Optimal Channel-Switching Strategies in Multi-channel Wireless Networks

by

Qingsi Wang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering: Systems)
in The University of Michigan
2014

Doctoral Committee:

Professor Mingyan Liu, Chair
Assistant Professor Jacob Abernethy
Associate Professor Achilleas Anastasopoulos
Professor Demosthenis Teneketzis

To my parents and grandparents

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor Professor Mingyan Liu, who offered me the great opportunity to become a Wolverine and guided me through years of study and work, as an excellent mentor and friend. I would also like to thank all my committee members, Professors Demosthenis Teneketzis, Achilleas Anastasopoulos and Jacob Abernethy, who have also given me so many enlightening lectures with so much intellectual fun, which is one of the most important parts of my graduate education.

I am also thankful to all my friends in Ann Arbor and in China, for the million times of chatting on random topics, for all the Chinese and exotic food we shared, for the "tons of damage" we dealt in the League of Legends, for all the sweet bitter time in life we spent together. I definitely want to thank the lovely town Ann Arbor. The suffering piles of snows in the depressing winters make the warmth from the culture, the people, and the green summer sunshine of this city feel more precious and adorable than any other places I have visited, and I am glad I have left a five-year mark along the dimension of time on this spatial spot, in a universe that I will often revisit in my heart.

Last but not least, I am grateful to my parents and grandparents for their eternal support and encouragement. I am regretful for years of lack of verbal communication with them but I always know that all the conquering of hardship in these years would mean, without their love, nothing.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

Optimal Channel-Switching Strategies in Multi-channel Wireless Networks

by

Qingsi Wang

Chair: Professor Mingyan Liu

The dual nature of scarcity and under-utilization of spectrum resources, as well as recent advances in software-defined radio, led to extensive study on the design of transceivers that are capable of opportunistic channel access. By allowing users to dynamically select which channel(s) to use for transmission, the overall throughput performance and the spectrum utilization of the system can in general be improved, compared to one with a single channel or more static channel allocations. The reason for such improvement lies in the exploitation of the underlying temporal, spatial, spectral and congestion diversity. In this dissertation, we focus on the channel-switching/hopping decision of a (group of) legitimate user(s) in a multi-channel wireless communication system, and study three closely related problems: 1) a jamming defense problem against a no-regret learning attacker, 2) a jamming defense problem with minimax (worst-case) optimal channel-switching strategies, and 3) the throughput optimal strategies for a group of competing users in IEEE 802.11-like medium access schemes.

For the first problem we study the interaction between a user and an attacker from a learning perspective, where an online learner naturally adapts to the available information on the adversarial environment over time, and evolves its strategy with certain payoff guarantee. We show how the user can counter a strong learning attacker with knowledge on its learning rationale, and how the learning technique can itself be considered as a countermeasure with no such prior information. We further consider in the second problem the worst-case optimal strategy for the user without prior information on the attacking pattern, except that the attacker is subject to a resource constraint, which models its energy consumption and replenishment process. We provide explicit characterization for the optimal strategies and show the most damaging attacker, interestingly, behaves randomly in an i.i.d. fashion. In the last problem, we consider a group of competing users in a non-adversarial setting. We place the interaction among users in the context of IEEE 802.11-like medium access schemes, and derive decentralized channel allocation for overall throughput improvement. We show the typically rule-of-thumb load balancing principle in spectrum resource sharing can be indeed throughput optimal.

CHAPTER I

Introduction

1.1 Motivation and Overview

Advances in software-defined radio in recent years have motivated numerous studies on building agile, channel-aware transceivers that are capable of sensing instantaneous channel quality [1, 2, 3]. With this opportunity comes the challenge of making effective opportunistic channel access and transmission scheduling decisions, as well as designing supporting system architectures. In this research, we focus on the channel-hopping decision of a (group of) legitimate user(s) in multi-channel wireless communication systems. By allowing users to dynamically select which channel to use for transmission, we aim to improve the system performance (e.g., throughput) compared to a system with a single channel or more static channel allocations, leveraging the potential temporal, spatial, spectral and congestion diversity. In particular, stemming from the assumption on motivating incentives and performance criteria of channel switching, we study three closely related problems: 1) a jamming defense problem against a no-regret learning attacker, 2) a jamming defense problem with minimax (worst-case) optimal channel-switching strategies, and 3) the throughput optimal strategies for a group of competing users in IEEE 802.11-like medium access schemes.

