, "Opportunistic spectral usage: Bounds and a multi-band CSMA/CA protocol," *IEEE/ACM Trans. Networking*, vol. 15, no. 3, pp. 533-545, June 2007.

[91] L. Lai and H. El Gamal, "The water-filling game in fading multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 5, pp. 2110-2122, May 2008.

[92] L. Li and A. J. Goldsmish, "Capacity and optimal resource allocation for fading broadcaset channels-Part I: Ergodic capacity," *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 1083-1102, Mar. 2001.

[93] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[94] N. Jindal and A. Lozana, "Fading models and metrics for contemporary wireless systems," in *Proc. 44th Asilomar Conf. on Signals, Systems, and Computers*, pp. 625-629, Nov. 2010.

[95] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization.* Athena Scientific, 2003.

[96] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer: New York, 1999.

[97] J. Acharya and R. D. Yates, "Dynamic spectrum allocation for uplink users with heterogeneous utilities," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1405-1413, Mar. 2009.

[98] J.-W. Lee, R. R. Mazumdar, and N. B. Shroff, "Opportunistic power scheduling for dynamic multi-server wireless systems," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1506-1515, June 2006.

[99] Y. Song, C. Zhang, and Y. Fang, "Harnessing traffic uncertainties in wireless mesh networks - a stochastic optimization approach," *ACM Mobile Networks and Applications (MONET)*, vol. 14, no. 2, pp. 124-133, Feb. 2009.

[100] X. Wang and N. Gao, "Stochastic resource allocation over fading multiple access and broadcast channels," *IEEE Trans. Inform. Theory*, vol. 56, no. 5, pp. 2382-2391, May 2010.

Appendix A

Proof of Lemma 3.4

Proof. In algorithm 3.1, Step 2 and Step 3 are repeated by $\left(kM - \sum_{n=1}^{N} z_n\right)$ times, referred to as $\left(kM - \sum_{n=1}^{N} z_n\right)$ rounds. After round i is completed, we denote the value of q_n as $q_n^{(i)}$, $n \in \mathcal{N}$. Then algorithm 3.1 can be represented equivalently in an alternative way:

Algorithm A.1 Equivalence to algorithm 3.1.

1: If $kM - \sum_{n=1}^{N} z_n < 0$, problem 3.3 is infeasible for the given k, return. Otherwise initialize $q_n^{(0)} = 0, n \in \mathcal{N}$ and set round index $i = 0$.
2: **while** $i < kM - \sum_{n=1}^{N} z_n$ **do**
3:     Set $i \leftarrow i + 1$
4:     **for** $n \in \mathcal{N}$ **do**
5:         If $n = \arg\max_{1 \le n \le N} \left[s_n \left(q_n^{(i-1)} + 1 \right) - s_n \left(q_n^{(i-1)} \right) \right]$, then $q_n^{(i)} = q_n^{(i-1)} + 1$; otherwise, $q_n^{(i)} = q_n^{(i-1)}$.
6: Output $\left\{ q_n^{(kM - \sum_{n=1}^{N} z_n)}, n \in \mathcal{N} \right\}$.

We denote the channel selected in round i as $c(i)$:

$$c(i) = \arg\max_{1 \le n \le N} \left[s_n \left(q_n^{(i-1)} + 1 \right) - s_n \left(q_n^{(i-1)} \right) \right].$$

If we record the value of $U(\{k_n\},\{\varepsilon_n\})|_{P_n^d(k_n,\varepsilon_n)=P_{th}}$ as

$$U(\{k_n\},\{\varepsilon_n\})|_{P_n^d(k_n,\varepsilon_n)=P_{th}} \overset{q_n=k_n-z_n}{=} \begin{array}{l}\sum\limits_{n=1}^{N}\Big(\Pr(\mathcal{H}_n^0)\big(1-P_n^f(q_n,P_{th})\big)R_n^0 \\ +\Pr(\mathcal{H}_n^1)(1-P_{th})R_n^1\Big)\end{array} \tag{A-1}$$

after each round, totally we have $\left(kM-\sum\limits_{n=1}^{N}z_n\right)$ values, denoted $H(1), H(2), ...,$ $H\left(kM-\sum\limits_{n=1}^{N}z_n\right)$, respectively. Here $P_n^f(q_n,P_{th})$ is given in (3.26).

