OUTAGE CAPACITY AND THROUGHPUT MAXIMIZATION USING THEORETICAL AND LEARNING-BASED APPROACHES

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

ELECTRICAL AND ELECTRONICS ENGINEERING

By

Saad Masrur

July 2023

Outage Capacity and Throughput Maximization Using Theoretical and Learning-Based Approaches
By Saad Masrur
July 2023

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Sinan Gezici (Advisor)

Aykut Koç

Serkan Sarıtaş

Approved for the Graduate School of Engineering and Science:

Orhan Arıkan
Director of the Graduate School

ABSTRACT

OUTAGE CAPACITY AND THROUGHPUT MAXIMIZATION USING THEORETICAL AND LEARNING-BASED APPROACHES

Saad Masrur
M.S. in Electrical and Electronics Engineering
Advisor: Sinan Gezici
July 2023

This thesis explores two research problems in wireless communications: the optimal channel switching and randomization problem in flat-fading Gaussian noise channels, and channel selection and switching approaches based on the upper confidence bound (UCB) bandit algorithm. In the first part of the thesis, the optimal channel switching and randomization problem is formulated and its solution is characterized for flat-fading Gaussian noise channels with the aim of outage capacity maximization under average power and outage probability constraints. For the single user scenario, it is proved that the optimal solution can always be realized by performing one of the following strategies: (1) Transmission over a single channel with no randomization. (2) Channel switching between two channels with no randomization. (3) Randomization between two parameter sets over a single channel. Hence, the solution can easily be obtained by considering only these three strategies. However, for the multiuser scenario, obtaining the optimal solution can have very high computational complexity. Therefore, an algorithm is proposed to calculate an approximately optimal channel switching and randomization solution (with adjustable approximation accuracy) based on the solution of a linearly constrained linear optimization problem. In the second part of the thesis, we consider the case of unknown channel statistics at the transmitter, and propose channel selection and channel switching approaches based on the UCB bandit algorithm for communications between a transmitter and a receiver over a block fading channel. In the absence of channel switching in a block, we propose a UCB bandit algorithm for selecting the best channel among the possible set of channels for maximizing the number of correctly received symbols per unit of time. In the presence of channel switching, we first define a set of virtual channels by considering all possible channel pairs with various power levels and time sharing factors. Then, a UCB bandit algorithm is utilized to determine the best virtual channel; hence, to find the optimal channel switching strategy. Also, a low complexity version of this algorithm is proposed for efficient convergence to the optimal solution when a high number of virtual channels exists. In addition, for comparison purposes, theoretical limits are presented when the channel statistics are available at the transmitter. Simulation results indicate that the proposed UCB bandit algorithms can achieve very close performance to theoretical limits over a sufficiently large number of blocks, and make benefits of channel switching be realized.

Keywords: Channel switching, outage capacity, time-sharing, power allocation, flat-fading channel, block-fading channel, learning, bandit algorithm, throughput.

ÖZET

KURAMSAL VE ÖĞRENME TABANLI YAKLAŞIMLAR KULLANARAK KESİNTİ KAPASİTESİ VE VERİ HIZI MAKSİMİZASYONU

Saad Masrur
Elektrik ve Elektronik Mühendisliği, Yüksek Lisans
Tez Danışmanı: Sinan Gezici
Temmuz 2023

Bu tez, telsiz iletişimdeki iki araştırma problemini incelemektedir: düz sönümlü Gauss gürültü kanallarında optimal kanal değiştirme ve rastgeleleştirme problemi ve üst güven sınırı (ÜGS) haydut algoritmasına dayalı kanal seçimi ve değiştirme yaklaşımları. Tezin ilk bölümünde, ortalama güç ve kesinti olasılığı kısıtlamaları altında, kesinti kapasitesini maksimize etmek amacıyla, düz sönümlemeli Gauss gürültü kanalları için optimal kanal değiştirme ve rastgeleleştirme problemi formüle edilmiş ve çözümü karakterize edilmiştir. Tek kullanıcı senaryosu için aşağıdaki stratejilerden birinin gerçekleştirilmesiyle optimum çözümün her zaman gerçekleştirilebileceği kanıtlanmıştır: (1) Rastgeleleştirme olmadan tek bir kanal üzerinden iletim. (2) Rastgeleleştirme olmadan iki kanal arasında kanal değiştirmesi. (3) Tek bir kanal üzerinden iki parametre seti arasında rastgeleleştirme. Dolayısıyla, yalnızca bu üç strateji dikkate alınarak çözüm kolayca elde edilebilmektedir. Ancak, çok kullanıcılı senaryo için, optimum çözümü elde etmek çok yüksek hesaplama karmaşıklığına sahip olabilir. Bu nedenle, doğrusal olarak kısıtlanmış bir doğrusal optimizasyon probleminin çözümüne dayalı olarak, optimale yakın bir kanal değiştirme ve rastgeleleştirme çözümünü (ayarlanabilir yaklaşık doğrulukla) hesaplamak için bir algoritma önerilmiştir. Tezin ikinci bölümünde, vericideki bilinmeyen kanal istatistiklerini göz önünde bulundurarak, blok sönümlemeli bir kanal üzerinden bir verici ile bir alıcı arasındaki iletişim için ÜGS haydut algoritmasına dayalı kanal seçimi ve kanal değiştirme yaklaşımları önermekteyiz. Bir blokta kanal değiştirmenin olmadığı durumda, birim zaman başına doğru alınan sembol sayısını maksimize etmek için olası kanal dizileri arasından en iyi kanalı seçmek için bir ÜGS haydut algoritması önermekteyiz. Kanal değiştirme varlığında, önce çeşitli güç seviyelerine ve zaman paylaşım faktörlerine sahip tüm olası kanal çiftlerini göz önünde bulundurarak bir dizi sanal kanallar tanımlanmaktadır. Ardından, en iyi sanal kanalı belirlemek, yani en uygun kanal değiştirme stratejisini bulmak, için bir ÜGS haydut algoritması kullanılmaktadır. Ayrıca, çok sayıda sanal kanal olduğunda optimum çözüme verimli yakınsama için bu algoritmanın düşük karmaşıklıklı bir versiyonu önerilmiştir. Ek olarak, karşılaştırma amacıyla, vericide kanal istatistikleri mevcut olduğunda kuramsal sınırlar sunulmaktadır. Benzetim sonuçları, önerilen ÜGS haydut algoritmalarının, yeterince büyük sayıda blok üzerinde, kuramsal sınırlara çok yakın performans elde edebildiğini ve kanal değiştirmenin faydalarının gerçeklenebildiğini göstermektedir.

Anahtar sözcükler: Kanal değiştirme, kesinti kapasitesi, zaman paylaşımı, güç tahsisi, düz sönümlemeli kanal, blok sönümlemeli kanal, öğrenme, haydut algoritması, verimlilik.

Acknowledgement

I would like to first and foremost thank my advisor Prof. Sinan Gezici. His experience, knowledge and motivation for math was a great inspiration for me. His door was always open and he always gently responded to my spontaneous visits. I will miss those visits and conversations about research and life. It was a pleasure and honor for me to work with him.

I am very thankful to Asst. Prof. Aykut Koç and Asst. Prof. Serkan Sarıtaş for agreeing to serve on my dissertation examination committee.

I am immensely grateful to my parents and siblings for their unconditional love, unwavering support, and the countless sacrifices they have made to help me reach this milestone in my academic journey. Their guidance and belief in me have been a constant source of strength and motivation.

I would also like to extend my heartfelt thanks to all my friends and colleagues for their valuable contributions, inspiring discussions, and unwavering encouragement.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Wireless communication has become an integral part of our daily lives, enabling seamless connectivity and data transmission across various devices and networks. With the ever-increasing demand for faster and more reliable communication, optimizing the performance of wireless channels has emerged as a critical research area. This thesis aims to explore and address key challenges in communication optimization, focusing on two distinct aspects: the optimal channel switching and randomization problem, and channel selection and switching approaches using the upper confidence bound (UCB) bandit algorithm.

In Chapter 2, the focus is on optimal channel switching and randomization [1], which revolves around maximizing communication capacity while considering average power and outage probability constraints. In the single user scenario, several strategies are identified, ranging from transmission over a single channel with no randomization to channel switching between two channels or randomization between two parameter sets over a single channel. These strategies provide insights into achieving the optimal solution for single user systems. However, in multiuser scenarios, the computational complexity increases significantly, necessitating the development of approximate solutions that balance computational efficiency with performance.

Chapter 3 of this thesis focuses on channel selection and switching in the presence of unknown channel statistics at the transmitter [2]. Traditional approaches rely on knowledge of channel statistics to make informed decisions. However, in situations where channel statistics are unavailable, the UCB bandit algorithm offers an intelligent solution. By considering various power levels, time sharing factors, and virtual channels, the algorithm identifies the optimal channel switching strategy, maximizing the number of correctly received symbols per unit of time. Additionally, a low complexity version of the UCB bandit algorithm is proposed to facilitate efficient convergence to the optimal solution in scenarios with a large number of virtual channels.

Through the exploration of these two research problems, this thesis contributes to the advancement of communication optimization techniques in wireless systems. The subsequent chapters delve into detailed analyses, methodologies, and experimental results, shedding light on the optimal channel switching and randomization problem as well as the channel selection and switching approaches using the UCB bandit algorithm. By addressing these challenges, this research aims to enhance the performance and reliability of wireless communication systems, ultimately meeting the ever-growing demands of our interconnected world.

1.1 Optimal Channel Switching and Randomization over Flat-Fading Channels for Outage Capacity Maximization

1.1.1 Literature Review

Performance of communication systems can be enhanced via various time-sharing approaches such as channel switching and parameter randomization [3]–[11]. In channel switching, a transmitter and a receiver perform time-sharing among different channels by communicating over only one channel at a given time [3], [4], [5]. In this way, improvements can be achieved in terms of the average probability of error [3], throughput [12], or channel capacity [4]–[6]. For example, as shown in [3], switching between two channels with a certain time-sharing factor can be necessary in some cases for attaining the minimum average probability of error in an average power constrained binary communication system. In addition, for maximizing the average Shannon capacity between a transmitter and a receiver under average and peak power constraints and in the presence of Gaussian channels, an optimal approach is to implement channel switching between at most two different channels [4]. Apart from channel switching, parameter randomization can also enhance performance of communication systems by employing different parameter values for certain fractions of time; i.e., by performing time-sharing among different parameter sets [7]–[11]. For example, power randomization was carried out in [7] for minimizing the outage probability in a flat block-fading Gaussian channel under an average transmit power constraint. In the context of jamming against digital modulation, the authors of [8] showed that the optimal jamming signal distribution has at most two signal levels along any signaling dimension.

In this thesis, we propose the problem of optimal channel switching and randomization for maximizing the outage capacity between a transmitter and a receiver in the presence of average power and outage probability constraints. The channels between the transmitter and the receiver are modeled as flat-fading and additive Gaussian noise channels, and the channel distribution information (CDI) of each channel is assumed to be available at both the transmitter and the receiver. On the other hand, the channel state information (CSI) of the channels is available only at the receiver. By employing the outage capacity [13] as a performance metric, we derive optimal channel switching and randomization strategies. The outage capacity is an important and practical metric, based on which communication systems can be designed to support a certain number of users with target data rates.

1.1.2 Contributions

Chapter 2 of this thesis explores optimal channel switching and randomization strategies to maximize communication capacity, considering average power and outage probability constraints. The main contributions and novelty of that chapter can be summarized as follows:

- The problem of optimal channel switching and randomization according to the outage capacity metric is proposed for the first time in the literature. (In [4], the optimal channel switching problem was studied based on the Shannon capacity metric in the absence of randomization or fading. In [7], optimal power randomization was performed for minimizing the outage probability in the absence of channel switching.)
- It is shown (in Proposition 1) that an optimal solution to minimize the average outage capacity can be implemented as one of the three strategies: (1) Transmission over a single channel with no randomization. (2) Channel switching between two channels with no randomization. (3) Randomization between two parameter sets over a single channel.
- For the first time in the literature, we propose and solve the optimal channel switching and randomization problem over flat-fading channels for multiuser systems (i.e., multiple transmitter and receiver pairs) with the aim of maximizing the total outage capacity of users. Unlike those in [4, 6, 7], an optimization theoretic approach is employed to provide a solution in the multiuser scenario.

1.2 Channel Switching Strategies based on UCB Bandit Algorithm for Throughput Maximization with Unknown Channel Statistics

1.2.1 Literature Review

Spectrum usage is increasing with the technological advancements in the field of wireless communications. Since spectrum is a scarce resource that needs to be managed efficiently, its inefficient utilization poses many problems for potential users. Among various approaches, time sharing and parameter randomization can facilitate more efficient usage of a given portion of the spectrum in terms of metrics such as channel capacity and throughput [3], [4], [14], [1], [15]. For example, for a given set of channels (frequency bands), time sharing among different channels; i.e., channel switching, can lead to various performance improvements [3], [4], [14], [1]. When channel coefficients or statistics are known, such channel switching and parameter randomization approaches can be implemented in optimal manners as in [3], [4], [14], [1], [15]. However, in practice, channel coefficients and their distributions may not be known beforehand, and they evolve in complex manners over time. Therefore, there is a need to learn and track the dynamic nature of channels in order to utilize them efficiently. The problem is further complicated due to the involvement of the exploration versus exploitation trade off [16], as the only way to gather information about channels is to use them. In this manuscript, by considering the channel switching problem as sequential decision making under uncertainty, we employ a multi-armed bandit (MAB) formulation for throughput maximization in presence of unknown channel statistics.

MAB is an efficient model for sequential decision problems [17]–[19], which has been extensively used in the literature for spectrum sensing and channel selection in cognitive radio systems [20]–[23]. In [24], a model with two users and two channels is considered by employing a Markov chain, where it is assumed that the transition probabilities are known. In [25]–[27], multiuser cognitive radio systems are modeled via MAB by assuming no coordination among users. In addition, time-division fair sharing (TDFS) is employed in the algorithm proposed in [25], which requires pre-agreement among the users with the assumption that the number of users is fixed. A sublinear regret is obtained in [26] by using the ε-greedy algorithm along with a collusion avoiding mechanism. In [27], the musical chair algorithm is used but again the regret is sub-linear. In both [26] and [27], the knowledge about the number of users is not required. When the number of users exceeds the number of the channels, the preceding approaches may not be used, and the problem can be addressed as in [28].

Practical scenarios can also be considered under the MAB framework for channel sensing and selection problems in cognitive radio networks, where secondary users aim to utilize the spectrum of primary users when it is available. For example, the problem of spectrum selection is modeled as a certain type of distributed MAB in [29] in order to address the distributed spectrum selection problem in cognitive networks in presence of imperfect sensing of secondary users. In addition, for single primary user and multi-channel problems, the spectrum selection problem is studied in [30]–[32] from various perspectives. Namely, secondary users perform cooperative spectrum sensing and share information about successful transmissions that they have made previously. Due to cooperation, the order of complexity and sensing time increase depending upon the network size. In [33], the spectrum selection problem with no cooperation among secondary users is modeled via a multi-armed multi-player bandit by incorporating a forgetting factor to the formulation to take into account successful transmissions in the past. A multi-player multi-armed bandit algorithm is also considered in [34] based on error-correction coding with collision-dependent reward distributions with the no-sensing model, and it is shown that the choice of coding schemes has a significant effect on the regret performance. In [35], the problem of allocating the radio channels to links in a wireless network is formulated using linear bandit problems, and analyses are performed under both stochastic and adversarial settings, showing that the proposed algorithms outperform the existing ones in terms of the regret upper bound. In [36], the multi-channel access problem is formulated as a non-cooperative game, where the switching ability of the secondary user is curbed by limiting its channel switching distance. Then, a multiagent Q-learning algorithm is proposed for this non-cooperative game, which requires no prior information. In addition, the distributed learning and channel access problem is investigated in [37] for a cognitive network with multiple secondary users, where the channel availability statistics are initially unknown to the secondary users and are estimated via sensing decisions. A number of policies are proposed for distributed learning of channel availability statistics and channel access of multiple secondary users in a cognitive radio network.

In this manuscript, we consider a transmitter and a receiver, each equipped with a single RF chain, that aim to enhance communication performance via channel switching. In particular, we investigate the optimal channel switching problem for maximization of the number of symbols that are correctly received per unit of time (equivalently, throughput) by considering unknown channel statistics at the transmitter. We propose channel selection and channel switching approaches based on the upper confidence bound (UCB) bandit algorithm for communications over block fading channels. We consider two approaches. In Approach 1, we assume that channel switching is performed only between different blocks (i.e., no channel switching in a given block), and propose a UCB bandit algorithm for selecting the best channel among the possible set of channels for maximizing the number of correctly received symbols per unit of time. In Approach 2, channel switching can be performed in blocks, as well. For this case, we first introduce a set of virtual channels by considering all possible channel pairs with various power levels and time sharing factors, and then employ a UCB bandit algorithm to determine the best virtual channel, therefore, to obtain the optimal channel switching strategy. We also propose a low complexity version of this algorithm for efficient convergence to the optimal solution when a high number of virtual channels exists. Furthermore, for comparison purposes, we present theoretical limits when the channel statistics are known at the transmitter. Although bandit algorithms have been extensively used for spectrum allocation and channel selection in the literature, e.g., [22], the proposed problem formulations in this manuscript are different since channel switching strategies (randomization/time sharing of power levels and channel indices) are considered in our work. Also, even though the optimal channel switching strategies are investigated in various work in the literature such as [4, 14, 38], they all assume that the transmitter has either the channel state information (CSI) or the channel distribution information (CDI) for all channels. For example, in [4], the optimal channel switching technique is derived for maximization of average Shannon capacity by considering perfect CSI for all channels. A similar problem is treated in [38] by taking into account channel switching delays, and optimal channel strategies are obtained for maximizing the average capacity under the assumption of perfect CSI. In [14], the average number of correctly received symbols is maximized via optimal channel switching under average power and cost constraints by assuming the availability of either perfect CSI or CDI for all channels. Therefore, to the best of our knowledge, our work is the first one in the literature that investigates the benefits of channel switching when the transmitter does not have any information about channels.