Existing work on jamming defense problems typically either assumes stationary or heuristic behavior of one side and examines countermeasures of the other, or assumes both sides to be strategic which leads to a game theoretical framework. Results from the former often lack robustness against changes in the adversarial behavior, while those from the latter may be difficult to justify due to the implied full information (either as realizations or as distributions) and rationality, both of which may be limited in practice. In the first jamming defense problem of this research, we take a different approach by assuming an intelligent attacker that is adaptive to the information available to it and is capable of learning over time with performance guarantee via repeated interaction. In the second problem, we revisit the jamming defense assuming an attacker that is subject to a resource (e.g. power) constraint with possibly a replenishment process, and meanwhile, no prior statistical information on the attacking pattern is known to the user. We consider the minimax optimal strategy of the user in a multi-stage interaction, i.e., the worst-case optimality result. In most commercial applications, the interference to a user is not from a malicious attacker but rather competing legitimate users. Our last problem then considers the interaction among the group of users and we seek efficient decentralized channel allocation for overall throughput enhancement of the widely deployed IEEE 802.11-like medium access schemes.

1.2 Literature Review

In this section, we provide a brief review of existing work related to our problem formulation and solution techniques. The first two problems in our study concern the channel-switching decision of a legitimate user in the presence of jamming attacks. In the first problem, we model the adaptive reasoning and decision-making of the jammer using no-regret online learning algorithms, which provides considerable performance guarantee for the attacker. We then consider in the second problem a more vicious attacker, which is a strict competitor in a zero-sum sense in terms of payoffs of the competing user and attacker, and we develop a worst-case optimal learning technique as the solution for the user. Our last problem will be placed in the context of medium access in multi-channel IEEE 802.11 WLANs, and we examine the decentralized channel-switching decisions in a population of users for throughput improvement. We hence in the following summarize literature related to jamming defense problems, online learning theory and modeling of throughput performance in IEEE 802.11 WLANs.

Jamming defense. The jamming defense problem can be considered in a general context as an instance of the pursuit-evasion problem, and the decision-making of both sides have been studied with extensive efforts. A big part of the literature considers stationary or heuristic behavior of one side and examines corresponding countermeasures of the other. Examples include [4, 5, 6, 7] and the references therein, that assume a stationary target (the evader) hiding in any of a set of locations with known prior probabilities. Variants of this model include, e.g., [8] that uses a random prior probability of hiding in a given location, and [9] where the detection probability is random with known distribution. Search problems with a moving evader have also been extensively studied. However, the evasion is typically either independent of the pursuer's activity, or heuristically given without clearly defined rationale or performance guarantee, see e.g., [10], where the evader's motion is given by a discrete-time Markov chain independent of the pursuer's activity, and [11] for a similar, continuous-time formulation. For applied studies in jamming and defense, see e.g., [12, 13] for a collection of specific attack mechanisms and anti-jamming measures. Examples also include using stronger error detection, correction, and spreading codes at the physical layer [14, 15, 16, 17], exploring the vulnerability in the rate adaptation mechanism of IEEE 802.11 [18], and multi-channel jamming using a single cognitive ratio [19]. Interestingly, jamming can also be used by legitimate users to achieve physical layer security in the presence of an eavesdropper, see e.g., [20, 21, 22].