Recall that algorithm 3.1 leads to an optimal solution of problem 3.2. So the optimal objective value of problem 3.2 (with a specific k), i.e., $U^*(k)$, is given by

$$U^*(k)=H\left(kM-\sum_{n=1}^{N}z_n\right). \tag{A-2}$$

In (A-1), the second term in the summation on the right-hand side of the equality is a fixed value. So we have

$$\begin{aligned}&H(i)-H(i-1)\\ =\;&\Pr\left(\mathcal{H}_{c(i)}^0\right)\left(1-P_{c(i)}^f\left(q_{c(i)}^{(i)},P_{th}\right)\right)R_{c(i)}^0\\ &-\Pr\left(\mathcal{H}_{c(i)}^0\right)\left(1-P_{c(i)}^f\left(q_{c(i)}^{(i-1)},P_{th}\right)\right)R_{c(i)}^0\\ =\;&s_{c(i)}\left(q_{c(i)}^{(i)}\right)-s_{c(i)}\left(q_{c(i)}^{(i)}-1\right).\end{aligned} \tag{A-3}$$

We first prove the property (1): $U^*(k)$ is an increasing function.

Similar to (3.27) and (3.28), it can be proved that $s_n(q_n)$ is an increasing and concave function with respect to q_n. So we have

$$s_{c(i)}\left(q_{c(i)}^{(i)}\right)-s_{c(i)}\left(q_{c(i)}^{(i)}-1\right)>0$$

and thus

$$H(i)-H(i-1)>0,\quad i>1.$$

This means $H(i)$ is an increasing function. Further, based on (A-2), we have

$$U^*(k)-U^*(k-1)=H\left(kM-\sum_{n=1}^{N}z_n\right)-H\left((k-1)M-\sum_{n=1}^{N}z_n\right)>0.$$

This means that $U^*(k)$ is an increasing function.

Now we proceed to prove property (2) of $U^*(k)$. We first compare $[H(i) - H(i-1)]$ with $[H(i+1) - H(i)]$. The difference is given by

$$\begin{aligned}
&[H(i) - H(i-1)] - [H(i+1) - H(i)] \\
&= \left[s_{c(i)}\left(q_{c(i)}^{(i)}\right) - s_{c(i)}\left(q_{c(i)}^{(i)} - 1\right)\right] - \left[s_{c(i+1)}\left(q_{c(i+1)}^{(i+1)}\right) - s_{c(i+1)}\left(q_{c(i+1)}^{(i+1)} - 1\right)\right].
\end{aligned} \tag{A-4}$$

We have two scenarios as follows.

- If $c(i) = c(i+1) = j$: We have $q_j^{(i+1)} = q_j^{(i)} + 1$. From (A-4), we have

$$\begin{aligned}
&[H(i) - H(i-1)] - [H(i+1) - H(i)] \\
&= \left[s_j\left(q_j^{(i)}\right) - s_j\left(q_j^{(i)} - 1\right)\right] - \left[s_j\left(q_j^{(i)} + 1\right) - s_j\left(q_j^{(i)}\right)\right] \\
&\overset{(a)}{>} 0
\end{aligned}$$

  where (a) follows from the fact that $s_n(q_n)$ is a concave function with respect to q_n.