1.2.2 Contributions

In Chapter 3, we delve into the exploration of channel selection and switching in the presence of unknown channel statistics at the transmitter with the help of the UCB bandit algorithm. The main contributions of that chapter can be listed as follows:

- For the first time in the literature, we formulate the channel switching problem over block fading channels under average and peak power constraints and in the absence of channel statistics at the transmitter as a MAB problem, and propose UCB bandit algorithms for its solution (Algorithms 1–3).
- Although the UCB bandit algorithm can directly be applied when channel switching is performed only between blocks (i.e., no channel switching in a given block), we propose generation of *virtual channels* to be able to employ the UCB bandit algorithm in the presence of channel switching in blocks. In this way, the optimal channel switching strategy to maximize the number of correctly received symbols per unit of time can be obtained (Algorithm

2), which can lead to significant improvements over the case of no channel switching in a block.

- To achieve faster and more efficient convergence, a UCB bandit algorithm with channel elimination (UCB-CE) is proposed (Algorithm 3), which is useful when the number of virtual channels is high.

- To provide performance benchmarks, we present theoretical limits when the channel statistics are known at the transmitter (Section 3.3).

In addition, we provide simulation results to show that the proposed UCB bandit based algorithms converge to the theoretical limits for large numbers of blocks.

1.3 Organization of the Thesis

This thesis is organized as follows. Chapter 2 focuses on the investigation of optimal channel switching and randomization over flat-fading channels to maximize outage capacity. In Chapter 3, the thesis explores channel switching strategies based on the UCB bandit algorithm to achieve throughput maximization when channel statistics are unknown. Then, concluding remarks and future research directions are discussed in Chapter 4.

Chapter 2

Optimal Channel Switching and Randomization over Flat-Fading Channels for Outage Capacity Maximization

In this chapter, we propose a problem formulation related to optimal channel switching and randomization for maximizing the outage capacity in wireless communication systems. The channels between the transmitter and the receiver are modeled as flat-fading channels with additive Gaussian noise, and the channel distribution information (CDI) is assumed to be available at both ends. However, the channel state information (CSI) is only known at the receiver. By considering average power and outage probability constraints, we derive optimal strategies for channel switching and randomization based on the outage capacity metric. We also extend our study to multiuser systems, addressing the maximization of total outage capacity for multiple transmitter-receiver pairs. Our approach employs optimization theory to provide a solution in the multiuser scenario. The findings in this chapter contribute to the literature by proposing the problem of optimal channel switching and randomization for outage capacity maximization and introducing novel solutions for both single-user and multiuser scenarios.

2.1 System Model

We consider the presence of $K \geq 2$ different channels (frequency bands) for communication between a transmitter and a receiver, which can perform channel switching (time-sharing) among these K channels to enhance the capacity of the communication system. As described in [4], during channel switching, only one channel is utilized for the communication between the transmitter and the receiver at any given time, and the transmitter informs the receiver about the occupied channel for synchronization purposes.

In this work, the channels are modeled as flat-fading and additive Gaussian noise channels with various bandwidths and constant power spectral density levels. In particular, for channel i, B_i and $N_i/2$ denote, respectively, the bandwidth and the constant power spectral density level of the additive Gaussian noise, where $i \in \{1, \ldots, K\}$. Also, g_i represents the channel gain (i.e., the magnitude square of the channel coefficient) related to channel i between the transmitter and the receiver. It is assumed for each $i \in \{1, \ldots, K\}$ that g_i is a continuous random variable, and the support of its probability density function (PDF) is an interval. In addition, $g_1, \ldots, g_K$ are modeled as independent random variables. For the information about the channel gains at the transmitter and the receiver, it is assumed that the CDI of each channel is available at both the transmitter and the receiver; however, the CSI of the channels is available only at the receiver.[1] In this setting, the outage capacity can be employed as a well-suited performance metric [13].

In addition to channel switching, we also consider randomization of transmit power and outage probability for each channel by considering up to $L \geq 2$ different values per channel. In this regard, we employ *time-sharing factors* denoted by

[1]The CSI of the channels can be obtained at the receiver based on pilot (training) based channel estimation approaches. Since flat block-fading channels are considered, channel estimation can be performed at the beginning of each block. Also, CDI can be obtained based on previous channel estimates and/or statistical channel models related to the operating environment. In particular, by utilizing previous channel estimates, the receiver can form statistical knowledge about the channel parameters, and share this knowledge with the transmitter via a feedback mechanism.

$\lambda_{i,j}$ for $i = 1, \dots, K$ and $j = 1, \dots, L$, which satisfy $\sum_{i=1}^{K} \sum_{j=1}^{L} \lambda_{i,j} = 1$ and $\lambda_{i,j} \geq 0$ for all i and j. Namely, $\lambda_{i,j}$ corresponds to the fraction of time when channel i is used with the jth set of parameters (i.e., power level and outage probability pairs) for that channel, which can be denoted as $\boldsymbol{\theta}_{i,j}$ in general. For example, suppose that $K = 3$, $L = 2$, $\lambda_{1,1} = 0.2$, $\lambda_{1,2} = 0.3$, $\lambda_{2,1} = 0.15$, $\lambda_{2,2} = 0$, $\lambda_{3,1} = 0.1$, and $\lambda_{3,1} = 0.25$. Then, over a communication duration of T seconds, channel 1 is used with parameter $\boldsymbol{\theta}_{1,1}$ during the time interval of $[0, 0.2T]$ sec., channel 1 is used with parameter $\boldsymbol{\theta}_{1,2}$ during $[0.2T, 0.5T]$ sec., channel 2 is used with parameter $\boldsymbol{\theta}_{2,1}$ during $[0.5T, 0.65T]$ sec., channel 3 is used with parameter $\boldsymbol{\theta}_{3,1}$ during $[0.65T, 0.75T]$ sec., and channel 3 is used with parameter $\boldsymbol{\theta}_{3,2}$ during $[0.75T, T]$ sec. Channel switching and randomization can provide performance improvements as they facilitate achievement of convex combinations of the performance metric (outage capacity) via time-sharing.

2.2 Maximization of Average Outage Capacity Under Average Power and Outage Probability Constraints

The aim is to perform optimal channel switching and randomization for the maximization of the average outage capacity under average power and outage probability constraints. The outage capacity is defined as the maximum data rate that can be transmitted over a channel with a certain probability of outage (i.e., no proper decoding). Considering the time interval corresponding to the time-sharing factor $\lambda_{i,j}$, the outage capacity of channel i can be expressed as [13]

$$(1 - \varepsilon_{i,j}) B_i \log_2 (1 + \gamma_{i,j}) , \tag{2.1}$$

where $\varepsilon_{i,j}$ is the outage probability and $\gamma_{i,j}$ denotes the target SNR level below which channel i will be in outage. The outage probability $\varepsilon_{i,j}$ specifies the probability that the SNR level is below the target SNR level $\gamma_{i,j}$, in which case proper decoding cannot be performed. The outage probability can be calculated as [13]

$$\varepsilon_{i,j} = \mathbb{P}\left(\frac{g_i P_{i,j}}{N_i B_i} < \gamma_{i,j}\right) = F_{g_i}\left(\frac{\gamma_{i,j} N_i B_i}{P_{i,j}}\right), \tag{2.2}$$

where $P_{i,j}$ represents the power level and F_{g_i} denotes the cumulative distribution function (CDF) of g_i.

From (2.2), the outage capacity in (2.1) can be stated as

$$C_i(P_{i,j}, \varepsilon_{i,j}) = (1 - \varepsilon_{i,j}) B_i \log_2\left(1 + \frac{P_{i,j} F_{g_i}^{-1}(\varepsilon_{i,j})}{N_i B_i}\right), \tag{2.3}$$

where $F_{g_i}^{-1}$ denotes the inverse CDF of g_i. For channel i, $P_{i,j}$ and $\varepsilon_{i,j}$ (or, equivalently, $\gamma_{i,j}$) can be regarded as design parameters that are subject to power limits and acceptable levels of outage probability, respectively. Accordingly, the optimal channel switching and randomization problem for outage capacity maximization is proposed as follows:

$$\max_{\{\lambda_{i,j}, P_{i,j}, \varepsilon_{i,j}\}_{i,j=1}^{K,L}} \sum_{i=1}^{K} \sum_{j=1}^{L} \lambda_{i,j} C_i(P_{i,j}, \varepsilon_{i,j}) \tag{2.4a}$$

$$\text{subject to} \sum_{i=1}^{K} \sum_{j=1}^{L} \lambda_{i,j} P_{i,j} \leq P_{\text{av}}, \tag{2.4b}$$

$$\sum_{i=1}^{K} \sum_{j=1}^{L} \lambda_{i,j} \varepsilon_{i,j} \leq \varepsilon_{\text{av}}, \tag{2.4c}$$

$$P_{i,j} \in [0, P_{\text{pk}}], \ \forall i \in \mathcal{S}_{\text{c}}, \ \forall j \in \mathcal{S}_{\text{r}}, \tag{2.4d}$$

$$\varepsilon_{i,j} \in [0, \varepsilon_{\text{pk}}], \ \forall i \in \mathcal{S}_{\text{c}}, \ \forall j \in \mathcal{S}_{\text{r}}, \tag{2.4e}$$

$$\sum_{i=1}^{K} \sum_{j=1}^{L} \lambda_{i,j} = 1, \tag{2.4f}$$

$$\lambda_{i,j} \geq 0, \ \forall i \in \mathcal{S}_{\text{c}}, \ \forall j \in \mathcal{S}_{\text{r}}, \tag{2.4g}$$

where $\mathcal{S}_{\text{c}} \triangleq \{1, \ldots, K\}$, $\mathcal{S}_{\text{r}} \triangleq \{1, \ldots, L\}$, P_{av} is the average power constraint, ε_{av} is the average outage probability constraint, P_{pk} is the peak power constraint, and ε_{pk} is peak outage probability constraint. Due to practical reasons, it is assumed that $P_{\text{av}} < P_{\text{pk}}$. Although the problem in (2.4) is a challenging *non-convex* optimization problem over a $3KL$-dimensional space in general, we characterize its solution, denoted by $\{\lambda_{i,j}^*, P_{i,j}^*, \varepsilon_{i,j}^*\}_{i,j=1}^{K,L}$, in the following proposition.

Proposition 1: *Consider the following problem:*

$$\max_{\{\nu,P_1,P_2,\varepsilon_1,\varepsilon_2\}} \nu\, C_{\max}(P_1,\varepsilon_1) + (1-\nu)C_{\max}(P_2,\varepsilon_2) \tag{2.5a}$$

$$\text{subject to } \nu P_1 + (1-\nu)P_2 \leq P_{\text{av}}\,, \tag{2.5b}$$

$$\nu\varepsilon_1 + (1-\nu)\varepsilon_2 \leq \varepsilon_{\text{av}}\,, \tag{2.5c}$$

$$P_1 \in [0, P_{\text{pk}}],\ P_2 \in [0, P_{\text{pk}}]\,, \tag{2.5d}$$

$$\varepsilon_1 \in [0, \varepsilon_{\text{pk}}],\ \varepsilon_2 \in [0, \varepsilon_{\text{pk}}],\ \nu \in [0,1]\,, \tag{2.5e}$$

where

$$C_{\max}(P,\varepsilon) = \max_{i\in\mathcal{S}_{\text{c}}} C_i(P,\varepsilon)\,. \tag{2.6}$$

Let $(\nu^*, P_1^*, P_2^*, \varepsilon_1^*, \varepsilon_2^*)$ *denote the solution of* (2.5), *and let* ℓ *and* m *be defined as*

$$\ell = \arg\max_{i\in\mathcal{S}_{\text{c}}} C_i\big(P_1^*, \varepsilon_1^*\big),\ \ m = \arg\max_{i\in\mathcal{S}_{\text{c}}} C_i\big(P_2^*, \varepsilon_2^*\big)\,. \tag{2.7}$$

Then, the solution of (2.4) *can be specified as one of the following strategies:*

- ***Conventional Strategy– Single Channel with no Randomization:*** *If* $\nu^* = 1$ *or if* $\ell = m$ *and* $(P_1^*, \varepsilon_1^*) = (P_2^*, \varepsilon_2^*)$, *then a solution of* (2.4) *can be stated as* $\lambda^*_{\ell,1} = 1$, $P^*_{\ell,1} = P_1^*$, $\varepsilon^*_{\ell,1} = \varepsilon_1^*$, *and* $\lambda^*_{i,j} = 0$ *for all* $(i,j) \neq (\ell,1)$.[2] *Similarly, if* $\nu^* = 0$, *a solution of* (2.4) *can be stated as* $\lambda^*_{m,1} = 1$, $P^*_{m,1} = P_2^*$, $\varepsilon^*_{m,1} = \varepsilon_2^*$, *and* $\lambda^*_{i,j} = 0$ *for all* $(i,j) \neq (m,1)$. *Namely, one of the channels is used exclusively without any randomization.*

- ***CS2 Strategy– Channel Switching between Two Channels with no Randomization:*** *If* $\ell \neq m$ *and* $\nu^* \in (0,1)$, *then a solution of* (2.4) *can be stated as* $\lambda^*_{\ell,1} = \nu^*$, $P^*_{\ell,1} = P_1^*$, $\varepsilon^*_{\ell,1} = \varepsilon_1^*$, $\lambda^*_{m,1} = 1-\nu^*$, $P^*_{m,1} = P_2^*$, $\varepsilon^*_{m,1} = \varepsilon_2^*$, *and* $\lambda^*_{i,j} = 0$ *for all* $(i,j) \notin \{(\ell,1),(m,1)\}$. *That is, channel switching is performed between two channels without any randomization over each channel.*

- ***Rand2 Strategy– Randomization between two Power-Outage Probability pairs over a Single Channel:*** *If* $\ell = m$, $\nu^* \in (0,1)$,

[2] *It is noted that the second index in the subscripts is chosen as* 1 *arbitrarily as it does not affect the performance of the solution.*

and $(P_1^*, \varepsilon_1^*) \neq (P_2^*, \varepsilon_2^*)$*, then a solution of* (2.4) *can be stated as* $\lambda_{\ell,1}^* = \nu^*$, $P_{\ell,1}^* = P_1^*$, $\varepsilon_{\ell,1}^* = \varepsilon_1^*$, $\lambda_{\ell,2}^* = 1 - \nu^*$, $P_{\ell,2}^* = P_2^*$, $\varepsilon_{\ell,2}^* = \varepsilon_2^*$*, and* $\lambda_{i,j}^* = 0$ *for all* $(i,j) \notin \{(\ell,1),(\ell,2)\}$.[3] *That is, randomization (time-sharing) is performed between two different parameter sets over a single channel (i.e., without channel switching).*

Proof: By introducing a vector variable as $\boldsymbol{\theta}_{i,j} = [P_{i,j}, \varepsilon_{i,j}]$, the problem in (2.4) can be stated as

$$\max_{\{\lambda_{i,j},\boldsymbol{\theta}_{i,j}\}_{i,j=1}^{K,L}} \sum_{i=1}^{K}\sum_{j=1}^{L} \lambda_{i,j} C_i(\boldsymbol{\theta}_{i,j}) \tag{2.8a}$$

$$\text{subject to} \sum_{i=1}^{K}\sum_{j=1}^{L} \lambda_{i,j}\boldsymbol{\theta}_{i,j} \leq [P_{\text{av}}, \varepsilon_{\text{av}}]\,, \tag{2.8b}$$

$$\boldsymbol{\theta}_{i,j} \in \mathcal{T}\,,\ \forall i \in \mathcal{S}_{\text{c}},\ \forall j \in \mathcal{S}_{\text{r}}\,, \tag{2.8c}$$

$$\sum_{i=1}^{K}\sum_{j=1}^{L} \lambda_{i,j} = 1,\ \lambda_{i,j} \geq 0,\ \forall i \in \mathcal{S}_{\text{c}}, \forall j \in \mathcal{S}_{\text{r}}, \tag{2.8d}$$

where $\mathcal{T} \triangleq [0, P_{\text{pk}}] \times [0, \varepsilon_{\text{pk}}]$. To characterize the solution of (2.8), we first present the following problem:

$$\max_{p(\boldsymbol{\theta})} \int C_{\max}(\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{2.9a}$$

$$\text{subject to} \int \boldsymbol{\theta} p(\boldsymbol{\theta}) d\boldsymbol{\theta} \leq [P_{\text{av}}, \varepsilon_{\text{av}}]\,, \tag{2.9b}$$

$$\boldsymbol{\theta} \in \mathcal{T}\,,\ \int p(\boldsymbol{\theta}) d\boldsymbol{\theta} = 1\,,\ p(\boldsymbol{\theta}) \geq 0\,,\ \forall \boldsymbol{\theta} \tag{2.9c}$$

where $C_{\max}(\boldsymbol{\theta}_{i,j}) = \max_{i \in \mathcal{S}_{\text{c}}} C_i(\boldsymbol{\theta}_{i,j})$ as in (2.6), and $p(\boldsymbol{\theta})$ denotes the PDF of $\boldsymbol{\theta}$. The problem in (2.9) provides an upper bound on (2.8) since it employs the maximum of C_i's in its objective function and more generic weighting coefficients.