The interacting attacker and defender are also often assumed to be strategic, leading to a game theoretical framework. A typical method is to use differential games [23] to capture the continuous evolution; in fact, the pursuit-evasion problem bears the genesis of differential games. See also [24, 25, 26] for texts and examples of differential games and their application in the pursuit-evasion problem. For the particular jamming defense problem, the two-player interaction can be in terms of respective power control or channel selection strategies. Examples include a non-zero-sum game formulation when transmission costs are incurred to both the jammer and the user [27], a random access game [28], a differential game between a mobile jammer and mobile users [29], a Stackelberg game [30], and a zero-sum game framework [31]. Typically the existence of Nash equilibrium strategies is investigated and the equilibrium strategies are identified if they exist under the respective game formulation. We note that existing results in general focus on analyzing the one-stage game, while the multi-stage or the repeated case is often elusive in analysis and replaced with various approximated problems, e.g. [32].

Online learning. Sequential decision-making in an environment, which generates as feedback reward or penalty with uncertain, models the basic feature of a variety of situations in everyday life and engineering applications. Online learning concerns the development of an adaptive and systematic decision procedure (or an algorithm), given past decisions and feedbacks, so as to optimize the utility of the learner. In this research, we mainly utilize formulation and results from the online learning theory in the so-called "adversarial setting", which assumes no prior statistical knowledge on the environment (See [33] for a recent text). Given the unknown but non-stochastic nature of the environment, the performance of a learning algorithm is typically evaluated using the notion of regret, defined as the difference in reward or loss obtained between a suitably defined strategy in hindsight and that accumulated by the algorithm over time. There have been abundant results in realizing (asymptotic) no-regret learning algorithms that yield (order-optimal) regret sublinear in time (see [34, 35, 36, 37] as examples for a line of continuing efforts, see [38] for an in-depth survey and references therein). In the context of jamming defense problem, the no-regret performance guarantee translates into sublinear "missing" of effective jamming or interference-free transmission opportunities, compared to a certain class of strategies in hindsight.

The rationale behind online learning differs from the strategic reasoning behind game-theoretical analysis, in that typically neither the learner nor the environment (possibly a multitude of other learner with conflicting interest) has unbounded rationality. However, when all interacting parties adopt learning techniques, the asymptotic interplay may exhibit well-defined equilibria, that are closely related to game-theoretical solution concepts. thus providing an alternative interpretation of game theoretical results. In fact, the generic characterization on the learning limit using no-regret algorithms would be the convergence to the set of (coarse) correlated equilibria [39] [40], which is however usually weak in most applications given a broad set of equilibrium points. As to the convergence of learning to Nash Equilibrium (NE), it has been shown no-regret dynamics may not converge to NE in general games [41], with however a few affirmative results in special cases [42, 43].

IEEE 802.11 DCF. The medium access control in IEEE 802.11-based WLANs utilizes the standardized Distributed Coordination Function (DCF) [44], a contention-based MAC protocol. DCF has been very extensively studied in the literature, ranging from throughput performance in the saturated regime [45, 46] and the non-saturated regime [47, 48], to its rate region [49, 50], to channel assignment in multi-channel WLANs [51, 52], to name a few. We consider decentralized channel-switching strategies for a group of contending users in a multi-channel system under DCF, and to evaluate the throughput performance of a given scheme for a variety of network traffic conditions, we use the notion of its stability region, which is given as the set of all sustainable combinations of incoming traffic loads at users. To the best of our knowledge, however, none has studied multi-channel WLANs in the context of stability region. Works most relevant to ours include ones on the stability region of slotted Aloha (e.g., [53]) and the rate region of 802.11 DCF [49, 50].

1.3 Organization and Main Contributions

The first two chapters focus on the optimal channel switching as jamming defense. With respect to different set of assumptions on the adversarial behavior of the attacker and optimality criteria, we consider two problems, respectively.