- If $c(i) \neq c(i+1)$: In round i, $c(i) = \underset{1 \leq n \leq N}{\arg\max}\left[s_n\left(q_n^{(i-1)} + 1\right) - s_n\left(q_n^{(i-1)}\right)\right]$. This means that

$$s_{c(i)}\left(q_{c(i)}^{(i-1)} + 1\right) - s_{c(i)}\left(q_{c(i)}^{(i-1)}\right) \geq s_{c(i+1)}\left(q_{c(i+1)}^{(i-1)} + 1\right) - s_{c(i+1)}\left(q_{c(i+1)}^{(i-1)}\right). \tag{A-5}$$

  Since $q_{c(i)}^{(i)} = q_{c(i)}^{(i-1)} + 1$, we have

$$s_{c(i)}\left(q_{c(i)}^{(i-1)} + 1\right) - s_{c(i)}\left(q_{c(i)}^{(i-1)}\right) = s_{c(i)}\left(q_{c(i)}^{(i)}\right) - s_{c(i)}\left(q_{c(i)}^{(i)} - 1\right). \tag{A-6}$$

  Since $q_{c(i+1)}^{(i+1)} = q_{c(i+1)}^{(i)} + 1 = q_{c(i+1)}^{(i-1)} + 1$, we have

$$s_{c(i+1)}\left(q_{c(i+1)}^{(i-1)} + 1\right) - s_{c(i+1)}\left(q_{c(i+1)}^{(i-1)}\right) = s_{c(i+1)}\left(q_{c(i+1)}^{(i+1)}\right) - s_{c(i+1)}\left(q_{c(i+1)}^{(i+1)} - 1\right). \tag{A-7}$$

  From (A-4) - (A-7), it can be seen that

$$[H(i) - H(i-1)] - [H(i+1) - H(i)] \geq 0.$$

Therefore, in either of the above scenarios, we always have

$$[H(i) - H(i-1)] - [H(i+1) - H(i)] \geq 0, \quad i > 1$$

and it follows that

$$H(i) - H(i-1) \geq H(i+M) - H(i+M-1), \quad i > 1. \tag{A-8}$$

Then, we have

$$\begin{aligned}
&U^*(k) - U^*(k-1) \\
&= H\left(kM - \sum_{n=1}^{N} z_n\right) - H\left((k-1)M - \sum_{n=1}^{N} z_n\right) \\
&= \sum_{i=(k-1)M - \sum_{n=1}^{N} z_n + 1}^{kM - \sum_{n=1}^{N} z_n} \Big(H(i) - H(i-1)\Big) \\
&\overset{\text{from } (A-8)}{\geq} \sum_{i=kM - \sum_{n=1}^{N} z_n + 1}^{(k+1)M - \sum_{n=1}^{N} z_n} \Big(H(i) - H(i-1)\Big) \\
&= U^*(k+1) - U^*(k).
\end{aligned} \tag{A-9}$$

This completes the proof.

□

Appendix B

Primary-Free Probability Update Procedure for Channel b_l $(1 \leq l \leq k-1)$ when User 1 is Selecting a_k

We have four possible situations as follows.

- If $b_l \notin \mathcal{A}$: then $\theta^*_{b_l} = \theta_{b_l}$;
- If $b_l \in \mathcal{A}$, and $g(b_l, \boldsymbol{A}) < l$: Since user 1 should sense channel b_l prior to user 2, and user 1 has not stopped at channel b_l, we have $\theta^*_{b_l} = 0$;
- If $b_l \in \mathcal{A}$, and $g(b_l, \boldsymbol{A}) = l$ (i.e., $a_l = b_l$): Channel b_l has the same sensing position (i.e., the l^{th} position) with both users.
  - If either `FAIL_THEN_QUIT` or `COLLIDE` strategy is adopted as the contention resolution strategy, apparently there should be primary activity in channel b_l, because otherwise user 1 should have stopped at its l^{th} sensing. So $\theta^*_{b_l} = 0$.
  - If `FAIL_THEN_CONTINUE` is adopted as the contention resolution strategy, since user 1 has not stopped at channel b_l, either of the following two events should happen: i) there are primary activities in channel b_l; ii) there is no primary activity in channel b_l, user 1 and 2 both proceed to sense channel b_l in their l^{th}