The problems in the form of (2.9) have been investigated in various contexts in the literature; e.g., [9]–[11]. By adopting a similar approach, we define set $\mathcal{U}$ and set $\mathcal{W}$ as follows: $\mathcal{U} = \left\{(\boldsymbol{\theta}, C_{\max}(\boldsymbol{\theta})) \text{ for all } \boldsymbol{\theta} \in \mathcal{T}\right\}$ and $\mathcal{W} =$

[3] *The second indices in the subscripts are chosen as* 1 *and* 2 *arbitrarily.*

$\{ (\int \boldsymbol{\theta} p(\boldsymbol{\theta}) d\boldsymbol{\theta}, \int C_{\max}(\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta})$ for all $\int p(\boldsymbol{\theta}) d\boldsymbol{\theta} = 1,\ p(\boldsymbol{\theta}) \geq 0,\ \boldsymbol{\theta} \in \mathcal{T} \}$. It is noted that $\mathcal{W}$ contains the solution of (2.9) since it consists of all possible values of $\int C_{\max}(\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$ and $\int \boldsymbol{\theta} p(\boldsymbol{\theta}) d\boldsymbol{\theta}$ subject to the constraints in (2.9c). Also, via the arguments in [9]–[11], it can be shown that $\mathcal{W}$ is equal to the convex hull of $\mathcal{U}$; i.e., $\mathcal{W} = \text{hull}(\mathcal{U})$. Therefore, as a result of Carathéodory's theorem [39, 40], any element of $\mathcal{W}$ can be expressed as a convex combination of $\dim(\mathcal{U}) + 1 = 3$ elements in $\mathcal{U}$, where $\dim(\mathcal{U}) = 2$ since $\mathcal{U} \subset \mathbb{R}^2$. In addition, as the maximizer of (2.9) must reside on the boundary of $\text{hull}(\mathcal{U})$, the solution of (2.9) can be achieved by a convex combination of $\dim(\mathcal{U}) = 2$ elements in $\mathcal{U}$ by Carathéodory's theorem [39, 40]. Hence, an optimal $p(\boldsymbol{\theta})$ can be specified as $p(\boldsymbol{\theta}) = \nu\delta(\boldsymbol{\theta} - \boldsymbol{\theta}_1) + (1 - \nu)\delta(\boldsymbol{\theta} - \boldsymbol{\theta}_2)$, where $\delta(\cdot)$ denotes the Dirac delta function. By inserting this specific $p(\boldsymbol{\theta})$ expression into (2.9), we obtain the following problem:

$$\max_{\{\nu, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2\}} \ \nu\, C_{\max}(\boldsymbol{\theta}_1) + (1-\nu) C_{\max}(\boldsymbol{\theta}_2) \tag{2.10a}$$

$$\text{subject to } \nu\boldsymbol{\theta}_1 + (1-\nu)\boldsymbol{\theta}_2 \leq [P_{\text{av}}, \varepsilon_{\text{av}}]\,, \tag{2.10b}$$

$$\boldsymbol{\theta}_1 \in \mathcal{T}\,,\ \boldsymbol{\theta}_2 \in \mathcal{T}\,,\ \nu \in [0,1]\,, \tag{2.10c}$$

which is guaranteed to achieve the same maximum value as (2.9). It is noted that the problem in (2.10) is the same as (2.5) in Proposition 1 as $\boldsymbol{\theta}_1 = [P_1, \varepsilon_1]$ and $\boldsymbol{\theta}_2 = [P_2, \varepsilon_2]$. Let $(\nu^*, \boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^*)$ denote the solution of (2.10), and let ℓ and m be defined as in (2.7), where $\boldsymbol{\theta}_1^* = [P_1^*, \varepsilon_1^*]$ and $\boldsymbol{\theta}_2^* = [P_2^*, \varepsilon_2^*]$. The maximum value of (2.10) (equivalently, of (2.9)) can be achieved by the problem in (2.4) via the conventional strategy, the CS2 strategy, or the Rand2 strategy, as specified in the proposition. Since the problem in (2.9) is an upper bound on the problem in (2.4), and the maximum value of (2.9) can be achieved by (2.4) via the conventional, CS2, or Rand2 strategies, it is concluded that the solution of (2.4) can be characterized by the strategies specified in the proposition. ■

Based on Proposition 1, the solution of (2.4) can be obtained by searching over three possible strategies, which significantly reduces the computation complexity of the problem. Namely, instead of a search over a $3KL$-dimensional space, the optimal solution can be obtained via a four-dimensional search as specified in the following: First, it is noted that $C_{\max}(P, \varepsilon)$ in (2.6) is a monotone increasing function of $P > 0$ for any $\varepsilon \in [0, 1]$ since each $C_i(P, \varepsilon)$ in (2.3) is monotone increasing with respect to P. Therefore, any approach with $\nu P_1 + (1-\nu)P_2 < P_{\text{av}}$ cannot be a solution of (2.5) since it can always be improved by increasing at least one of the power levels (as $P_{\text{av}} < P_{\text{pk}}$). Hence, $\nu P_1 + (1-\nu)P_2 = P_{\text{av}}$ must be satisfied in (2.5). By utilizing this equality, the problem in (2.5) can be simplified as

$$\max_{\{P_1, P_2, \varepsilon_1, \varepsilon_2\}} \frac{P_{\text{av}} - P_2}{P_1 - P_2} C_{\max}(P_1, \varepsilon_1) + \frac{P_1 - P_{\text{av}}}{P_1 - P_2} C_{\max}(P_2, \varepsilon_2) \tag{2.11a}$$

$$\text{subject to } \frac{P_{\text{av}} - P_2}{P_1 - P_2} \varepsilon_1 + \frac{P_1 - P_{\text{av}}}{P_1 - P_2} \varepsilon_2 \leq \varepsilon_{\text{av}} \,, \tag{2.11b}$$

$$P_1 \in (P_{\text{av}}, P_{\text{pk}}] \,, \;\; P_2 \in [0, P_{\text{av}}] \,, \tag{2.11c}$$

$$\varepsilon_1 \in [0, \varepsilon_{\text{pk}}] \,, \;\; \varepsilon_2 \in [0, \varepsilon_{\text{pk}}] \,. \tag{2.11d}$$

Let $(P_1^*, P_2^*, \varepsilon_1^*, \varepsilon_2^*)$ denote the solution of (2.11). Then, ν^* is obtained as $\nu^* = (P_{\text{av}} - P_2^*)/(P_1^* - P_2^*)$. Also, ℓ and m are calculated as in (2.7). Then, the solution of (2.4) corresponds to the use of the conventional, CS2, or Rand2 strategies, as specified in Proposition 1 based on the parameters ν^*, P_1^*, P_2^*, ε_1^*, ε_2^*, ℓ, and m. (To specify the computational complexity of solving (2.11) via exhaustive search, suppose that ε_1 and ε_2 are discretized with a step size of Δ_0, and P_1 and P_2 are discretized with step sizes of Δ_1 and Δ_2, respectively. Then, the objective function in (2.11) should be evaluated about $\varepsilon_{\text{pk}}^2 P_{\text{av}}(P_{\text{pk}} - P_{\text{av}})/(\Delta_0^2 \Delta_1 \Delta_2)$ times to find the solution.)

Remark 1: Proposition 1 also implies that the use of both channel switching and randomization is not needed to achieve the solution of (2.4). In other words, when optimal channel switching is performed as in the CS2 strategy, randomization over any of the channels does not bring any additional benefits. Similarly, when the optimal solution is achieved via randomization over a single channel as in the Rand2 strategy, the average outage capacity cannot be increased further via channel switching.

2.3 Extension to Multiuser Systems

In this section, we consider a multiuser system with U transmitter-receiver pairs, i.e., users, which aim to communicate over the K available channels with the ability of performing channel switching and randomization. The aim is to maximize the total outage capacity of the users under average power and outage probability constraints by jointly optimizing the parameters of all the users in a centralized fashion. Accordingly, the following problem is proposed (cf. (2.4)):

$$\max_{\{\lambda_{i,j}^{(u)},P_{i,j}^{(u)},\varepsilon_{i,j}^{(u)}\}_{i,j,u=1}^{K,L,U}} \sum_{u=1}^{U}\sum_{i=1}^{K}\sum_{j=1}^{L}\lambda_{i,j}^{(u)}C_i^{(u)}\big(P_{i,j}^{(u)},\varepsilon_{i,j}^{(u)}\big) \tag{2.12a}$$

$$\text{subject to } \sum_{i=1}^{K}\sum_{j=1}^{L}\lambda_{i,j}^{(u)}P_{i,j}^{(u)} \le P_{\rm av}^{(u)},\ \forall u\in\mathcal{S}_{\rm u}\,, \tag{2.12b}$$

$$\sum_{i=1}^{K}\sum_{j=1}^{L}\lambda_{i,j}^{(u)}\varepsilon_{i,j}^{(u)} \le \varepsilon_{\rm av}^{(u)},\ \forall u\in\mathcal{S}_{\rm u}\,, \tag{2.12c}$$

$$P_{i,j}^{(u)} \in \big[0,P_{\rm pk}^{(u)}\big],\ \forall i\in\mathcal{S}_{\rm c},\ \forall j\in\mathcal{S}_{\rm r},\ \forall u\in\mathcal{S}_{\rm u}\,, \tag{2.12d}$$

$$\varepsilon_{i,j}^{(u)} \in \big[0,\varepsilon_{\rm pk}^{(u)}\big],\ \forall i\in\mathcal{S}_{\rm c},\ \forall j\in\mathcal{S}_{\rm r},\ \forall u\in\mathcal{S}_{\rm u}\,, \tag{2.12e}$$

$$\sum_{i=1}^{K}\sum_{j=1}^{L}\lambda_{i,j}^{(u)} \le 1,\ \forall u\in\mathcal{S}_{\rm u}\,, \tag{2.12f}$$

$$\lambda_{i,j}^{(u)} \ge 0\,,\ \forall i\in\mathcal{S}_{\rm c},\ \forall j\in\mathcal{S}_{\rm r},\ \forall u\in\mathcal{S}_{\rm u}\,, \tag{2.12g}$$

$$\sum_{u=1}^{U}\sum_{j=1}^{L}\lambda_{i,j}^{(u)} \le 1,\ \forall i\in\mathcal{S}_{\rm c}\,, \tag{2.12h}$$

where $\mathcal{S}_{\rm u} \triangleq \{1,\dots,U\}$, and the capacity function and the parameters are defined as in (2.4) with the addition of superscript (u) for denoting the user index. The constraint in (2.12h) is required in multiuser scenarios to make sure that the total time-sharing factor over each channel does not exceed one.

The problem in (2.12) is very challenging to solve in general since its solution cannot be reduced to a set of three strategies as in Section 2.2 and exhaustive search has prohibitive complexity. However, we can employ a convex relaxation approach [9], [41] to approximate the non-convex problem in (2.12) with a convex problem with the ability to adjust the approximation accuracy. To this aim, instead of the continuum of values for the $P_{i,j}^{(u)}$ and $\varepsilon_{i,j}^{(u)}$ terms, we consider a set of N_u possible (known) values for them specified as $\big(P_{i,j}^{(u)}, \varepsilon_{i,j}^{(u)}\big) \in \big\{(\tilde{P}_1^{(u)}, \tilde{\varepsilon}_1^{(u)}), \ldots, (\tilde{P}_{N_u}^{(u)}, \tilde{\varepsilon}_{N_u}^{(u)})\big\}$ for each $u \in \mathcal{S}_{\mathrm{u}}$. Accordingly, we define the following vectors:

$$\tilde{\boldsymbol{\lambda}} = \left[\tilde{\boldsymbol{\lambda}}^{(1)} \cdots \tilde{\boldsymbol{\lambda}}^{(U)}\right] \text{ and } \tilde{\boldsymbol{C}} = \left[\tilde{\boldsymbol{C}}^{(1)} \cdots \tilde{\boldsymbol{C}}^{(U)}\right], \tag{2.13}$$

where

$$\begin{aligned} \tilde{\boldsymbol{\lambda}}^{(u)} &= \left[\lambda_{1,1}^{(u)} \cdots \lambda_{1,N_u}^{(u)} \cdots\cdots \lambda_{K,1}^{(u)} \cdots \lambda_{K,N_u}^{(u)}\right], & (2.14)\\ \tilde{\boldsymbol{C}}^{(u)} &= \left[C_1^{(u)}(\tilde{P}_1^{(u)}, \tilde{\varepsilon}_1^{(u)}) \cdots C_1^{(u)}(\tilde{P}_{N_u}^{(u)}, \tilde{P}_{N_u}^{(u)}) \cdots\cdots\right. \\ &\qquad \left. C_K^{(u)}(\tilde{P}_1^{(u)}, \tilde{\varepsilon}_1^{(u)}) \cdots C_K^{(u)}(\tilde{P}_{N_u}^{(u)}, \tilde{P}_{N_u}^{(u)})\right] & (2.15) \end{aligned}$$

for $u \in \mathcal{S}_{\mathrm{u}}$. Based on these definitions, (2.12) is approximated as

$$\max_{\tilde{\boldsymbol{\lambda}}} \ \tilde{\boldsymbol{C}}^T \tilde{\boldsymbol{\lambda}} \tag{2.16a}$$

$$\text{subject to } \mathbf{R}\tilde{\boldsymbol{\lambda}} \preceq \left[P_{\mathrm{av}}^{(1)} \cdots P_{\mathrm{av}}^{(U)}\right]^T, \tag{2.16b}$$

$$\mathbf{E}\tilde{\boldsymbol{\lambda}} \preceq \left[\varepsilon_{\mathrm{av}}^{(1)} \cdots \varepsilon_{\mathrm{av}}^{(U)}\right]^T, \tag{2.16c}$$

$$\mathbf{B}\tilde{\boldsymbol{\lambda}} \preceq \mathbf{1}_U, \ \tilde{\boldsymbol{\lambda}} \succeq \mathbf{0}, \ \mathbf{G}\tilde{\boldsymbol{\lambda}} \preceq \mathbf{1}_K, \tag{2.16d}$$

with $\mathbf{1}_X$ representing a column vector of ones with X elements, $\mathbf{R} = \mathrm{blkdiag}\Big(\mathbf{1}_K^T \otimes \tilde{\boldsymbol{P}}^{(1)}, \ldots, \mathbf{1}_K^T \otimes \tilde{\boldsymbol{P}}^{(U)}\Big)$, $\mathbf{E} = \mathrm{blkdiag}\Big(\mathbf{1}_K^T \otimes \tilde{\boldsymbol{\varepsilon}}^{(1)}, \ldots, \mathbf{1}_K^T \otimes \tilde{\boldsymbol{\varepsilon}}^{(U)}\Big)$, $\mathbf{B} = \mathrm{blkdiag}\Big(\mathbf{1}_{KN_1}^T, \ldots, \mathbf{1}_{KN_U}^T\Big)$, and $\mathbf{G} = \Big[\mathbf{I}_K \otimes \mathbf{1}_{N_1}^T \cdots \mathbf{I}_K \otimes \mathbf{1}_{N_U}^T\Big]$, where blkdiag specifies a block diagonal matrix constructed by the given matrices, $\otimes$ denotes the Kronecker product, $\mathbf{I}_K$ is the $K \times K$ identity matrix, $\tilde{\boldsymbol{P}}^{(u)} = \big[\tilde{P}_1^{(u)} \cdots \tilde{P}_{N_u}^{(u)}\big]$, and $\tilde{\boldsymbol{\varepsilon}}^{(u)} = \big[\tilde{\varepsilon}_1^{(u)} \cdots \tilde{\varepsilon}_{N_u}^{(u)}\big]$ for $u \in \mathcal{S}_{\mathrm{u}}$.

It is noted that (2.16) is a linearly constrained linear optimization problem. Therefore, it can be solved rapidly via linear or convex optimization algorithms such as the simplex or the interior point method [41]. Although (2.16) is an approximation to (2.12), the approximation accuracy can be enhanced by increasing the number of possible values for the $\big(P_{i,j}^{(u)}, \varepsilon_{i,j}^{(u)}\big)$ pairs, i.e., by increasing $N_1, \ldots, N_U$, with the cost of higher computational complexity. (It is noted that the problem in (2.16) has polynomial complexity in the number of variables, which is equal to $K\sum_{u=1}^{U} N_u$ [41].)

Remark 2: It is also possible to add fairness constraints to (2.12) in the form of $\sum_{i=1}^{K}\sum_{j=1}^{L}\lambda_{i,j}^{(u)}C_i^{(u)}\big(P_{i,j}^{(u)},\varepsilon_{i,j}^{(u)}\big) \geq \varsigma_u \ \forall u \in \mathcal{S}_{\mathrm{u}}$. In that case, the solution can be obtained by an approximate problem as in (2.16) by adding linear constraints related to fairness.