Chapter II: Optimal Channel Switching as Jamming Defense - Part I: Against a No-Regret Learning Attacker [54, 55]

Assuming an online learning attacker, we investigate two cases depending on the knowledge of the user. In the first case we assume the user is aware of the type of learning algorithm used by the attacker, while in the second case it does not have such information and thus must try to learn. We show that the optimal policies in the first case have a greedy nature. This result is then used to assess the performance of the learning algorithms that both sides employ in the second case, which is shown to be mutually optimal and there is no loss for either side compared to the case when the user knows perfectly the adaptive pattern used by the adversary and responses optimally. Based on the above result, we also show that if in addition to the channel-switching decisions the user also needs to perform transmission power control, there is one-way decoupling of the joint control of channel selection and power control. That is, the power control can be independent from the channel selection.

Chapter III: Optimal Channel Switching as Jamming Defense - Part II: Against a Resource-Replenishing Attacker with Minimax Optimality [56]

The worst-case optimality criterion leads to a repeated zero-sum game theoretical framework as our main solution technique; however, this framework does not originate from the assumption on the rationality of the jammer or its motivating payoff, but rather the learner's (user's) objective of optimizing achievable payoff unilaterally. Interestingly, we show that the most damaging attacker for the user can be given as an adversary who behaves in an i.i.d. manner in the multi-stage interaction. Based on this, we provide the explicit characterization of the optimal channel-switching strategies of the user, which is determined by the induced random walk of the adversarial behavior. In addition to the jamming defense, our framework is also applicable to other competitive game problems with finite action spaces.

The interference to a given user is usually a result of competition of other legitimate users in commercial applications, instead of attack from a malicious entity. The next two chapters then consider the problem of using channel switching as spectrum resource sharing technique to mitigate interference among users and aim to promote the throughput performance.

Chapter IV: Throughput Optimal Channel Switching in Random Access - Part I: Intuition from Slotted Aloha

This chapter aims to provide theoretical preparation and insights, using slotted Aloha and population game based simplification, for throughput optimal switching in IEEE 802.11 Distributed Coordination Function (DCF), a widely deployed industrial standard medium access scheme that we study in Chapter V. The retrospective thinking by replacing DCF by Aloha will continue to be vital for our analysis on the would-be more involved DCF, and the intuition we shall conclude from this chapter will be shown to be consistent with our results for DCF.

Chapter V: Throughput Optimal Channel Switching in Random Access - Part II: in IEEE 802.11 WLANs [57]

We observe that in a multi-channel wireless system, an opportunistic channel/spectrum access scheme that solely focuses on channel quality sensing measured by received SNR may induce users to use channels that, while providing better signals, are more congested. Ultimately the notion of channel quality should include both the signal quality and the level of congestion, and a good multi-channel access scheme should take both into account in deciding which channel to use and when. Motivated by this, we focus on the congestion aspect and examine what type of dynamic channel-switching schemes may result in the best system throughput performance. This will be evaluated using the notion of stability region of a scheme. This is because more effective resource allocation and sharing can achieve a lower overall congestion level, thus expanding the range of sustainable arrival rates and resulting in a larger stability region. The scheme with the largest such region is commonly known as the throughput optimal scheme. We derive the stability region of a multi-user multi-channel Wireless Local Area Network (WLAN) system and determine the throughput optimal channel-switching scheme within a certain class of schemes.

Chapter VI: Conclusion and Future Work

We summarize the main contributions of this dissertation and discuss the topics that can be further pursued based on this research.

1.4 Notation Convention

For time-varying vector quantities, we typically reserve the subscript for its entry indices, and use the superscript for time and other annotations. For example, we use

$w^t = (w_1^t, w_2^t, \ldots, w_n^t)$ to denote a time-varying probability distribution vector over the set $\{1, 2, \ldots, n\}$, where w_k^t is the probability mass assigned to k. The convention of using the superscript for time also applies to general quantities when there are multiple indices pertaining to them except for a few instances. The main notation used in each chapter is summarized in Appendix D.

CHAPTER II

Optimal Channel Switching as Jamming Defense - Part I: Against a No-Regret Learning Attacker

The vulnerabilities of wireless networks to security attacks given their broadcast nature, and accordingly how to build resilient defenses, have been subjects of extensive research. In the next two chapters, we utilize channel switching as a defense strategy in multi-channel systems against jamming, which is a common type of Denial-of-Service (DoS) attacks, and we investigate the optimal response from a legitimate user to a jamming attacker in repeated interaction.