sensing, and user 2 wins. Define y as the scenario that event i) happens, and x as the scenario that either event i) or event ii) happens. The updated primary-free probability $\theta^*_{b_l}$ is given by $\theta^*_{b_l} = 1 - \Pr(y|x)$, which means $\Pr(y|x) = 1 - \theta^*_{b_l}$. We also have $\Pr(y) = 1 - \theta_{b_l}$, $\Pr(x|y) = 1$, and

$$\Pr(x) = 1 - \theta_{b_l} + \theta_{b_l} \left(\Pi_{j=1}^{l-1} (1 - \theta^*_{b_j}) \right) \cdot 0.5$$

where the term $\Pi_{j=1}^{l-1}(1 - \theta^*_{b_j})$ means the probability that user 2 does not stop at its first $l-1$ sensings, and the term 0.5 means user 2 has a probability 0.5 to win in the contention with user 1. Based on Bayesian formula $\Pr(y|x) = \frac{\Pr(y)\Pr(x|y)}{\Pr(x)}$, we can get

$$\theta^*_{b_l} = \theta_{b_l} \frac{0.5\Pi_{j=1}^{l-1}(1 - \theta^*_{b_j})}{1 - \theta_{b_l} + 0.5\theta_{b_l}\Pi_{j=1}^{l-1}(1 - \theta^*_{b_j})}. \quad \text{(B-1)}$$

- If $b_l \in \mathcal{A}$, and $g(b_l, \boldsymbol{A}) > l$: User 2 should sense channel b_l prior to user 1. Since user 1 has not stopped at channel b_l, either of the following two events should happen: i) there are primary activities in channel b_l; ii) there is no primary activity in channel b_l, and user 2 proceeds to sense channel b_l and transmits in channel b_l. Similarly we can get

$$\theta^*_{b_l} = \theta_{b_l} \frac{\Pi_{j=1}^{l-1}(1 - \theta^*_{b_j})}{1 - \theta_{b_l} + \theta_{b_l}\Pi_{j=1}^{l-1}(1 - \theta^*_{b_j})}. \quad \text{(B-2)}$$

Appendix C

Convexity Proof of Problem 6.1

To prove convexity of problem 6.1, we need to prove that the objective function is concave with respect to $\{p_n\}$ and $\{x_n\}$. Since the objection function in problem 6.1 is separable with link index n, it is equivalent to prove that function $x \ln\left(1 + \frac{ph}{x}\right)$ is concave with respect to p and x. We investigate the Hessian matrix of function $x \ln\left(1 + \frac{ph}{x}\right)$, $\mathbf{H}$, which can be obtained as

$$\mathbf{H} = \frac{h^2}{x^3(1 + \frac{hp}{x})^2} \begin{bmatrix} -p^2 & px \\ px & -x^2 \end{bmatrix}$$

with eigenvalues values $-\frac{h^2(x^2+p^2)}{x^3(1+\frac{hp}{x})^2}$ and 0. With all the eigenvalues not larger than 0, it can be seen that matrix $\mathbf{H}$ is a semi-negative definite matrix. Therefore, the function $x \ln\left(1 + \frac{ph}{x}\right)$ is concave. This completes the proof.

Appendix D

Proof of Lemma 6.2

A similar proof method to that in Lemma 6.1 is used.

We use proof by contradiction. When the constraint in (6.25a) is inactive, suppose $|\mathcal{A}| > 1$, and $i^\dagger, i^\ddagger \in \mathcal{A}$. Similar to (6.22), we have

$$\frac{\Gamma^*}{\frac{\lambda^*}{g_{i^\dagger}} + \frac{\mu_{i^\dagger} h_{i^\dagger}}{g_{i^\dagger}}} = \frac{\Gamma^*}{\frac{\lambda^*}{g_{i^\ddagger}} + \frac{\mu_{i^\ddagger} h_{i^\ddagger}}{g_{i^\ddagger}}} \tag{D-1}$$

and $\Gamma^* > 0$.