2.4 Simulation Results

In this section, numerical examples are presented to corroborate the theoretical results by considering two settings. In the first setting, as in [4], we consider $K = 3$ channels with the following bandwidths and noise levels: $B_1 = 1\,\mathrm{MHz}$, $B_2 = 5\,\mathrm{MHz}$, $B_3 = 10\,\mathrm{MHz}$, $N_1 = 10^{-12}\,\mathrm{W/Hz}$, $N_2 = 10^{-11}\,\mathrm{W/Hz}$, and $N_3 = 10^{-11}\,\mathrm{W/Hz}$. In this setting, there can exist one user or two users in the system. We model the channels as independent Rayleigh fading with the channel gains being exponentially distributed with the following CDFs and PDFs: $F_{g_i^{(u)}}(g) = 1 - e^{-\alpha_i^{(u)}g}$ and $p_{g_i^{(u)}}(g) = \alpha_i^{(u)}e^{-\alpha_i^{(u)}g}$ if $g \geq 0$ for $i \in \{1,2,3\}$ and $u \in \{1,2\}$. The parameters are given by $\alpha_1^{(1)} = 0.25$, $\alpha_2^{(1)} = 1$, and $\alpha_3^{(1)} = 0.75$ for user 1, and $\alpha_1^{(2)} = 0.95$, $\alpha_2^{(2)} = 0.35$, and $\alpha_3^{(2)} = 0.3$ for user 2. Also, the peak power and peak outage constraints are set as $P_{\mathrm{pk}} = 10P_{\mathrm{av}}$ and $\varepsilon_{\mathrm{pk}} = 10\varepsilon_{\mathrm{av}}$.

We first assume that only user 1 exists in the system and investigate the outage capacity maximization problem in (2.4), the solution of which is specified by Proposition 1. In Figure. 2.1, the average outage capacity achieved by the solution of (2.4) (which is obtained via (2.11)) is plotted with respect to the average power constraint, P_{av} for $\varepsilon_{\mathrm{av}} = 0.01$ (labeled as 'Proposed (single user)' in the figure). In addition, the average outage capacities achieved by the conventional strategy, which employs the best channel all the time at the average power limit (and with the corresponding optimal outage probability) are presented for comparison purposes (labeled as 'Conventional (single user)').[4] It

[4]This corresponds to solving (2.5) over ε_1 by setting $\nu = 1$ and $P_1 = P_{\mathrm{av}}$.

is observed from Figure. 2.1 that employing the best channel all the time (i.e., conventional strategy) is not always optimal, in accordance with Proposition 1. For example, when $P_{\mathrm{av}} = 10^{-2}\,\mathrm{mW}$, the proposed (optimal) solution employs channel 1 with a time-sharing factor of 0.33444 and a power level of 0.029901 mW (corresponding to an outage probability of 0.029901) and does not send any power in the remaining duration (i.e., employs zero power with a time-sharing factor of 0.66556). This can be considered as a special case of the Rand2 Strategy (or, CS2 Strategy) in Proposition 1 with $P_2^* = 0$, which results in an average outage capacity of 0.71742 Mbps. (Such an on-off strategy can never be optimal according to the Shannon capacity metric in [4] due to its concavity.) As another example, when $P_{\mathrm{av}} = 5\,\mathrm{mW}$, the proposed (optimal) solution corresponds to the CS2 Strategy, which uses channel 1 with a time-sharing factor of 0.7023 and a power level of 1.5561 mW, and channel 3 with a time-sharing factor of 0.2977 and a power level of 13.125 mW, leading to an average outage capacity of 10.292 Mbps. However, for the same setting, the conventional strategy employs channel 1 exclusively, and achieves an outage capacity of 7.5817 Mbps. On the other hand, when $P_{\mathrm{av}} = 100\,\mathrm{mW}$, the proposed strategy corresponds to the conventional strategy, which utilizes channel 3 exclusively. It is also important to mention that the turning points in Figure. 2.1 occur when a strategy starts employing a different channel.

Next, we consider the presence of both user 1 and user 2 in the system described above, and obtain the proposed channel switching and randomization strategy in Section 2.3. To implement the proposed strategy, we first generate a set of N_u possible values of $\left(P_{i,j}^{(u)}, \varepsilon_{i,j}^{(u)}\right)$ for $u \in \{1, 2\}$ by forming a vector of power levels consisting of 51 values between 0 and P_{pk} (with equal spacing) and similarly a vector of outage probabilities consisting of 51 values between 0 and $\varepsilon_{\mathrm{pk}}$ (hence, $N_u = 2601$ for $u \in \{1, 2\}$).[5] Then, the proposed strategy is obtained by solving (2.16) (labeled as 'Proposed (multiuser)' in Figure. 2.1). For comparison purposes, we also illustrate the conventional strategy in Figure. 2.1 (labeled as 'Conventional (multiuser)'), which performs sequential assignments of users to channels over which they achieve the highest outage capacities by considering all

[5]Increasing N_u further does not lead to notable performance improvement.

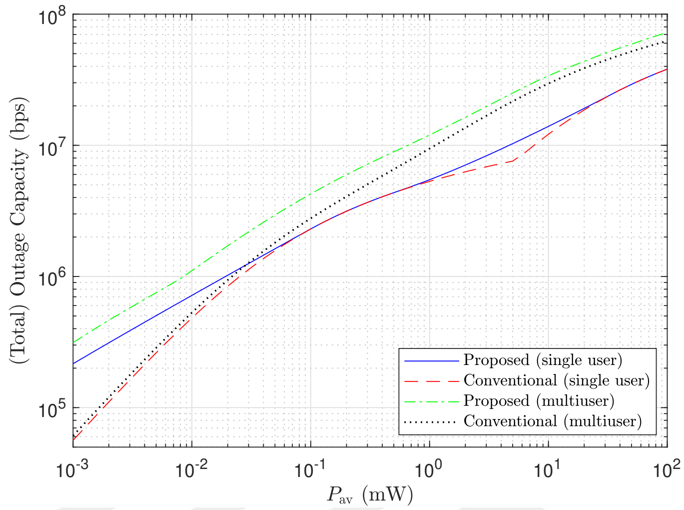

Figure 2.1: Average outage capacity versus P_{av} achieved by the solution of (2.4) (labeled as 'Proposed (single user)') and by the Conventional Strategy in Section 2.2 for the single user scenario. Also, the total average outage capacity is plotted versus P_{av} for the multiuser scenario with $U = 2$ users considering both the proposed approach based on (2.16) and the conventional approach.

possible assignment orders for the users. (In the conventional strategy, at most one user is assigned to a channel, and the channel is completely utilized by a user once it is assigned; i.e., no time-sharing.) It is observed from Figure. 2.1 that the proposed approach for the multiuser scenario outperforms the conventional strategy for all values of P_{av}, especially for low P_{av}'s.

Next, we investigate the effects of N_u, the number of possible values of $\left(P_{i,j}^{(u)}, \varepsilon_{i,j}^{(u)}\right)$ selected for solving (2.16), on the performance of the proposed strategy. For this purpose, in Figure. 2.2, we plot the total average outage capacity versus P_{av} for the setting described above ($U = 2$ users and $K = 3$ channels) by considering different values of N_u, where $\varepsilon_{\text{av}} = 0.01$. It can be observed that after a certain value of N_u, the solution of eqn. (2.16) converges. This implies that with a sufficiently high value of N_u (around 1000), the solution of (2.16) becomes very close to that of the original problem in (2.12). Hence, the proposed strategy based on (2.16) is expected to perform very close to the optimal solution in practice.

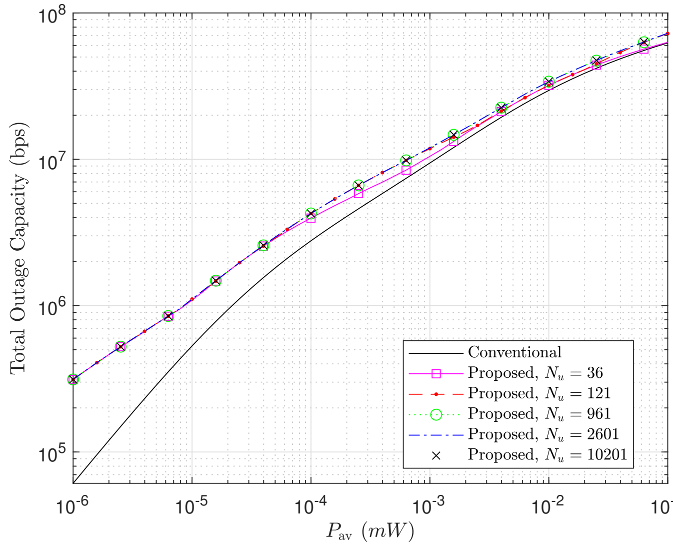

Figure 2.2: Total average outage capacity versus P_{av} achieved by the solution of (16) (labeled as 'Proposed') and by the conventional strategy considering different values of N_u, where $U = 2$ and $\varepsilon_{\text{av}} = 0.01$.

Finally, we consider the second setting with $K = 5$ channels, which have the following bandwidths and noise levels: $B_1 = 1\,\text{MHz}$, $B_2 = 5\,\text{MHz}$, $B_3 = 10\,\text{MHz}$, $B_4 = 15\,\text{MHz}$, $B_5 = 20\,\text{MHz}$, $N_1 = 10^{-12}\,\text{W/Hz}$, $N_2 = 10^{-11}\,\text{W/Hz}$, $N_3 = 10^{-11}\,\text{W/Hz}$, $N_4 = 10^{-12}\,\text{W/Hz}$, and $N_5 = 10^{-11}\,\text{W/Hz}$. The number of users, U, is varied from 1 to 10. The channel gain parameters for the users (i.e., $\alpha_i^{(u)}$'s) are generated uniformly and independently over $[0.1, 1]$ in MATLAB with seed 1 (namely, "$0.1 + 0.9 * \text{rand}(10, 5)$"). As in the first setting, $P_{\text{pk}} = 10P_{\text{av}}$ and $\varepsilon_{\text{pk}} = 10\varepsilon_{\text{av}}$. Also, the N_u possible values of $\left(P_{i,j}^{(u)}, \varepsilon_{i,j}^{(u)}\right)$ are generated as in the previous paragraph for each u.

In Figure. 2.3, the average outage capacity per user is plotted versus the number of users for three different values of P_{av}, where $\varepsilon_{\text{av}} = 0.01$. It is observed that the proposed strategy based on (2.16) outperforms the conventional strategy in all cases, especially for low values of P_{av}. It is also noted that as the number of users increases, the average outage capacity per user tends to decrease in general since there exist limited resources. However, this trend is not monotone since the channel parameters of the users are generated randomly (i.e., a new user with favorable channel characteristics may improve the average outage capacity per

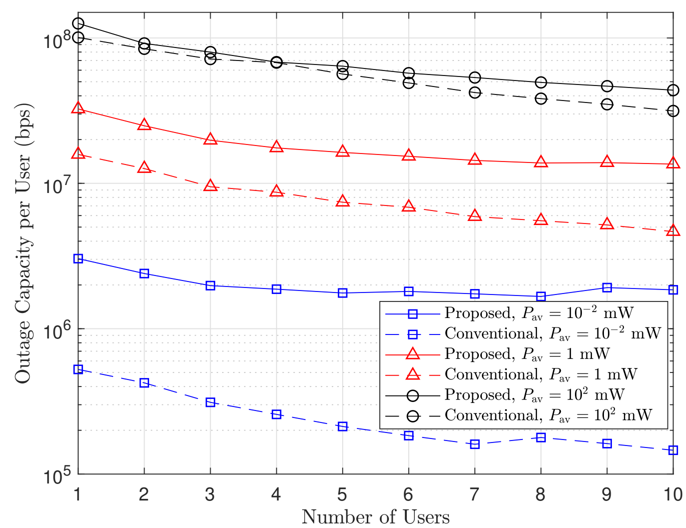

Figure 2.3: Average outage capacity per user versus the number users for the proposed approach based on (2.16) and the conventional approach, where $K = 5$ and $\varepsilon_{\text{av}} = 0.01$.

user).

To provide examples, we present the tables for $U = 4$ users with $N_u = 2601$, $\varepsilon_{\text{av}} = 0.01$, and $K = 5$, as specified above. Table 2.1 depicts the solution of the proposed strategy while Table 2.2 presents the solution of the conventional strategy, where $P_{\text{av}} = 10^{-2}$ mW. For this setting, the proposed strategy significantly outperforms the conventional strategy by achieving a per-user outage capacity of 1.8702 Mbps while the conventional strategy yields 0.25746 Mbps. It is observed from Table 2.1 that when P_{av} is small, the proposed strategy employs a special case of randomization or channel switching in which it uses a high power or a high value of ε with a small time sharing factor to have high outage capacity. For this special case, the user does not utilize channels with a complete time sharing factor of 1 (i.e., employ an on-off technique), and it beats the conventional approach. On other hand, in the conventional approach (i.e., Table 2.2) the users utilize the channels completely without any switching. Namely, users 3, 2, 4, and 1 employ channels 1, 4, 2, and 5, respectively.

The solutions of the proposed and conventional strategies are presented in

Tables 2.3 and 2.4 for $P_{\text{av}} = 1\,\text{mW}$. For this case, again the proposed approach beats the conventional approach by yielding a per-user outage capacity of 17.5198 Mbps while the conventional approach achieves 8.6856 Mbps. It can be observed in Table 2.3 that user 3 performs channel switching while user 4 employs randomization, while the other users are adopting the on-off strategy (i.e., a special case of randomization or channel switching). By looking at Table 2.4, it becomes clear that the conventional approach assigns channels 4, 5, 1, and 3, to users 2, 4, 1, and 3, respectively. From Tables 2.3 and 2.5, the effects of channel switching as a result of the proposed strategy are observed clearly.

Finally, the conventional and proposed approaches are summarized in Tables 2.5 and 2.6 for $P_{\text{av}} = 10^2\,\text{mW}$. In this setting, the difference between the outage capacities per-user obtained by the proposed and conventional strategies is not significant as compared to the previous cases. The proposed strategy achieves an outage capacity per-user of 68.1306 Mbps while the conventional approach yields 67.6801 Mbps.

Table 2.1: Solution of the proposed formulation in (2.16) for $P_{\text{av}} = 10^{-2}\,\text{mW}$, $U = 4$ and $K = 5$.

User	Channel	Power (W)	ε	Outage Capacity (Mbps)	λ
1	4	9.1837×10^{-5}	0.0918	27.8673	0.1089
2	4	1×10^{-4}	0.1000	17.5831	0.1000
3	1	1.8367×10^{-5}	0.0184	02.0987	0.5444
4	4	1×10^{-4}	0.1000	15.4579	0.1000

Table 2.2: Solution of the conventional strategy for $P_{\text{av}} = 10^{-2}\,\text{mW}$, $U = 4$ and $K = 5$.

User	Channel	Power (W)	ε	Outage Capacity (Mbps)
3	1	P_{av}	0.0100	0.5898
2	4	P_{av}	0.0100	0.3236
4	2	P_{av}	0.0100	0.09531
1	5	P_{av}	0.0100	0.02109

Table 2.3: Solution of the proposed formulation in (2.16) for $P_{\text{av}} = 1\,\text{mW}$, $U = 4$ and $K = 5$.

User	Channel	Power (W)	ε	Outage Capacity (Mbps)	λ
1	4	2.2×10^{-3}	0.0224	62.3172	0.4455
2	4	3.5×10^{-3}	0.0347	60.4371	0.2882
3	1	4.0816×10^{-4}	0.0041	4.1269	0.8739
3	5	5.1×10^{-3}	0.0510	42.9648	0.1261
4	4	3.7×10^{-3}	0.0367	58.8222	0.1599
4	4	3.9×10^{-3}	0.0388	60.8403	0.1064

Table 2.4: Solution of the conventional strategy for $P_{\text{av}} = 1\,\text{mW}$, $U = 4$ and $K = 5$.

User	Channel	Power (W)	ε	Outage Capacity (Mbps)
2	4	P_{av}	0.0100	19.8194
4	5	P_{av}	0.0100	6.3232
1	1	P_{av}	0.0100	5.4493
3	3	P_{av}	0.0100	3.1511

Table 2.5: Solution of the proposed formulation in (2.16) for $P_{\text{av}} = 10^2\,\text{mW}$, $U = 4$ and $K = 5$.

User	Channel	Power (W)	ε	Outage Capacity (Mbps)	λ
1	4	0.1020	0.0102	126.7395	0.98
2	2	0.0408	0.0041	12.4518	0.5800
2	4	0.1429	0.0143	120.7345	0.0200
2	5	0.1837	0.0184	88.8274	0.4000
3	2	0.0408	0.0041	18.2917	0.4000
3	5	0.1224	0.0122	88.6174	0.1000
3	5	0.1429	0.0143	96.8927	0.5000
4	3	0.0816	0.0082	33.3077	0.1000
4	3	0.1020	0.0102	39.1149	0.9000

Table 2.6: Solution of the conventional strategy for $P_{\text{av}} = 10^2\,\text{mW}$, $U = 4$ and $K = 5$.

User	Channel	Power (W)	ε	Outage Capacity (Mbps)
2	4	P_{av}	0.0100	107.8067
4	5	P_{av}	0.0100	92.8237
3	3	P_{av}	0.0100	46.3619
1	2	P_{av}	0.0100	23.7373

2.5 Concluding Remarks

This chapter addressed the problem of optimal channel switching and randomization for maximizing the outage capacity in flat-fading Gaussian noise channels under average power and outage probability constraints. Firstly, the problem of optimal channel switching and randomization based on the outage capacity metric was proposed, which is a novel contribution to the literature. Through extensive simulations, the performance of different strategies was evaluated, and the results consistently demonstrated the effectiveness of the proposed approaches in achieving higher outage capacity compared to conventional strategies. For the single-user scenario, it was shown that the optimal solution can be achieved through three strategies: transmission over a single channel with no randomization, channel switching between two channels with no randomization, or randomization between two parameter sets over a single channel. The numerical examples provided compelling evidence of the superiority of these optimal strategies, showcasing higher outage capacity compared to alternative methods. Moreover, we investigated the optimal channel switching and randomization problem for multiuser systems, considering multiple transmitter and receiver pairs. An algorithm was developed to calculate an approximately optimal solution based on a linearly constrained linear optimization problem. The numerical simulations conducted on multiuser scenarios confirmed the effectiveness of the proposed algorithm, showing significant improvements in outage capacity compared to traditional methods. The numerical examples provided in the study served to validate the theoretical results and reinforce the superiority of the proposed approach. These results consistently demonstrated higher outage capacity and improved performance compared to conventional strategies, providing empirical evidence to support the claims made in the chapter.