2.1 Problem Formulation

We shall formulate the problem in a general context of two-player repeated game[1]. At time (or round) t, the row and the column player respectively choose distributions (mixed strategies) w^t and u^t over their action spaces $\mathcal{A}_w$ and $\mathcal{A}_u$. We assume both the action spaces $\mathcal{A}_w$ and $\mathcal{A}_u$ are finite, and let m and n be their respective cardinality. An action $i_t \in \mathcal{A}_w$ is then realized per w^t independently for the row player, and so does an action j_t for the column player. The row player receives a payoff $M_w(i_t, j_t)$, where M_w is a $n \times m$ payoff matrix for the row player, and $M_w(i_t, j_t)$ denotes the (i_t, j_t)

[1]We use the word *game* in a broad sense, without necessarily implying a game-theoretical strategic analysis, unless otherwise stated.

entry of M_w. Similarly, we denote by M_u the payoff matrix for the column player, who observes a payoff $M_u(i_t, j_t)$. We denote by $\mathcal{I}_w^t$ (resp. $\mathcal{I}_u^t$) the informational state of the row player (resp. the column player) at time t, which consists of all information or a sufficient statistic available to the player for decision making, and by g_w^t (resp. g_u^t) the decision rule at time t, that is, $w^t = g_w^t(\mathcal{I}_w^t)$ and $u^t = g_u^t(\mathcal{I}_u^t)$. Let $g_w = (g_w^1, g_w^2, \ldots)$ be the decision policy of the row player, which is the collection of decision rules, and let the space of all policies be $\mathcal{G}_w$. We similarly define the decision policy and the policy space for the column player. The above setup can accommodate a variety of two-player game problems depending on the structure of the payoff matrices. For our application in jamming defense, we let the row player be the user and the column player be the attacker (or the adversary interchangeably). Also, $\mathcal{A}_w = \mathcal{A}_u = [n] := \{1, 2, \ldots, n\}$ denotes the index set of n channels of the system. Without otherwise stated, we assume in this chapter:

Assumption II.1.

1. *$M_w(i,j) = \mathbf{1}_{i \neq j}$ and $M_u(i,j) = \mathbf{1}_{i=j}$, where $\mathbf{1}_{\{\cdot\}}$ is the indicator function.*

2. *Both the user and the attacker have perfect observation and recall of the opponent's and their own past actions.*

The key question is given by how to model the interaction between the two players and their behavior. In principle, each player can consider the optimization problems

$$\underset{g_v \in \mathcal{G}_v}{\text{maximize}} \quad \mathbb{E}\left\{\sum_{t=1}^{T} (w^t)^\top M_v u^t\right\}, \tag{2.1}$$

and

$$\underset{g_v \in \mathcal{G}_v}{\text{maximize}} \quad \liminf_{T \to \infty} \mathbb{E}\left\{\frac{1}{T}\sum_{t=1}^{T} (w^t)^\top M_v u^t\right\}, \tag{2.2}$$

where "v" denotes either "w" or "u", the superscript $\top$ denotes the transpose, T is a finite time horizon, and the expectation is taken with respect to any randomness involved in the evaluation of the total mean payoff. We note that while on one hand these problems are always well-defined up to a subjective belief on the opponent's behavior, it can be difficult to be justify the optimization with respect to (w.r.t.) this belief. In particular, when one has no prior knowledge on the opponent's behavior, this belief can be arbitrary. So the key boils down to the knowledge model on the opponent's behavior. Throughout this chapter, we assume:

Assumption II.2. *The attacker has no information on the decision-making rationale of the user. Then, instead of optimizing w.r.t. an arbitrary belief, the attacker is an online learner and adopts a no-regret learning algorithm as its policy.*

We thus investigate two cases depending on the knowledge of the user. In the first case we assume the user is aware of the type of learning algorithm used by the attacker, and the reasoning process of the user can then be given by the single-sided optimization problems (2.1) and (2.2). This assumption on the knowledge is unrealistically strong, but this case will serve as our baseline. With a bit abuse of language, this case will be referred to as the *known case*. In the second case, or the *unknown case*, it does not have such prior information, and there is hence a symmetry in the amount of information for both sides. We then assume the user also behaves as a learner, and we consider the asymptotic interaction.