When the constraint in (6.25a) is inactive, $\lambda^* = 0$ according to (6.26c). Note that $\mu_{i^\dagger}$ and $\mu_{i^\ddagger}$ are fixed for the problem in (6.26). In this case, equation (D-1) holds with probability zero since $h_{i^\dagger}$ and $h_{i^\ddagger}$ are independent random variables.

When the constraint in (6.25a) is active, $\lambda^* \geq 0$. Suppose $|\mathcal{A}| > 2$, and $i^\dagger, i^\ddagger, i' \in \mathcal{A}$. We have

$$\frac{\Gamma^*}{\frac{\lambda^*}{g_{i^\dagger}} + \frac{\mu_{i^\dagger} h_{i^\dagger}}{g_{i^\dagger}}} = \frac{\Gamma^*}{\frac{\lambda^*}{g_{i^\ddagger}} + \frac{\mu_{i^\ddagger} h_{i^\ddagger}}{g_{i^\ddagger}}} = \frac{\Gamma^*}{\frac{\lambda^*}{g_{i'}} + \frac{\mu_{i'} h_{i'}}{g_{i'}}}. \tag{D-2}$$

Note that $\mu_{i^\dagger}$, $\mu_{i^\ddagger}$, and $\mu_{i'}$ are fixed for the problem in (6.26). It can be seen that equation (D-2) holds with probability zero since $g_{i^\dagger}$, $g_{i^\ddagger}$, $g_{i'}$, $h_{i^\dagger}$, $h_{i^\ddagger}$, and $h_{i'}$ are independent random variables.

This completes the proof.

Appendix E

Proof of Lemma 6.3

According to (6.26d), for $\forall n \in \mathcal{A}$, we have $\delta_n^* = 0$. Since $\mathrm{SNR}_n = \frac{p_n g_n}{x_n}$, equation (6.26a) can be written as

$$\ln\left(1 + \mathrm{SNR}_n\right) - \frac{\mathrm{SNR}_n}{1 + \mathrm{SNR}_n} = \Gamma^*. \tag{E-1}$$

The left-hand side of (E-1) is in the form of $S(x)$ (defined in Chapter 6.3.1), which is a monotonic increasing function. If $i, j \in \mathcal{A}$, then both i and j satisfy (E-1), which means $\mathrm{SNR}_i = \mathrm{SNR}_j$.

Appendix F

Proof of Lemma 6.6

For $\forall i \in \mathcal{A}_1$ and $\forall j \in \mathcal{A}_2$, it can be seen that $\Delta_i^* = 0$ and $\Delta_j^* = 0$ according to equation (6.30d) and $\mu_j^* = 0$ according to (6.30c). Therefore, for i and j, equation (6.30b) can be rewritten as

$$\frac{g_i}{1 + \text{SNR}_i} = \lambda + \mu_i^* h_i \tag{F-1}$$

and

$$\frac{g_j}{1 + \text{SNR}_j} = \lambda \tag{F-2}$$

respectively. Together with the conclusion of Lemma 6.5 and the fact that $\mu_i^* \geq 0$, it is apparent that $g_i \geq g_j$ from equations (F-1) and (F-2).

Appendix G

Proof of Lemma 6.7

We use proof by contradiction. For ease of presentation, assume in the optimal solution of the problem in (6.31) we have $1 \in \mathcal{B}$ and $2 \in \mathcal{A}$ and $g_1 > g_2$. Then $p_1 = 0$, and $p_2 \in (0, \frac{I_2^{ST}}{h_2}]$. In the optimal solution, the total utility achieved on channels 1 and 2 is

$$\ln(1 + p_2 g_2) - \lambda p_2. \tag{G-1}$$

It can be seen that there exist $p_1^\dagger$ and $p_2^\dagger$ such that $0 < p_1^\dagger \leq \frac{I_1^{ST}}{h_1}$, $0 \leq p_2^\dagger < p_2$, and $p_1^\dagger + p_2^\dagger = p_2$. If the transmission power on channels 1 and 2 is $p_1^\dagger$ and $p_2^\dagger$, respectively, the total achieved utility on channels 1 and 2 is

$$\ln\left(1 + p_1^\dagger g_1 + p_2^\dagger g_2\right) - \lambda p_2 \tag{G-2}$$

which is definitely larger than the utility in (G-1) since $g_1 > g_2$.