Overall, the findings from the numerical simulations reaffirm the potential of time-sharing techniques such as channel switching and parameter randomization in enhancing the performance of communication systems. The validated results highlight the practical applicability of the proposed strategies and emphasize their effectiveness in maximizing the outage capacity under average power and outage probability constraints. As an important direction for future work, channel switching delays can be considered when determining the optimal channel switching and randomization strategies.

Chapter 3

Channel Switching Strategies based on UCB Bandit Algorithm for Throughput Maximization with Unknown Channel Statistics

Wireless communication systems are constantly evolving, leading to an increasing demand for efficient spectrum utilization. To address the challenges of limited spectrum resources, methods such as time sharing and parameter randomization have been proposed to enhance channel capacity and throughput. However, in practical scenarios, channel coefficients and statistics are often unknown, making it difficult to optimize spectrum utilization. In this context, we present a novel approach based on the upper confidence bound (UCB) bandit algorithm for channel selection and channel switching in block fading channels. Our proposed methods consider the unknown channel statistics at the transmitter and aim to maximize the number of correctly received symbols per unit of time. We investigate two scenarios: one without channel switching within blocks and another with channel switching. For the former, we propose a UCB bandit algorithm to select the best channel among a set of possibilities, while for the latter, we introduce virtual channels and employ a UCB bandit algorithm to determine the optimal channel switching strategy. Additionally, we propose a low-complexity version of the algorithm to ensure efficient convergence when dealing with a large number of virtual channels. Theoretical limits are provided for comparison, and simulation results demonstrate the effectiveness of the proposed UCB bandit algorithms in achieving near-optimal performance and leveraging the benefits of channel switching. Our work contributes to the field by addressing the challenges of unknown channel statistics and offering efficient solutions for maximizing throughput in wireless communication systems.

3.1 System Model

Suppose that a transmitter and a receiver, each equipped with a single RF chain, communicate over $K \geq 2$ available channels (frequency bands) by performing channel switching [4]. During channel switching, only one channel is employed for communication at a given time, and the transmitter and the receiver switch among the available channels in a coordinated fashion. The aim is to perform optimal channel switching in order to maximize the average number of correctly received symbols [14].

The channels are modeled as flat-fading and additive white Gaussian noise channels, where σ_i^2 represents the spectral density level of the zero-mean Gaussian noise for the ith channel between the transmitter and the receiver. Also, g_i denotes the channel gain (that is, the magnitude square of the channel coefficient) related to channel i for $i \in \{1, \ldots, K\}$, and $g_1, \ldots, g_K$ are modeled as independent random variables. Let $p_{g_i}(\cdot)$ denote the probability density function (PDF) of the channel gain for the ith channel. A block fading scenario is assumed for the channels in the sense that g_i remains constant over each block (corresponding to a large number of symbols) and then changes to a different value independently according to the fading distribution. It is assumed that the channels have a common block interval denoted by T_b and fixed channel gains are observed over an interval of $[iT_b, (i+1)T_b]$ for each integer i. Regarding the information about the channel gains at the transmitter and the receiver, it is assumed that the CSI of each channel is available at the receiver (i.e., the receiver performs channel estimation perfectly) but the transmitter does not know the CSI at the time of transmission. However, the receiver informs the transmitter about the channel gains (equivalently, the SNRs) via feedback. In particular, the transmitter learns about the channel gain(s) at the end of each block, and can utilize this information for developing a strategy for transmission in upcoming blocks. (The transmitter cannot utilize the CSI of a block for the transmission during that block since it does not have that information at the time of transmission.)

For evaluating the performance, we consider the number of symbols that are correctly received per unit of time. Suppose that channel i is selected and a transmit power of P is employed for transmission through that channel. In that case, the outage occurs when

$$\frac{Pg_i}{\sigma_i^2} < \gamma_{\min} \tag{3.1}$$

where $\gamma_{\min}$ represents the minimum SNR for non-outage [13]. When the system is not in outage, the symbols are received correctly with the following probability:

$$1 - a\,Q\left(\sqrt{bPg_i/\sigma_i^2}\right) \tag{3.2}$$

where Q represents the Q-function, and a and b are parameters that depend on the modulation type [13]. Therefore, the number of symbols that are correctly received is proportional to

$$\mathbb{I}\left(Pg_i/\sigma_i^2 \geq \gamma_{\min}\right)\left(1 - a\,Q\left(\sqrt{bPg_i/\sigma_i^2}\right)\right) \tag{3.3}$$

where $\mathbb{I}$ denotes the indicator function.

Consider a number of fading blocks denoted by N_{b}. The aim is to maximize the average number of correctly received symbols per unit of time over the possible strategies, which can be formulated via (3.3) as

$$\max_{\{\lambda_k^{(j)}, P_k^{(j)}\}_{j=1,k=1}^{N_b,K}} \frac{1}{N_b}\sum_{j=1}^{N_b}\sum_{k=1}^{K}\lambda_k^{(j)}\mathbb{I}\left(P_k^{(j)}g_k^{(j)}/\sigma_k^2 \geq \gamma_{\min}\right)\left(1 - a\,Q\left(\sqrt{bP_k^{(j)}g_k^{(j)}/\sigma_k^2}\right)\right) \tag{3.4}$$

$$\text{subject to} \sum_{k=1}^{K}\lambda_k^{(j)}P_k^{(j)} \leq P_{\text{av}}\,, \ \sum_{k=1}^{K}\lambda_k^{(j)} = 1\,,\ \forall j \tag{3.5}$$

$$\lambda_k^{(j)} \geq 0\,,\ P_k^{(j)} \in [0, P_{\text{pk}}]\,,\ \forall j, k \tag{3.6}$$

where $P_k^{(j)}$ is the power level employed for channel k during the jth block, $\lambda_k^{(j)}$ is the time sharing factor corresponding to channel k for the jth block, P_{av} and P_{pk} specify the average and peak power limits, respectively (with $P_{\text{av}} < P_{\text{pk}}$), and $g_k^{(j)}$ represents the realization of the channel gain for channel k during the jth block (i.e., a realization generated according to the PDF $p_{g_k}(\cdot)$ during block j). It is assumed that realizations are independent over different blocks.

For solving the optimization problem given in (3.4)–(3.6), two different approaches can be considered:

Approach 1: Consider the use of a fixed power level with no channel switching in a given block; that is, $P_k^{(j)} = P_{\text{av}}$ for all $j \in \{1, \ldots, N_{\text{b}}\}$ and $k \in \{1, \ldots, K\}$. Then, the problem in (3.4)–(3.6) reduces to

$$\max_{\{s(j)\}_{j=1}^{N_b}} \frac{1}{N_b}\sum_{j=1}^{N_b}\mathbb{I}\left(P_{\text{av}}g_{s(j)}^{(j)}/\sigma_{s(j)}^2 \geq \gamma_{\min}\right)\left(1 - a\,Q\left(\sqrt{bP_{\text{av}}g_{s(j)}^{(j)}/\sigma_{s(j)}^2}\right)\right) \tag{3.7}$$

where $s(j) \in \{1, \ldots, K\}$ denotes the channel selected for transmission during the jth block. This approach corresponds to the *optimal single-channel strategy* since only one channel is employed (at the average power limit) in each block.

Approach 2: In this approach, it is allowed to perform channel switching in each block by using different power levels with certain time sharing factors. As shown in the literature [4], [14], [7] (also see Section 3.3), channel switching between two channels is sufficient to maximize the performance under an average power constraint. Therefore, we consider the selection of at most two different channels in each block and formulate the following simplified version of (3.4)–(3.6) in Approach 2:

$$\max_{\{s_\ell(j),P_\ell^{(j)},\tilde{\lambda}_\ell^{(j)}\}_{j=1,\ell=1}^{N_b,2}} \frac{1}{N_b}\sum_{j=1}^{N_b}\sum_{\ell=1}^{2}\tilde{\lambda}_\ell^{(j)}\mathbb{I}\left(\frac{P_\ell^{(j)}g_{s_\ell(j)}^{(j)}}{\sigma_{s_\ell(j)}^2}\geq\gamma_{\min}\right) \times\left(1-a\,Q\left(\sqrt{bP_\ell^{(j)}g_{s_\ell(j)}^{(j)}/\sigma_{s_\ell(j)}^2}\right)\right) \tag{3.8}$$

$$\text{subject to } \sum_{\ell=1}^{2}\tilde{\lambda}_\ell^{(j)}P_\ell^{(j)}\leq P_{\text{av}},\ \sum_{\ell=1}^{2}\tilde{\lambda}_\ell^{(j)}=1,\ \forall j \tag{3.9}$$

$$\tilde{\lambda}_\ell^{(j)}\geq 0\,,\ P_\ell^{(j)}\in[0,P_{\text{pk}}]\,,\ \forall j,\ell \tag{3.10}$$

In (3.8)–(3.10), $s_1(j)$ and $s_2(j)$ represent the channel indices selected for transmission during the jth block, $P_1^{(j)}$ and $P_2^{(j)}$ correspond to the power levels employed for the selected channels during the jth block, and $\tilde{\lambda}_1^{(j)}$ and $\tilde{\lambda}_2^{(j)}$ denote the time sharing (channel switching) factors for the jth block. In other words, during the jth block, channel $s_1(j)$ is used with power level $P_1^{(j)}$ and time sharing factor $\tilde{\lambda}_1^{(j)}$, and channel $s_2(j)$ is employed with power level $P_2^{(j)}$ and time sharing factor $\tilde{\lambda}_2^{(j)}$ (i.e., $1-\tilde{\lambda}_1^{(j)}$). This approach (Approach 2) is called the *optimal channel switching strategy* in the remainder of the manuscript.

Remark 1: Since the transmitter does not know the channel gains during the transmission of a block (learns it at the end of the block via feedback),[1] it is not possible for the transmitter to solve the problems in (3.4)–(3.10) directly. The transmitter can evaluate the objective functions after choosing its strategy. For example, in Approach 1, at the beginning of the kth block, the transmitter is able to calculate

$$\sum_{j=1}^{k-1}\mathbb{I}\left(P_{\text{av}}g_{s(j)}^{(j)}/\sigma_{s(j)}^2\geq\gamma_{\min}\right)\left(1-a\,Q\left(\sqrt{bP_{\text{av}}g_{s(j)}^{(j)}/\sigma_{s(j)}^2}\right)\right) \tag{3.11}$$

Therefore, the transmitter needs to make sequential decisions. Accordingly, a learning based approach is proposed in the following.

[1]When only one channel is used in a block, the receiver estimates the SNR of that channel and sends it to the transmitter as a feedback at the end of the block. When channel switching is performed between two channels in a block (as in (3.8)–(3.10)), the receiver estimates the SNR for each channel and sends them to the receiver via feedback.

3.2 Bandit Formulation: Unknown Channel Statistics at Transmitter

In this section, we propose a learning based algorithm to solve the problems formulated in (3.7) and (3.8)–(3.10). Our aim is to maximize the rate of correctly received symbols in the communication system by performing channel switching when the channel statistics are not known at the transmitter at the time of transmission. In this situation, a well-suited approach to achieve the maximum rate of correctly received symbols by balancing exploration and exploitation of channels is the multi-armed bandit (MAB) algorithm [42]. Therefore, we formulate the proposed channel switching problem as an MAB problem having non-stationary rewards, in which online policy update can be performed. In particular, we apply the upper confidence bounds (UCB) MAB algorithm [43], which converges to the best channel selection/switching strategy and provides a high rate of correctly received symbols. The details are presented in the following sections.

3.2.1 Computing the Reward

Defining a suitable reward function is crucial to fulfill the objective of the proposed problems in (3.7) and (3.8)–(3.10) when the channel statistics are not known at the transmitter at the time of transmission. In general, a reward function provides a measure of performance, which is defined by the state and the action taken at a given state. In the considered setting, after selecting a particular strategy during the jth block, a reward is obtained depending on the quality of communication through that channel. We utilize the number of correctly received symbols as a quality metric due to its practical importance for communication systems. Accordingly, for the problem in (3.7), that is, for Approach 1, the reward function for the jth block is defined as

$$\tilde{r}_j(s(j)) = \begin{cases} 1 - a\,Q\left(\sqrt{bP_{\mathrm{av}} g_{s(j)}^{(j)}/\sigma_{s(j)}^2}\right), & \text{if } \frac{P_{\mathrm{av}} g_{s(j)}^{(j)}}{\sigma_{s(j)}^2} \geq \gamma_{\min} \\ 0\,, & \text{otherwise} \end{cases} \tag{3.12}$$

Namely, the reward is zero if there occurs outage and is given by the probability of correct decision (i.e., the rate of correctly received symbols) otherwise.

For Approach 2 specified by (3.8)–(3.10), the reward function for the jth block is given by

$$r_j\big(s_1(j), s_2(j), P_1^{(j)}, P_2^{(j)}, \tilde{\lambda}_1^{(j)}, \tilde{\lambda}_2^{(j)}\big) = \tag{3.13}$$
$$\begin{cases} \sum_{\ell=1}^{2} \tilde{\lambda}_\ell^{(j)} \left(1 - a\,Q\left(\sqrt{\frac{bP_\ell^{(j)} g_{s_\ell(j)}^{(j)}}{\sigma_{s_\ell(j)}^2}}\right)\right), & \text{if } \frac{P_\ell^{(j)} g_{s_\ell(j)}^{(j)}}{\sigma_{s_\ell(j)}^2} \geq \gamma_{\min} \\ 0\,, & \text{otherwise} \end{cases}$$

In this case, the reward is determined by the selected channels, the employed power levels, and the time sharing factors.

3.2.2 Proposed Strategies based on UCB Bandit Algorithm

To learn and adopt the policy which maximizes the total rewards defined in the preceding section, we apply the UCB bandit algorithm proposed by Lai and Robbins [43]. The UCB bandit algorithm is an advanced version of the greedy and epsilon-greedy algorithms, which is designed to cope with non-stationary environments. The UCB bandit algorithm is relies on the fact that there is a need to explore all feasible channels, but at the same time, it is equally important to exploit the best channels so that the maximum total reward can be achieved.

For **Approach 1** specified by the problem in (3.7), all the K channels operating at the average power limit P_{av} are considered as input bandits. The reward $\tilde{r}_j(s(j))$ for each block of transmission is calculated via (3.12) based on the selected channel $s(j) \in \{1, \ldots, K\}$ over that block. In this context, the action is to select a particular channel $s(j)$ during the jth block of transmission. The number of times a channel k is selected prior to the jth block of transmission is represented by

$$N_k^{(j)} = \sum_{l=1}^{j-1} \mathbb{I}\left(s(l) = k\right). \tag{3.14}$$

In contrast to the traditional greedy and epsilon-greedy algorithms, the UCB bandit algorithm balances the trade-off between exploration and exploitation after gathering information about the environment. At first, the algorithm just focuses on exploring all the available channels/bandits, and then it exploits the channel which gives the highest reward while exploring the other channels at certain times. The action taken by UCB for the jth block is [44]

$$s(j) = \underset{k \in \{1,\dots,K\}}{\arg\max} \left(H_k^{(j)} + \varrho \sqrt{\frac{\log j}{N_k^{(j)}}} \right) \tag{3.15}$$

where the extent of exploration is controlled by the confidence factor ϱ, and $H_k^{(j)}$ is the average (estimated) reward of selecting channel k before the transmission in the jth block. That is,

$$H_k^{(j)} = \frac{1}{N_k^{(j)}} \sum_{l=1}^{j-1} \tilde{r}_l(s(l))\, \mathbb{I}\left(s(l) = k\right) \tag{3.16}$$

The UCB bandit formula in (3.15) consists of two parts; the first part, $H_k^{(j)}$, is related to the exploitation while the second part is responsible for the exploration. The principle behind exploitation is 'optimism in the face of uncertainty', which means that if we do not know which channel is the best, then we choose the one which is currently giving the highest payoff. On the other hand, the exploration in the UCB bandit algorithm is specified by the second part of (3.15). The extent of exploration can be controlled by choosing different values of the hyper-parameter ϱ. A larger value of ϱ corresponds to more extensive exploration. The measure of uncertainty for a particular channel is also specified by this second part of the equation. Namely, $N_k^{(j)}$ becomes zero or exceedingly small if the channel has not been selected or selected very few times, respectively. As a result, this uncertainty term becomes large, assigning a large value to that action, which makes it more likely to be selected for the next time. Every time we select a channel k, $N_k^{(j)}$ increases, which in turn leads to a reduction in the uncertainty term. When $N_k^{(j)}$

increases, we make that action less likely to be selected because of exploration, but still, there are chances that a particular channel gets selected due to the high value of the exploitation term.

Due to the log function in the formulation in (3.15), when a particular channel is not selected, the uncertainty term increases at a slow pace. On other hand, the uncertainty term decreases quickly when a channel is being selected, due to the increase in $N_k^{(j)}$. Therefore, infrequently selected channels have a high value of the exploration term due to high uncertainty. The exploration term gradually decreases with the increase in the number of transmission blocks; then the channel is selected based on the exploitation term.