2.2 Optimal Channel Switching against a No-Regret Learning Attacker

Online learning techniques typically provides certain performance guarantee; in particular, we consider the so-called "no-regret" guarantee, which we elaborate as follows. Given any sequence $i_{[T]} := (i_1, i_2, \ldots, i_T)$ of the row player's realization of actions over T in hindsight and the decision policy g_u of the column player, define the (external) regret $R_u(T; i_{[T]}, g_u)$ of the column player with respect to $i_{[T]}$ and g_u

at the horizon T as

$$R_u(T; i_{[T]}) = \max_{j\in[n]} \sum_{t=1}^{T} e_{i^t}^\top M_u e_j - \mathbb{E}\left\{\sum_{t=1}^{T} e_{i^t}^\top M_u u^t\right\},$$

where e_k denotes the degenerate distribution with probability one on action k, and the expectation is taken with respect to the randomness induced by g_u. The regret measures the gap between the payoff using a given policy and that using the best static policy that always selects the same action. An online learning algorithm as a policy for the column player is called "no-regret" if the regret is sublinear in time or $R_u(T; i_{[T]}) = o(T)$, that is, its average payoff is no worse than that of any static policy for any realization of the opponent's actions over time.

2.2.1 Against Adaptive Attack with Known Patterns

We consider in this chapter the baseline case, and we assume

Assumption II.3. *The attacker uses the no-regret algorithm Hedge (or called Exponential Weights Algorithm) [35, 36], detailed in Table 2.1, which is known to user along with the initial condition.*

As to its no-regret guarantee, the performance of Hedge is formally characterized by the following results from [36, 37].

Theorem II.4 ([36, 37])**.** *If $a = 1 + \sqrt{2\ln(n)/T}$, then $R_u(T; i_{[T]}) \leq \sqrt{2T \ln n}$ for any $i_{[T]}$, where the expectation is w.r.t. the randomness in the actions taken by Hedge, and the diminishing rate of the average regret over time is order-optimal.*

Given the knowledge on the fact that the attacker is using Hedge and its initial condition, and further due to the user's perfect recall of past actions and observations, it thus maintains the correct belief about the evolution of the adversary's mixed strategy u^t determined by Hedge, and the information state of the user is simply

Table 2.1: The Hedge algorithm.

Hedge
Parameter: A real number $a > 1$.
Initialization: Set $G_k^0 := 0$ for all $k \in [n]$.
Repeat for $t = 1, 2, \dots, T$

1. Choose channel j^t according to the distribution $u^t = (u_1^t, u_2^t, \dots, u_n^t)$ on $[n]$, where
$$u_j^t = \frac{a^{G_j^{t-1}}}{\sum_{k=1}^n a^{G_k^{t-1}}}$$
2. Observe the action i^t taken by the user, and obtain the (reward) vector $(x_1^t, x_2^t, \dots, x_n^t)$, where $x_j^t = M_u(i^t, j)$.
3. Set $G_k^t = G_k^{t-1} + x_k^t$ for all $k \in [n]$.