This completes the proof.

Appendix H

Proof of Lemma 6.8

Before the proof, we have two properties for φ_j^*: If $\varphi_j^* = 0$, then $\varphi_{j+1}^* = 0$; If $0 < \varphi_j^* < \frac{I_{s_j}^{ST}}{h_{s_j}}$, then $\varphi_{j+1}^* = 0$. This is because, according to the three observations of (6.33), we have

$$\begin{aligned}
&\varphi_j^* = 0 \\
&\Leftrightarrow \tfrac{Bg_{s_j}}{\lambda} \le B + \sum_{n=1}^{j-1} \tfrac{I_{s_n}^{ST} g_{s_n}}{h_{s_n}} \\
&\Rightarrow \tfrac{Bg_{s_j}}{\lambda} \le B + \sum_{n=1}^{j} \tfrac{I_{s_n}^{ST} g_{s_n}}{h_{s_n}} \\
&\Rightarrow \tfrac{Bg_{s_{j+1}}}{\lambda} \le B + \sum_{n=1}^{j} \tfrac{I_{s_n}^{ST} g_{s_n}}{h_{s_n}} \\
&\Leftrightarrow \varphi_{j+1}^* = 0.
\end{aligned} \tag{H-1}$$

$$\begin{aligned}
&0 < \varphi_j^* < \tfrac{I_{s_j}^{ST}}{h_{s_j}} \\
&\Rightarrow \tfrac{Bg_{s_j}}{\lambda} < B + \sum_{n=1}^{j} \tfrac{I_{s_n}^{ST} g_{s_n}}{h_{s_n}} \\
&\Rightarrow \tfrac{Bg_{s_{j+1}}}{\lambda} < B + \sum_{n=1}^{j} \tfrac{I_{s_n}^{ST} g_{s_n}}{h_{s_n}} \\
&\Rightarrow \varphi_{j+1}^* = 0.
\end{aligned} \tag{H-2}$$

Then it can be concluded that, for Case $j < j^\dagger$, we have $\varphi^* = \frac{I_{s_j}^{ST}}{h_{s_j}}$. Based on the third observations for (6.33), it can be seen that, for Case $j < j^\dagger$, the maximal utility in Case j is less than the maximal utility in Case $j+1$. Therefore, the maximal utility in Case $j^\dagger$ is larger than the maximal utility in any prior case.

Next we prove that the maximal utility in Case $j^\dagger$ is larger than the maximal utility in Case $j^\dagger + 1$. Here we assume $\varphi_{j^\dagger}^* \in (0, \frac{I_{s_{j^\dagger}}^{ST}}{h_{s_{j^\dagger}}})$, as a similar proof can apply when $\varphi_{j^\dagger}^* = 0$.

From the observations for (6.33) and the fact that $\varphi^*_{j^\dagger} \in (0, \frac{I^{ST}_{s_{j^\dagger}}}{h_{s_{j^\dagger}}})$, it can be seen that $\varphi^*_{j^\dagger}$ is the solution of (6.33) in Case $j^\dagger$. It also means the maximal utility in Case $j^\dagger$ is larger than the utility when $p_{s_n} = \frac{I^{ST}_{s_n}}{h_{s_n}}$ for $n = 1, ..., j^\dagger - 1$, $p_{s_{j^\dagger}} = \frac{I^{ST}_{s_{j^\dagger}}}{h_{s_{j^\dagger}}}$, and $p_{s_{j^\dagger+1}} = p_{s_{j^\dagger+2}} = ... = p_{s_N} = 0$. The latter utility is exactly the utility of the optimal solution in Case $j^\dagger + 1$ (which is because, $\varphi^*_{j^\dagger+1} = 0$ from the first property of φ^*_j).