Based on the preceding descriptions, the pseudo-code of the proposed learning based channel selection algorithm for Approach 1 can be found in Algorithm 1. It is noted that the algorithm first explores all the K channels in the first K blocks, and then performs the channel selection based on (3.15) by considering the exploration and exploitation trade-off. The numerical results in Section 3.4 show that the proposed channel selection strategy for Approach 1 (i.e., Algorithm 1) gets very close to the optimal single-channel selection solution in the presence of known channel statistics at the transmitter after proper exploration.

For the *optimal channel switching strategy* referred to as **Approach 2** and specified by (3.8)–(3.10), the single channels operating at the average power limit cannot be considered as the only input bandits since channel switching between two channels at different power levels and with different time sharing factors can outperform the single channel strategy in terms of the rate of correctly received symbols as proposed in Section 3.1. For the purpose of defining suitable bandits for Approach 2, we generate *virtual channels* by considering all possible pairwise combinations of the K channels with discretized power levels and time sharing factors, as detailed in the following.

In the first step of virtual channel generation, $M = \binom{K}{2}$ possible combinations of K channels are generated. Let $\mathcal{S}_{\rm c}$ represent the list (set) of all channel pairs;

Algorithm 1 UCB Bandit Algorithm for Approach 1

Input:
  K: **number of channels (bandits)**
  N_b: **total number of blocks** ▷ $N_b > K$
  P_{av}: **average power limit**
  $\sigma_1^2, \ldots, \sigma_K^2$: **noise spectral density levels**
for $j \geq 1$ & $j \leq K$ **do**
    Select Action $s(j) = j$ ▷ select every bandit once
    Calculate reward $\tilde{r}_j(s(j))$ **using** (3.12)
    Update $N_k^{(j)}$ **using** (3.14), **where** $k = s(j)$
    Update $H_k^{(j)}$ **using** (3.16), **where** $k = s(j)$
end for
for $j \geq K+1$ & $j \leq N_b$ **do**
    Select Action $s(j) = \underset{k \in \{1,\ldots,K\}}{\arg\max} \left(H_k^{(j)} + \varrho \sqrt{\frac{\log j}{N_k^{(j)}}} \right)$
    Calculate reward $\tilde{r}_j(s(j))$ **using** (3.12)
    Update $N_k^{(j)}$ **using** (3.14), **where** $k = s(j)$
    Update $H_k^{(j)}$ **using** (3.16), **where** $k = s(j)$
end for

that is,

$$\mathcal{S}_{\text{c}} = \{(1,2), (1,3), \ldots, (K-1, K)\} \tag{3.17}$$

where $|\mathcal{S}_{\text{c}}| = M$. Then, in the second step, the power level parameter $P \in [0, P_{\text{pk}}]$ and the time sharing factor $\tilde{\lambda} \in [0,1]$ are discretized using the discretization parameter Δ. Namely, P and $\tilde{\lambda}$ are sampled uniformly with a step size of Δ starting from zero. Accordingly,

$$\zeta_p = \left\lfloor \frac{P_{\text{pk}}}{\Delta} + 1 \right\rfloor \tag{3.18}$$

power levels and

$$\zeta_\lambda = \left\lfloor \frac{1}{\Delta} + 1 \right\rfloor \tag{3.19}$$

time sharing factors are considered. For each channel pair in set $\mathcal{S}_{\text{c}}$ in (3.17), the power levels for the first channel denoted by P_1 and those for the second channel denoted by P_2 are determined as follows:

- P_1 is set to the discretized power levels and the corresponding P_2 values are found from

$$P_2 = \frac{P_{\text{av}} - \tilde{\lambda} P_1}{1 - \tilde{\lambda}} \tag{3.20}$$

for all discretized values of $\tilde{\lambda}$. Namely, $\zeta_p \zeta_\lambda$ different (P_1, P_2) pairs are generated, where ζ_p and ζ_λ are as in (3.18) and (3.19), respectively.

- P_2 is set to the discretized power levels and the corresponding P_1 values are found from

$$P_1 = \frac{P_{\text{av}} - P_2(1 - \tilde{\lambda})}{\tilde{\lambda}} \tag{3.21}$$

for all discretized values of $\tilde{\lambda}$. Again, $\zeta_p \zeta_\lambda$ different (P_1, P_2) pairs are generated.

In this way, all the combinations of the power levels and time sharing factors are considered by utilizing the fact that an optimal solution must operate at the power limit; i.e., $\tilde{\lambda} P_1 + (1 - \tilde{\lambda}) P_2 = P_{\text{av}}$ (as the objective function is a monotone increasing function of the power parameter). Overall, $2\zeta_p \zeta_\lambda$ different (P_1, P_2) pairs are generated. However, some of these power level pairs can violate the peak power constraint; hence, they are discarded. Let N_P denote the number of power level pairs that satisfy the peak power constraint. We can represent these feasible power level pairs and their corresponding time sharing factors as $\big(P_1(i), P_2(i), \tilde{\lambda}(i), 1 - \tilde{\lambda}(i)\big)$ for $i = 1, \ldots, N_P$. Since there are $M = \binom{K}{2}$ channel pairs as defined in (3.17), there exist MN_P virtual channels (equivalently, bandits), which can be stated as follows:

$$\begin{aligned} \mathcal{S}_V = \big\{ &\big(\mathcal{S}_{\text{c}}(m,1), \mathcal{S}_{\text{c}}(m,2), P_1(i), P_2(i), \tilde{\lambda}(i), 1 - \tilde{\lambda}(i)\big) \\ &\text{for } m = 1, \ldots, M, \ i = 1, \ldots, N_P \big\} \end{aligned} \tag{3.22}$$

where $\mathcal{S}_{\text{c}}(m,1)$ and $\mathcal{S}_{\text{c}}(m,2)$ represent the first and second channel indices, respectively, in the m channel pair in set $\mathcal{S}_{\text{c}}$.

The virtual channels in set $\mathcal{S}_V$ in (3.22) serve as bandits for Approach 2. For example, when the jth virtual channel is selected from set $\mathcal{S}_V$, which is denoted by $\mathcal{S}_V(j)$, the corresponding reward can be calculated from (3.13). By updating the set of channels in Algorithm 1, we can employ it for obtaining the solution of

Approach 2, as well. Namely, instead of defining the set of channels/bandits as $\{1, \ldots, K\}$ as in Approach 1, we consider the virtual channels in set $\mathcal{S}_V$ in (3.22) as the possible strategies (bandits) in Approach 2. After this step, the UCB bandit algorithm is implemented in its conventional manner to determine the optimal strategy. The proposed strategy for Approach 2 is presented in Algorithm 2. As noted from the algorithm, the set of virtual channels is generated first. Then, they are explored in the first MN_P blocks. After that, the UCB bandit algorithm is implemented for achieving a trade-off between exploration and exploitation.

Algorithm 2 UCB Bandit Algorithm for Approach 2

Input:
  K: **number of channels**
  N_b: **total number of blocks**
  P_{av}: **average power limit**
  P_{pk}: **peak power limit**
  Δ: **discretization parameter**
  $\sigma_1^2, \ldots, \sigma_K^2$: **noise spectral density levels**
Generate power level pairs based on (3.18)–(3.21)
Choose feasible power level pairs according to P_{pk}
Generate set $\mathcal{S}_V$ **of virtual channels (bandits) in** (3.22)
$\triangleright$ $|\mathcal{S}_V| = MN_P$ and $N_b > MN_P$
for $j \geq 1$ & $j \leq MN_P$ **do**
  Select Action $\mathcal{A} = \mathcal{S}_V(j)$ $\triangleright$ select every bandit once
  Calculate reward $r_j(\mathcal{A})$ **using** (3.13)
  Update $N_k^{(j)}$ **using** (3.14), **where** $k = \mathcal{A}$
  Update $H_k^{(j)}$ **using** (3.16), **where** $k = \mathcal{A}$
end for
for $j \geq MN_P + 1$ & $j \leq N_b$ **do**
  Select Action $\mathcal{A} = \underset{k \in \{1,\ldots,MN_P\}}{\arg\max} \left(H_k^{(j)} + \varrho \sqrt{\frac{\log j}{N_k^{(j)}}} \right)$
  Calculate reward $r_j(\mathcal{A})$ **using** (3.13)
  Update $N_k^{(j)}$ **using** (3.14), **where** $k = \mathcal{A}$
  Update $H_k^{(j)}$ **using** (3.16), **where** $k = \mathcal{A}$
end for

3.2.3 UCB Bandit Algorithm with Channel Elimination (UCB-CE)

The number of bandits is a key factor that plays a significant role in the performance and complexity of the proposed algorithms since there is a direct relation between the number of bandits and the number of iterations (blocks) required for achieving close-to-optimal performance. Namely, as the number of bandits increases, the algorithms need more iterations to determine the optimal channel selection or switching strategy by properly learning the channel statistics. In Approach 1, as the number of bandits is equal to the number of channels, it is usually not a very high number; hence, complexity is not a critical issue in that scenario. However, for Approach 2, the number of bandits can be very high since it is equal to the number of virtual channels, which are generated by considering all possible channel pairs with various power levels and time sharing factors. For such cases, we propose a modified version of the UCB bandit algorithm, named the *UCB bandit with channel elimination (UCB-CE)* algorithm, in order to achieve close to optimal performance with fewer iterations and lower complexity. The main idea behind the UCB-CE algorithm is to maximize the cumulative reward while sequentially removing some of the bandits (virtual channels). The sequential removal of bandits is performed based on the average rewards they yield over certain intervals. Namely, the bandits that result in low gains are eliminated after sequentially observing their behavior.

To describe the proposed UCB-CE algorithm in more detail, let $\bar{B}$ denote the total number of bandits and let the total number of blocks, N_b, be divided into N_{int} intervals of equal length Γ, resulting in Γ blocks in each interval. (If N_b is not an integer multiple of N_{int}, the last interval can have fewer blocks.) We also introduce an elimination factor parameter denoted by $\chi \in (0,1]$, and an elimination threshold parameter represented by η, which is a positive integer. Then, at the end of the first interval, the number of bandits is reduced from $\bar{B}$ to $\text{round}(\chi\bar{B})$ if $\text{round}(\chi\bar{B}) \geq \eta$; otherwise, it is kept as $\bar{B}$. Let $\bar{B}_1$ denote the number of bandits at the end of the first interval. Next, at the end of the second interval, the number of bandits is reduced from $\bar{B}_1$ to $\text{round}(\chi\bar{B}_1)$ if $\text{round}(\chi\bar{B}_1) \geq \eta$;

otherwise, it is kept as $\bar{B}_1$. Similarly, at the end of the third interval, the number of bandits is reduced from $\bar{B}_2$ to round$(\chi\bar{B}_2)$ if round$(\chi\bar{B}_2) \geq \eta$; otherwise, it is kept as $\bar{B}_2$, where $\bar{B}_2$ is the number of bandits at the end of the second interval. The elimination procedure continues in this manner until the end of the last interval; that is, as long as the number of bandits does not stay below η, $(1-\chi)$ fraction of them are eliminated after each interval. The bandits to be eliminated are determined according to the average rewards in (3.16). Let $\mathcal{Z}_{\mathrm{a}}$ denote the set of bandits, which initially consists of all the $\bar{B}$ bandits. First, as usual, all the bandits are selected once for exploration, and the corresponding rewards are calculated (assuming $\bar{B} < N_b$). Then, the UCB algorithm is implemented in the conventional manner until the Γ-th block. When the block index j becomes equal to Γ (i.e., at the end of the first interval), the bandits with the lowest round$((1-\chi)\bar{B})$ average rewards in set $\left\{H_k^{(j)},\ \text{for } k \in \mathcal{Z}_{\mathrm{a}}\right\}$ are eliminated from set $\mathcal{Z}_{\mathrm{a}}$ if round$(\chi\bar{B}) \geq \eta$; otherwise, set $\mathcal{Z}_{\mathrm{a}}$ remains unchanged.

Then, the UCB algorithm is implemented based on the updated set $\mathcal{Z}_{\mathrm{a}}$ of bandits until the end of the second interval. When the block index j becomes equal to 2Γ (i.e., at the end of the second interval), again the bandits with the lowest round$((1-\chi)|\mathcal{Z}_{\mathrm{a}}|)$ average rewards in set $\left\{H_k^{(j)},\ \text{for } k \in \mathcal{Z}_{\mathrm{a}}\right\}$ are eliminated from set $\mathcal{Z}_{\mathrm{a}}$ if round$(\chi|\mathcal{Z}_{\mathrm{a}}|) \geq \eta$; otherwise, set $\mathcal{Z}_{\mathrm{a}}$ remains unchanged. (Here, $|\mathcal{Z}_{\mathrm{a}}|$ represents the number of bandits.) The procedure continues in this fashion until the end of the last interval. The detailed implementation of the UCB-CE algorithm is presented in Algorithm 3 by considering Approach 2.

Algorithm 3 UCB Bandit with Channel Elimination (UCB-CE)

Input:
- K: **number of channels**
- N_b: **total number of blocks**
- P_{av}: **average power limit**
- P_{pk}: **peak power limit**
- Δ: **discretization parameter**
- $\sigma_1^2, \ldots, \sigma_K^2$: **noise spectral density levels**
- N_{int}: **number of intervals**
- χ: **elimination factor**
- η: **elimination threshold**

Generate power level pairs based on (3.18)–(3.21)

Choose feasible power level pairs according to P_{pk}

Generate set $\mathcal{S}_V$ **of virtual channels (bandits) in** (3.22)

$\triangleright$ $|\mathcal{S}_V| = MN_P$ and $N_b > MN_P$

Initialization:
- $\bar{B} = m = MN_p$, **total number of bandits initially** ($\bar{B} < N_b$)
- $\Gamma = \frac{N_b}{N_{\text{int}}}$, **number of blocks per interval (assume integer)**
- $\mathcal{Z}_{\text{a}} = \mathcal{S}_V$, **initializing available set of bandits**

Algorithm 3 (continued): UCB Bandit with Channel Elimination (UCB-CE)

if $MN_p \leq \max\{\eta, \mathrm{round}(\chi^{N_{\mathrm{int}}-1}\bar{B})\}$ **then**
  Run Algorithm 2 ▷ no need for bandit elimination
else
  for $j \geq 1$ & $j \leq \bar{B}$ **do**
    Select Action $\mathcal{A} = \mathcal{Z}_{\mathrm{a}}(j)$ ▷ select each bandit once
    Calculate reward $r_j(\mathcal{A})$ **using** (3.13)
    Update $N_k^{(j)}$ **using** (3.14), **where** $k = \mathcal{A}$
    Update $H_k^{(j)}$ **using** (3.16), **where** $k = \mathcal{A}$
  end for
  for $j \geq \bar{B}+1$ & $j \leq N_b$ **do**
    Select Action $\mathcal{A} = \arg\max_{k \in \mathcal{Z}_{\mathrm{a}}} \left(H_k^{(j)} + \varrho\sqrt{\frac{\log j}{N_k^{(j)}}} \right)$
    Calculate reward $r_j(\mathcal{A})$ **using** (3.13)
    Update $N_k^{(j)}$ **using** (3.14), **where** $k = \mathcal{A}$
    Update $H_k^{(j)}$ **using** (3.16), **where** $k = \mathcal{A}$
    $Counter+ = 1$
    if $Counter == \Gamma$ **then**
      if $\mathrm{round}(\chi m) \geq \eta$ **then**
        Determine the worst $\mathrm{round}((1-\chi)m)$ **bandits based on the set** $\left\{H_k^{(j)},\ \text{for } k \in \mathcal{Z}_{\mathrm{a}}\right\}$
        Remove the worst bandits from set $\mathcal{Z}_{\mathrm{a}}$
        $m = length(\mathcal{Z}_{\mathrm{a}})$ ▷ Update # bandits
      end if
      $Counter = 0$
    end if
  end for
end if

3.3 Performance Benchmarks: Known Channel Statistics at Transmitter

For comparison purposes, we also consider the scenario in which the channel statistics are known at the transmitter (the receiver has again the CSI of all the channels). Namely, the transmitter knows the PDFs of the channel gains, $p_{g_1}(\cdot), \ldots, p_{g_K}(\cdot)$. In this case, the following optimization problems can be considered.

3.3.1 Optimal Single-Channel Approach

When the channel statistics are known, the transmitter can calculate the average number of correctly received symbols for each channel and choose the best channel accordingly. This corresponds to the optimal single channel approach with known channel statistics, which is formulated as follows:

$$\max_{i\in\{1,\dots,K\}} \int_{\frac{\sigma_i^2\gamma_{\min}}{P_{\mathrm{av}}}}^{\infty} p_{g_i}(g)\left(1-a\,Q\left(\sqrt{bP_{\mathrm{av}}g/\sigma_i^2}\right)\right)dg \tag{3.23}$$

3.3.2 Optimal Channel Switching Approach

In the presence of channel switching and known channel statistics, the optimal channel switching strategy can be obtained from the following problem:

$$\max_{\{\lambda_i,P_i\}_{i=1}^K} \sum_{i=1}^{K}\lambda_i \int_{\frac{\sigma_i^2\gamma_{\min}}{P_i}}^{\infty} p_{g_i}(g)\left(1-a\,Q\left(\sqrt{\frac{bP_ig}{\sigma_i^2}}\right)\right)dg \tag{3.24}$$

$$\text{subject to } \sum_{i=1}^{K}\lambda_iP_i \leq P_{\mathrm{av}}\,, \quad P_i\in[0,P_{\mathrm{pk}}]\,,\ \forall i \tag{3.25}$$

$$\sum_{i=1}^{K}\lambda_i = 1\,, \quad \lambda_i\geq 0\,,\ \forall i \tag{3.26}$$

If we define

$$h_i(P_i) \triangleq \int_{\frac{\sigma_i^2\gamma_{\min}}{P_i}}^{\infty} p_{g_i}(g)\left(1-a\,Q\left(\sqrt{\frac{bP_ig}{\sigma_i^2}}\right)\right)dg \tag{3.27}$$

then (3.24)–(3.26) can be stated as

$$\max_{\{\lambda_i,P_i\}_{i=1}^K} \sum_{i=1}^{K} \lambda_i h_i(P_i) \tag{3.28}$$

$$\text{subject to } \sum_{i=1}^{K} \lambda_i P_i \le P_{\rm av}\,, \quad P_i \in [0, P_{\rm pk}]\,,\ \forall i \tag{3.29}$$

$$\sum_{i=1}^{K} \lambda_i = 1\,, \quad \lambda_i \ge 0\,,\ \forall i \tag{3.30}$$

The optimization problems in the form of (3.28)–(3.30) have been investigated in many studies such as [4, Eq. (2)], where it is shown that an optimal strategy can be implemented by channel switching between at most two different channels. Accordingly, the optimal solution can be obtained from the solution of

$$\max_{\{\lambda,P_1,P_2\}} \lambda\, h_{\max}(P_1) + (1-\lambda) h_{\max}(P_2) \tag{3.31a}$$

$$\text{subject to } \lambda P_1 + (1-\lambda) P_2 \le P_{\rm av}\,, \tag{3.31b}$$

$$P_1 \in [0, P_{\rm pk}]\,,\ P_2 \in [0, P_{\rm pk}]\,,\ \lambda \in [0,1] \tag{3.31c}$$

where $h_{\max}(P) \triangleq \max\{h_1(P), \ldots, h_K(P)\}$ [4, Sec. III-B].

As the presence of channel statistics is assumed in this section, the problems in (3.23) and (3.31) will be used to obtain theoretical limits on the performance of the proposed learning based algorithms in Section 3.2.

3.4 Simulation Results

In this section, we present numerical examples to investigate the performance of the proposed approaches. We consider $K = 2$ channels with noise level parameters being given by $\sigma_1^2 = 1$ and $\sigma_2^2 = 5$. Also, $P_{\rm av} = 7$, $P_{\rm pk} = 10$, and $\gamma_{\min} = 1$ are employed. The binary phase shift keying (BPSK) modulation is assumed, corresponding to parameters $a = b = 1$ in (3.2). In addition, the channels are modeled as independent Nakagami fading channels; that is, the gains are Nakagami distributed with parameters $m_1 = 0.5$ and $\Omega_1 = 1.5$ for

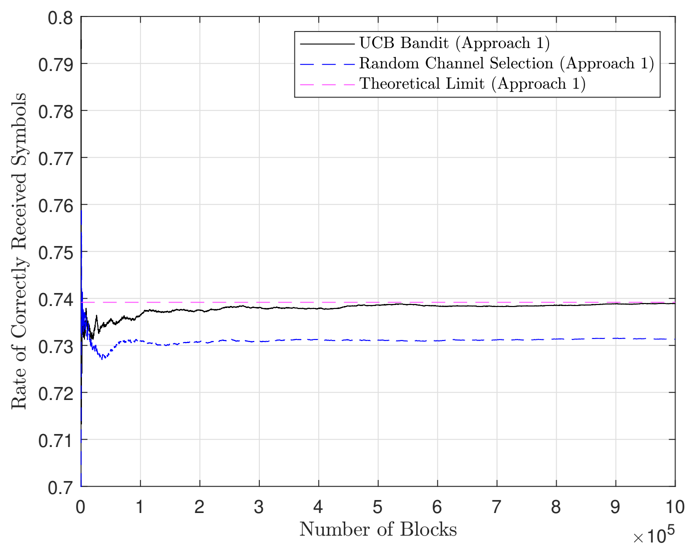

Figure 3.1: Rate of correctly received symbols achieved by the proposed UCB bandit approach (Algorithm 1), the random channel selection strategy, and the theoretical limit (assuming known channel statistics at the transmitter) for the single channel approach.

channel 1, and $m_2 = 10$ and $\Omega_2 = 1$ for channel 2. For the UCB-CE algorithm, the number of intervals is $N_{\text{int}} = 10$ and the number of required channels at the end is specified by $\chi = 10$.

In the simulations, we consider both the optimal single channel strategy corresponding to **Approach 1** and the optimal channel switching strategy referred to as **Approach 2**. For each strategy, we evaluate not only the performance of the proposed UCB bandit algorithm but also the theoretical performance limit in the presence of known channel statistics at the transmitter; namely, the solution of (3.23) for the single channel strategy and that of (3.31) for the channel switching strategy. For comparison purposes, we also assess the performance of the random channel selection approach, in which a channel is selected randomly in each block.

First, we focus on the single channel strategy, in which there exist two bandits corresponding to the number of channels ($K = 2$). The aim is to determine the optimal channel that yields the highest number of correctly received symbols

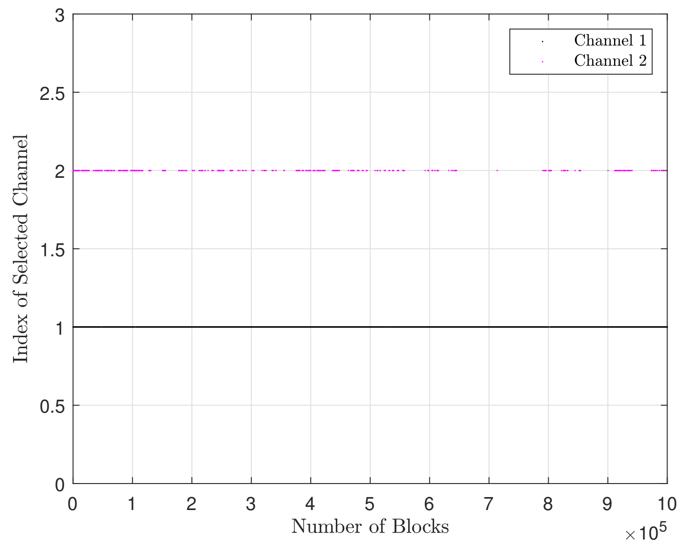

Figure 3.2: Indices of selected channels versus the block index for the proposed UCB bandit algorithm for Approach 1 (Algorithm 1).

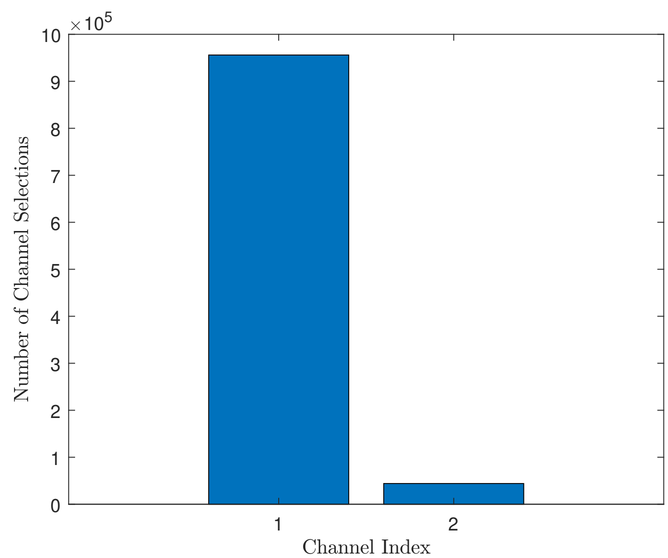

Figure 3.3: Number of times channels are selected by the proposed UCB bandit algorithm for Approach 1 (Algorithm 1).

by visiting the available channels. A total of 10^6 blocks are transmitted in the bandit problem, meaning that the channels are visited 10^6 times in total. In this scenario, the solution of (3.23) yields the theoretical limit. In Figure. 3.1, the rate of correctly received symbols is plotted against the block index for the proposed UCB bandit algorithm (Algorithm 1), the random channel selection approach, and the theoretical limit when the channel statistics are known at the transmitter. From (3.23), it is obtained that the optimal channel is channel 1, which yields 0.7392 for the maximum rate of correctly received symbols. Also, it is noted that the proposed UCB bandit algorithm is able to learn the best channel and converge to the optimal solution yielding 0.7379 for the rate of correctly received symbols. On the other hand, the random channel selection algorithm results in inferior performance, as expected. In Figure. 3.2, the channel indices selected by the proposed UCB bandit algorithm are shown with respect to the block index. It is noted that the proposed approach rapidly converges to the best channel (channel 1). However, channel 2 is still selected for the last few blocks due to the exploration nature of the UCB bandit algorithm. To illustrate the convergence of the proposed scheme further, a bar graph is presented in Figure. 3.3, which shows the number of times each channel is selected. This figure clearly illustrates that channel 1 is selected most of the time.

Next, we focus on the optimal channel switching strategy (**Approach 2**). In this case, virtual channels are considered as described in Section 3.2 in order to realize channel switching between channel pairs with various power levels and time sharing factors. Also, in the presence of known channel statistics at the transmitter, the theoretical limit is obtained from the solution of (3.31). In addition, the random channel selection approach is again implemented to provide a benchmark. We consider $P_{\mathrm{pk}} = 10$ and $\Delta = 0.1$, leading to 1150 virtual channels. The number of blocks transmitted in this case is equal to 10^8. The proposed UCB bandit algorithm (Algorithm 2) aims to find the optimal channel among these 1150 virtual channels to achieve the highest rate of correctly received symbols on average. From the theoretical formulation in (3.31), the optimal power levels are obtained as $P_1 = 4.0126$ and $P_2 = 8.9947$ with the time sharing factor of $\lambda = 0.4004$, yielding 0.7778 for the rate of correctly received

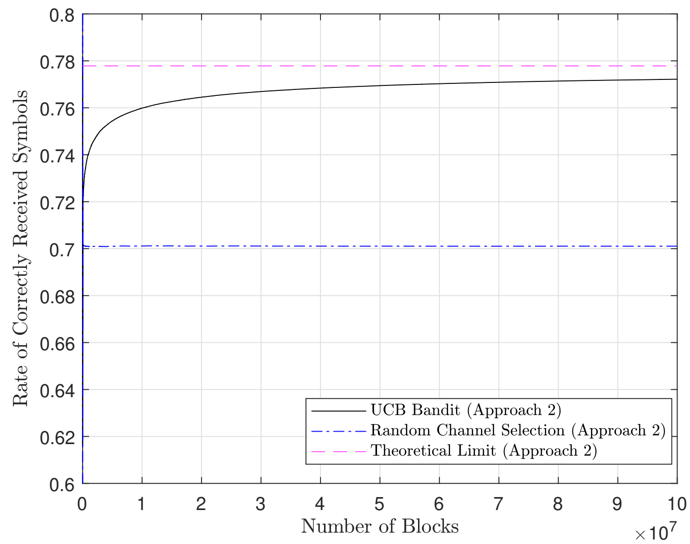

Figure 3.4: Rate of correctly received symbols achieved by the proposed UCB bandit algorithm for Approach 2 (Algorithm 2), together with the random channel selection strategy and the theoretical limit, which is obtained from (3.31).

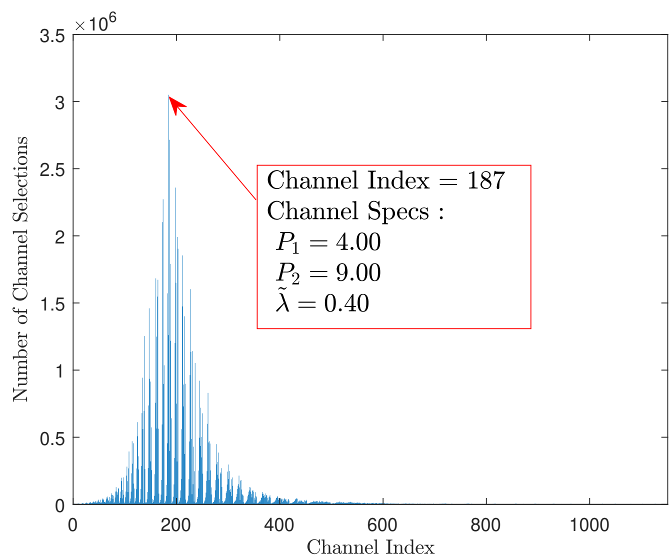

Figure 3.5: Number of times virtual channels are selected by the UCB bandit algorithm for Approach 2 (Algorithm 2).

symbols. In Figure. 3.4, we plot the rate of correctly received symbols versus the block index for the proposed UCB bandit algorithm (Algorithm 2), along with the theoretical limit and the random channel selection algorithm. It is seen that the proposed UCB bandit algorithm converges to the best strategy as the number of the blocks increases, and the rate of correctly received symbols becomes 0.7722 for 10^8 blocks (iterations), which is very close to the theoretical limit of 0.7778. Also, the bar graph in Figure. 3.5 illustrates the number of times each virtual channel is selected. As specified in the figure, the optimal virtual channel that is most frequently selected by the proposed UCB bandit algorithm corresponds to the solution of the theoretical formulation in (3.31) subject to quantization inaccuracies. Hence, the proposed UCB bandit algorithm for Approach 2 (Algorithm 2) learns the best channel switching strategy and converges to the optimal solution while significantly outperforming the random channel selection strategy.

The discretization parameter Δ plays an important role in determining the performance of the proposed UCB bandit algorithm for Approach 2 (Algorithm 2). On one hand, Δ should be sufficiently small so that the optimal combination of P_1, P_2, and λ parameters is included with sufficient accuracy in the set of virtual channels. On the other hand, as Δ decreases, the number of virtual channels increases, resulting in a requirement of more iterations to converge to the optimal solution. In Figure. 3.6, we plot the performance of the proposed UCB bandit algorithm versus Δ for different values of the number of blocks (iterations), where Δ is varied from 0.05 to 1. Also, the theoretical limits for the single channel and channel switching strategies are included in the figure. When Δ is 0.05, the total number of virtual channels is 4555, and when Δ is increased to 0.1, the total number of virtual channels decreases to 1084. The number of virtual channels decreases in the same fashion and becomes equal to 2 when Δ is equal to 1. It can be seen that when Δ is 0.05, the UCB bandit algorithm is unable to exceed the theoretical limit for the single channel strategy when the number of iterations is lower than 10^7 since it cannot find the best channels due to the insufficient number of iterations. However, as the number of iterations increases to 10^7 or higher, the UCB bandit algorithm beats the single

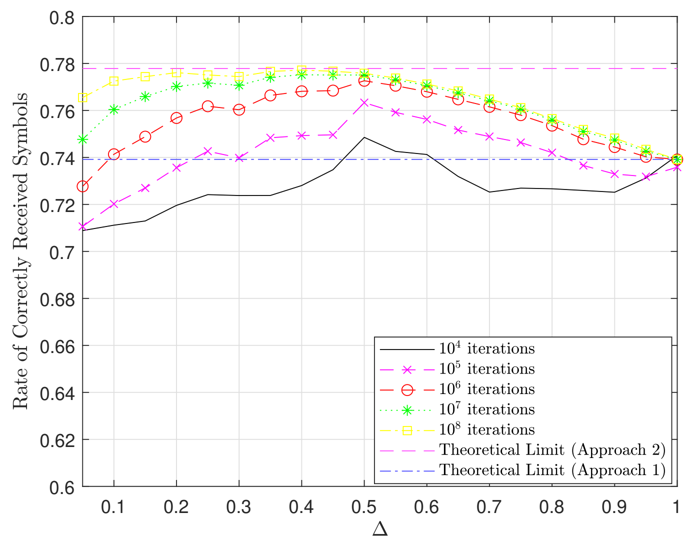

Figure 3.6: Rate of correctly received symbols achieved by the UCB bandit algorithm for Approach 2 (Algorithm 2) with respect to Δ, considering different values for the number of iterations, N_b.

channel strategy and starts getting close to the theoretical limits for the channel switching case, which is 0.7778. Furthermore, it is observed that as Δ and the number of iterations increase, the UCB starts getting close to the theoretical limit; for example, when Δ is equal to 0.5, the total number of virtual channels is 15 and in this case, for all the considered values of iterations, the UCB bandit algorithm outperforms the optimal single channel strategy, and approaches the channel switching theoretical limit when the number of iterations is above 10^7. For the values of Δ above 0.5, the performance of the UCB bandit algorithm starts deteriorating since the optimal solution cannot be well approximated by the virtual channels due to the large value of Δ. At last, when Δ is 1, the UCB bandit algorithm converges to the optimal single channel strategy for all values of the number of iterations, as expected.