Now we prove that the maximal utility in Case $j^\dagger + 1$ is larger than the maximal utility in Case $j^\dagger + 2$. Note that the optimal solution in Case $j^\dagger + 1$ is $\varphi^*_{j^\dagger+1} = 0$. This means the solution for (6.33) in Case $j^\dagger + 1$ is a negative value. In the optimal solution in Case $j^\dagger + 1$ (i.e., $p_{s_{j^\dagger+1}} = \varphi^*_{j^\dagger+1} = 0$), if the value of $p_{s_{j^\dagger+1}}$ increases from $\varphi^*_{j^\dagger+1} = 0$ to $\frac{I^{ST}_{s_{j^\dagger+1}}}{h_{s_{j^\dagger+1}}}$, the utility function decreases. On the other hand, in the optimal solution in Case $j^\dagger + 1$, when the value of $p_{s_{j^\dagger+1}}$ is changed to $\frac{I^{ST}_{s_{j^\dagger+1}}}{h_{s_{j^\dagger+1}}}$, it leads to the optimal solution in Case $j^\dagger + 2$ (according to the second property of φ^*_j). Therefore, the maximal utility in Case $j^\dagger + 1$ is larger than the maximal utility in Case $j^\dagger + 2$.

By induction, it can be proved that when $j > j^\dagger$, the maximal utility in Case j is larger than the maximal utility in Case $j + 1$. It follows that the maximal utility in Case $j^\dagger$ is larger than the maximal utility in any subsequent case.

Appendix I

Proof of Lemma 6.9

We first prove $|\mathcal{A}_2| = 0$ when the constraint in (6.34b) is inactive. We use proof by contradiction. Assume $|\mathcal{A}_2| \geq 1$, i.e., there exists a channel, denoted channel n, such that $p_n < \frac{I_n^{ST}}{h_n}$. It can be seen that the utility function of the sub-problem in (6.34) can be further improved if we increase p_n by the smaller value between $\left(P^{ST} - \sum_{n=1}^{N} p_n\right) (> 0)$ and $\left(\frac{I_n^{ST}}{h_n} - p_n\right) (> 0)$, since the utility function of the sub-problem in (6.34) is an increasing function with respect to p_n. Therefore, when constraint in (6.34b) is inactive, $\mathcal{A}_2 = \emptyset$.

Next we prove $\lambda^* > 0$ when the constraint in (6.34b) is active. Here we first show that there exists a channel, denoted channel n, such that $0 < p_n < \frac{I_n^{ST}}{h_n}$. We use proof by contradiction. Assume for any channel, say channel j, we have either $p_j = 0$ or $p_j = \frac{I_j^{ST}}{h_j}$. Denote the number of channels with zero power assignment as K. Without loss of generality, assume channels 1, ..., K are with zero power assignment. Therefore, we have

$$\begin{aligned} & p_j h_j = I^{ST}, j = K+1, K+2, ..., N \\ & \sum_{j=K+1}^{N} p_j = P^{ST}. \end{aligned} \tag{I-1}$$

The probability for (I-1) to hold is zero, since h_j's are independent random variables. Therefore, there exists a channel, denoted channel n, such that $0 < p_n < \frac{Q_n^{ST} h_n}{g_n}$. Together with (6.35d) and (6.35e), we have $\mu_n^* = 0$ and $\Delta_n^* = 0$.

Assume $\lambda^* = 0$. Since $\mu_n^* = 0$ and $\Delta_n^* = 0$, we have $x_n = 0$ according to (6.35b). This contradicts the fact that channel n is with positive power and bandwidth assignment. Therefore, we have $\lambda^* > 0$.

Then, we prove $|\mathcal{A}_2| \leq 1$ when the constraint in (6.34b) is active. The proof is similar to that of Lemma 6.4, and is omitted here.