In Figure. 3.7, we plot the rate of correctly received symbols versus Δ for the UCB-CE algorithm for the same setup as defined for Figure. 3.6. The parameters of the UCB-CE algorithm are taken as $N_{\text{int}} = 10$, $\chi = 0.3$ and $\eta = 5$. As discussed in Section 3.2.3, UCB-CE can outperform the UCB bandit algorithm by converging quickly and efficiently to optimal or close to optimal channels in

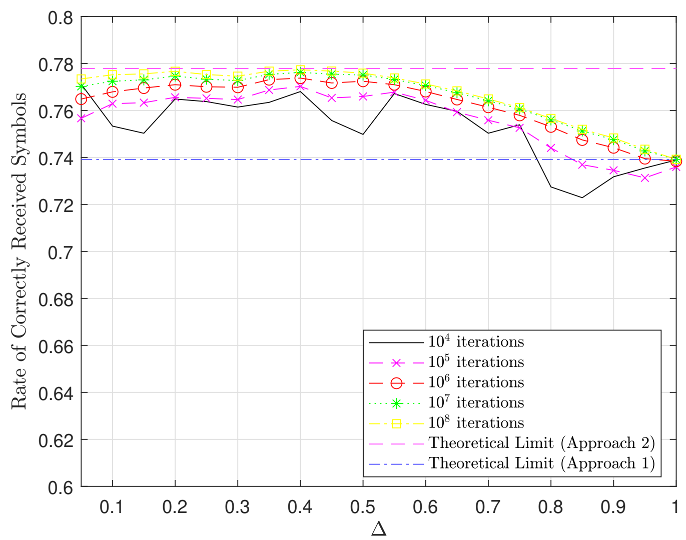

Figure 3.7: Rate of correctly received symbols achieved by the UCB-CE algorithm for Approach 2 (Algorithm 3) with respect to Δ, considering different values for the number of iterations, N_b.

cases with high numbers of bandits (virtual channels). As noted from Figure. 3.7, UCB-CE obtains high rates of correctly received symbols as compared to the UCB bandit algorithm for all considered numbers of iterations when Δ is sufficiently low. On the other hand, when Δ is high, the performance of both algorithms are very close since the number of channels is not very high in those cases. Another advantage of the UCB-CE algorithm is that it requires lower execution times since the number of channels/bandits decreases over the intervals according to the channel elimination parameter χ by considering the elimination threshold η (see Algorithm 3). To illustrate the execution time benefits of the UCB-CE algorithm, we present some results in Table 3.1 related to the execution times of the UCB-CE and UCB bandit algorithms for different numbers of iterations when Δ is 0.05, 0.1, and 0.5. As observed from the table, the execution times required by UCB-CE are significantly lower than those for UCB when the number of iterations is large and Δ is small. For example, when $\Delta = 0.05$, the number of virtual channels is 4555 (i.e., $MN_p = 4555$), and for 10^7 iterations (blocks), the UCB takes 500.04 seconds and yields 0.7478 for the rate of correctly received symbols. For the same setting, UCB-CE takes 83.23 seconds and results in 0.7702 for the rate of correctly received symbols. Hence, UCB-CE not only saves time but also rapidly converges towards the optimal solution, yielding a higher rate of correctly received symbols than the UCB algorithm. On the other hand, when $\Delta = 0.5$, the results in Table 3.1 show that the time taken by the UCB algorithm is slightly lower than that taken by UCB-CE since the number of virtual channels is small in that case (for $\Delta = 0.5$, the number of virtual channels is 15) and some extra steps in the UCB-CE algorithm becomes significant in the execution time. Therefore, the UCB-CE algorithm is most useful when there exist a large number of bandits (virtual channels).

Table 3.1: Execution times of the UCB and UCB-CE algorithms.

Δ	Number of Iterations	UCB execution time (sec)	UCB-CE execution time (sec)
0.05	10^4	4.8781×10^{-1}	1.5825×10^{-1}
	10^5	0.4841×10^1	0.0577×10^1
	10^6	5.1092×10^1	0.7224×10^1
	10^7	5.0004×10^2	0.8323×10^2
	10^8	4.8952×10^3	0.8184×10^3
0.1	10^4	1.2599×10^{-1}	0.0418×10^{-1}
	10^5	0.1175×10^1	0.0239×10^1
	10^6	1.2747×10^1	0.2348×10^1
	10^7	1.2288×10^2	0.2754×10^2
	10^8	1.1803×10^3	0.2515×10^3
0.5	10^4	0.0859×10^{-1}	0.1342×10^{-1}
	10^5	0.0075×10^1	0.0093×10^1
	10^6	0.0772×10^1	0.0960×10^1
	10^7	0.0744×10^2	0.0905×10^2
	10^8	0.0738×10^3	0.08849×10^3

3.5 Concluding Remarks

In this chapter, optimal single channel and channel switching strategies have been investigated for maximization of the number of correctly received symbols by considering unknown channel statistics at the transmitter. We have first focused on the scenario with single channel utilization in each block and utilized the UCB bandit algorithm for proposing a learning based strategy (Algorithm 1). Then, in the presence of channel switching capability in each block, we have introduced virtual channels corresponding to channel switching between all channel pairs with all possible power level pairs and time sharing factor. Consequently, we have proposed a UCB bandit based algorithm (Algorithm 2) for learning the best channel switching solution. In addition, for faster and more efficient convergence, the UCB-CE algorithm has been proposed when the number of virtual channels is high (Algorithm 3). For comparison purposes, the theoretical limits have also been presented when the channel statistics are known at the transmitter. Via simulation studies, it has been illustrated that the proposed UCB bandit based algorithms for both single channel and channel switching strategies converge to the theoretical limits for large numbers of blocks. In addition, it has been noted that channel switching can provide performance improvements over the single channel strategy in some cases. Therefore, even when the transmitter does not have the channel statistics, benefits of channel switching can be realized based on the proposed learning based strategies.

Chapter 4

Conclusion and Future Work

In this thesis, we first addressed, in Chapter 2, the problem of optimal channel switching and randomization for maximizing the outage capacity in flat-fading Gaussian noise channels, considering average power and outage probability constraints. For the single-user scenario, we proved that the optimal solution can be achieved through three strategies: transmission over a single channel with no randomization, channel switching between two channels with no randomization, or randomization between two parameter sets over a single channel. These strategies provide a simple and effective way to obtain the optimal solution. In the multiuser scenario, obtaining the optimal solution becomes computationally complex. To overcome this challenge, we proposed an algorithm based on linearly constrained linear optimization to calculate an approximately optimal channel switching and randomization solution with adjustable approximation accuracy. This approach allows us to maximize the total outage capacity of multiple transmitter and receiver pairs efficiently. Numerical examples were presented to validate the theoretical results. We demonstrated that the proposed strategies outperform conventional approaches, especially for low average power constraints. The results highlight the effectiveness of channel switching and randomization in improving the system's performance.

There exist several avenues for future research in this area. Firstly, extending the analysis to consider fading channels with different characteristics, such as Rician or Nakagami fading, would provide a more comprehensive understanding of system behavior in practical scenarios. Additionally, investigating the impact of different outage capacity metrics and performance measures regarding the optimal channel switching and randomization strategies could provide insights into the trade-offs between various system objectives, such as throughput, reliability, or energy efficiency. Furthermore, exploring adaptive algorithms that dynamically adjust the channel switching and randomization parameters based on the changing channel conditions and system requirements could enhance system adaptability and robustness. Finally, considering practical constraints and limitations, such as hardware impairments, non-ideal synchronization, or limited channel state information, would be valuable to evaluate the performance of the proposed strategies in realistic communication systems.

In conclusion, the optimal channel switching and randomization problem addressed in Chapter 2 provided valuable insights into enhancing communication system performance. The proposed strategies and algorithms offer effective solutions for both single-user and multiuser scenarios, with future research directions focusing on extending the analysis to more realistic scenarios and considering additional constraints and system objectives.

Chapter 3 of this thesis proposed channel selection and channel switching approaches based on the UCB bandit algorithm for communications over block fading channels, considering unknown channel statistics at the transmitter. The goal was to maximize the number of correctly received symbols per unit of time, also known as throughput. The proposed approaches addressed the challenges of channel switching and parameter randomization in the absence of prior knowledge about channel coefficients and their distributions. This study made several contributions. Firstly, it formulated the channel switching problem as a multi-armed bandit (MAB) problem, considering sequential decision making under uncertainty. This allowed the application of efficient MAB algorithms for throughput maximization. Two approaches were considered: Approach 1, where channel switching was performed only between blocks, and Approach 2, where channel switching was allowed within blocks as well.

For Approach 2, a set of virtual channels was introduced to determine the optimal channel switching strategy. The study proposed UCB bandit algorithms (Algorithms 1-3) to solve the channel selection and channel switching problems. Additionally, a low complexity version of the UCB bandit algorithm with channel elimination (UCB-CE) was proposed for efficient convergence when a high number of virtual channels existed. Theoretical limits were also presented for comparison when the channel statistics were known at the transmitter. Numerical examples were provided to evaluate the performance of the proposed approaches. Simulation results demonstrated that the proposed UCB bandit algorithms achieved close performance to the theoretical limits over a sufficiently large number of blocks. The algorithms effectively learned the best channels and channel switching strategies, outperforming the random channel selection approach. The convergence and effectiveness of the algorithms were illustrated through simulation results showing the rate of correctly received symbols and the channel selection patterns. Overall, this study contributed to the efficient utilization of spectrum resources in wireless communications by addressing the challenges of channel switching and parameter randomization in the presence of unknown channel statistics. The proposed UCB bandit algorithms provide practical solutions for maximizing throughput, even without prior knowledge of channel characteristics. Further research can explore extensions and variations of the proposed approaches to address additional complexities and practical scenarios in wireless communication systems.

In conclusion, the proposed channel selection and channel switching approaches based on the UCB bandit algorithm have shown promising results for maximizing the number of correctly received symbols per unit of time over block fading channels. However, there are several avenues for future research and extensions that would further enhance the understanding and applicability of these algorithms in practical communicatioFn systems. This study focused on a transmitter-receiver scenario with a single RF chain. Extending the proposed algorithms to multi-user scenarios, where multiple transmitters and receivers coexist in the same environment, would be of great interest. Investigating the performance of the algorithms in multi-user scenarios and exploring potential coordination strategies among users would provide insights into the scalability and applicability of the proposed approaches. While the results presented in this study were obtained through simulation experiments, future work could involve conducting real-world experiments or implementing the proposed algorithms on test-beds. Experimental validation would provide a more realistic assessment of the performance and practical feasibility of the proposed UCB bandit algorithms.

Bibliography

[1] S. Masrur and S. Gezici, "Optimal channel switching and randomization over flat-fading channels for outage capacity maximization," *IEEE Communications Letters*, vol. 27, no. 2, pp. 630–634, 2023.

[2] S. Masrur and S. Gezici, "Channel switching strategies based on UCB bandit algorithm for throughput maximization with unknown channel statistics." Under review, 2023.

[3] M. Azizoglu, "Convexity properties in binary detection problems," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1316–1321, July 1996.

[4] A. D. Sezer, S. Gezici, and H. Inaltekin, "Optimal channel switching strategy for average capacity maximization," *IEEE Trans. Commun.*, vol. 63, pp. 2143–2157, June 2015.

[5] B. Bozkurt, A. D. Sezer, S. Gezici, and T. Girici, "A game theoretic approach to channel switching in the presence of jamming," *IEEE Communications Letters*, vol. 25, no. 12, pp. 3927–3931, 2021.

[6] A. D. Sezer and S. Gezici, "Optimal channel switching in multiuser systems under average capacity constraints," *Digital Signal Processing*, vol. 63, pp. 145–154, 2017.

[7] B. Dulek, N. D. Vanli, S. Gezici, and P. K. Varshney, "Optimum power randomization for the minimization of outage probability," *IEEE Trans. Wirel. Commun.*, vol. 12, pp. 4627–4637, Sep. 2013.

[8] S. Amuru and R. M. Buehrer, "Optimal jamming against digital modulation," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 10, pp. 2212–2224, 2015.

[9] C. Goken, S. Gezici, and O. Arikan, "Optimal stochastic signaling for power-constrained binary communications systems," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 3650–3661, Dec. 2010.

[10] A. Patel and B. Kosko, "Optimal noise benefits in Neyman-Pearson and inequality-constrained signal detection," *IEEE Trans. Sig. Processing*, vol. 57, pp. 1655–1669, May 2009.

[11] S. Bayram and S. Gezici, "Stochastic resonance in binary composite hypothesis-testing problems in the Neyman-Pearson framework," *Digital Signal Processing*, vol. 22, pp. 391–406, May 2012.

[12] D. Xu, E. Jung, and X. Liu, "Optimal bandwidth selection in multi-channel cognitive radio networks: How much is too much?," in *IEEE Symp. New Frontiers in Dynamic Spectrum Access Networks*, Oct. 2008.

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

[14] M. F. Keskin, M. N. Kurt, M. E. Tutay, S. Gezici, and O. Arikan, "Maximization of average number of correctly received symbols over multiple channels in the presence of idle periods," *Digital Signal Processing*, vol. 54, pp. 95–118, 2016.

[15] S. Amuru and R. M. Buehrer, "Optimal jamming against digital modulation," *IEEE Trans. Inf. Forensics Security*, vol. 10, pp. 2212–2224, Oct. 2015.

[16] T. L. Lai, H. Robbins, *et al.*, "Asymptotically efficient adaptive allocation rules," *Advances in Applied Mathematics*, vol. 6, no. 1, pp. 4–22, 1985.

[17] S. Bubeck, N. Cesa-Bianchi, *et al.*, "Regret analysis of stochastic and nonstochastic multi-armed bandit problems," *Foundations and Trends in Machine Learning*, vol. 5, no. 1, pp. 1–122, 2012.

[18] Z. Wang, R. Zhou, and C. Shen, "Regional multi-armed bandits with partial informativeness," *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5705–5717, 2018.

[19] C. Shen, "Universal best arm identification," *IEEE Transactions on Signal Processing*, vol. 67, no. 17, pp. 4464–4478, 2019.

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

[21] M. T. Masonta, M. Mzyece, and N. Ntlatlapa, "Spectrum decision in cognitive radio networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1088–1107, 2012.

[22] L. Lai, H. El Gamal, H. Jiang, and H. V. Poor, "Cognitive medium access: Exploration, exploitation, and competition," *IEEE Transactions on Mobile Computing*, vol. 10, no. 2, pp. 239–253, 2010.

[23] Y. Xu, A. Anpalagan, Q. Wu, L. Shen, Z. Gao, and J. Wang, "Decision-theoretic distributed channel selection for opportunistic spectrum access: Strategies, challenges and solutions," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 1689–1713, 2013.

[24] H. Liu, B. Krishnamachari, and Q. Zhao, "Cooperation and learning in multiuser opportunistic spectrum access," in *ICC Workshops-2008 IEEE International Conference on Communications Workshops*, pp. 487–492, IEEE, 2008.

[25] K. Liu and Q. Zhao, "Distributed learning in multi-armed bandit with multiple players," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5667–5681, 2010.

[26] O. Avner and S. Mannor, "Concurrent bandits and cognitive radio networks," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 66–81, Springer, 2014.

[27] J. Rosenski, O. Shamir, and L. Szlak, "Multi-player bandits–a musical chairs approach," in *International Conference on Machine Learning*, pp. 155–163, PMLR, 2016.

[28] M. Bande and V. V. Veeravalli, "Multi-user multi-armed bandits for uncoordinated spectrum access," in *2019 International Conference on Computing, Networking and Communications (ICNC)*, pp. 653–657, IEEE, 2019.

[29] M. Di Felice, K. R. Chowdhury, and L. Bononi, "Learning with the bandit: A cooperative spectrum selection scheme for cognitive radio networks," in *2011 IEEE Global Telecommunications Conference-GLOBECOM 2011*, pp. 1–6, IEEE, 2011.

[30] C. Tekin and M. Liu, "Online learning in opportunistic spectrum access: A restless bandit approach," in *2011 Proceedings IEEE INFOCOM*, pp. 2462–2470, IEEE, 2011.

[31] H. Liu, K. Liu, and Q. Zhao, "Learning in a changing world: Restless multiarmed bandit with unknown dynamics," *IEEE Transactions on Information Theory*, vol. 59, no. 3, pp. 1902–1916, 2012.

[32] B. Li, P. Yang, J. Wang, Q. Wu, S. Tang, X.-Y. Li, and Y. Liu, "Almost optimal dynamically-ordered channel sensing and accessing for cognitive networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 10, pp. 2215–2228, 2013.

[33] S. K. Rashed, R. Shahbazian, and S. A. Ghorashi, "Spectrum decision in cognitive radio networks using multi-armed bandit," in *2015 5th International Conference on Computer and Knowledge Engineering (ICCKE)*, pp. 143–146, IEEE, 2015.

[34] C. Shi and C. Shen, "Multi-player multi-armed bandits with collision-dependent reward distributions," *IEEE Transactions on Signal Processing*, vol. 69, pp. 4385–4402, 2021.

[35] M. Lelarge, A. Proutiere, and M. S. Talebi, "Spectrum bandit optimization," in *2013 IEEE Information Theory Workshop (ITW)*, pp. 1–5, IEEE, 2013.

[36] H. Jiang, H. He, L. Liu, and Y. Yi, "Q-learning for non-cooperative channel access game of cognitive radio networks," in *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, IEEE, 2018.

[37] A. Anandkumar, N. Michael, A. K. Tang, and A. Swami, "Distributed algorithms for learning and cognitive medium access with logarithmic regret," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 731–745, 2011.

[38] A. D. Sezer and S. Gezici, "Average capacity maximization via channel switching in the presence of additive white Gaussian noise channels and switching delays," *IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 6228–6243, 2016.

[39] R. T. Rockafellar, *Convex Analysis.* Princeton, NJ: Princeton University Press, 1968.

[40] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization.* Boston, MA: Athena Specific, 2003.

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

[42] L. Zhou, "A survey on contextual multi-armed bandits," *arXiv preprint arXiv:1508.03326*, 2015.

[43] A. Garivier and O. Cappé, "The kl-ucb algorithm for bounded stochastic bandits and beyond," in *Proceedings of the 24th annual conference on learning theory*, pp. 359–376, JMLR Workshop and Conference Proceedings, 2011.

[44] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine learning*, vol. 47, no. 2, pp. 235–256, 2002.