given by $\mathcal{I}_w^t = u^t$. In principle, the finite-horizon problem (2.1) can be solved using standard dynamic programming. However, we will first try to argue intuitively what the optimal policy should behave like. Since Hedge has a sublinear regret for the attacker, if the users favors one channel, the attacker will eventually identify this most user-active channel and jam it at a rate linear in T and miss it at a rate no more than sublinear in T. It follows that the best strategy for the user is to transmit on each channel evenly, either deterministically or stochastically. This intuition indeed provides the precise solution to the infinite-horizon problem (2.2) as shown below. Let $\bar{r}_v(g_w, g_u, T) = \mathbb{E}\left\{\frac{1}{T}\sum_{t=1}^T (w^t)^\top M_v u^t\right\}$ for any pair of policies g_w and g_u of the user and the attacker, and let $\bar{r}_v(g_w, g_u) = \liminf_{T\to\infty} \bar{r}_v(g_w, g_u, T)$. For any sequence $i_{[T]}$ of the user's realization of actions over T, define the greedy policy by $w^t = e_{i^t}$ where $i^t \in \arg\min_{j\in[n]} u_j^t$, and we have the following result.

Theorem II.5. *$\bar{r}_w(g_w, g_u) \le \frac{n-1}{n}$ for any policy g_w when the attacker's policy g_u is given by Hedge, and the greedy policy of the user achieves this upper bound.*

Proof. Note that

$$\bar{r}_w(g_w, g_u, T) = \mathbb{E}\left\{\frac{1}{T}\sum_{t=1}^{T}(w^t)^\top M_w u^t\right\} = 1-\mathbb{E}\left\{\frac{1}{T}\sum_{t=1}^{T}(w^t)^\top M_u u^t\right\} = 1-\bar{r}_u(g_w, g_u, T)$$

for any pair of g_w and g_u. Therefore, when g_u is given by Hedge, using Theorem II.4 we have

$$\begin{aligned}
\bar{r}_w(g_w, g_u) &= 1 - \limsup_{T\to\infty}\mathbb{E}\left\{\frac{1}{T}\sum_{t=1}^{T}(w^t)^\top M_u u^t\right\} \\
&\leq 1 - \limsup_{T\to\infty}\mathbb{E}\left\{\frac{1}{T}\left(\max_{j\in[n]}\sum_{t=1}^{T} e_{i^t}^\top M_u e_j - \sqrt{2T\ln m}\right)\right\} \\
&= 1 - \limsup_{T\to\infty}\mathbb{E}\left\{\frac{1}{T}\max_{j\in[n]}\sum_{t=1}^{T} e_{i^t}^\top M_u e_j\right\} \leq \frac{n-1}{n},
\end{aligned}$$

where the expectation operator on the second line incorporates the randomness in the realization of $i_{[T]}$, and the last inequality is due to the fact that $\max_{j\in[n]}\sum_{t=1}^{T} e_{i^t}^\top M_u e_j \geq \frac{T}{n}$ for any $i_{[T]}$.

Since $\min_{j\in[n]} u_j^t \leq \frac{1}{n}$ for any u^t, we have $(w^t)^\top M_w u^t \geq \frac{n-1}{n}$ for any t using the greedy policy, which implies that the greedy policy achieves the maximum average payoff for the user. □

Without loss of generality, we will assume under the greedy policy ties are broken in favor of the lowest-indexed channel. Note that since the greedy policy always selects the channel least likely to be jammed, it eventually (in finite time) leads to equal weights over all channels even if the initial weights under Hedge is unequal. Once the weights are equal, the user's action is a simple round robin, using channel in the order $1, 2, \cdots, n$. The above proof also suggests that any policy that results in an equal frequency of presence on each channel has the same infinite-horizon average payoff, thus asymptotically optimal. It should be noted that these equi-occupancy polices are not necessarily optimal for the finite-horizon problem posed in (2.1). The greedy policy, however, is in fact also optimal over the finite horizon. Below we prove this result for a two-channel scenario so as to avoid letting technicalities obscure the main idea. The general case is stated in a theorem and the proof can be found in Appendix A.

Lemma II.6. *In a two-channel scenario, the optimal finite-horizon policy yields* $w^t = e_k$ *if* $u_k^t < 1/2$ *where* $k = 1, 2$*, and is indifferent between 1 and 2 when* $u_1^t = u_2^t = 1/2$*.*

Proof. For any policy, let $\Delta(t) := |G_1^t - G_2^t|$; this is the difference between the number of times channel 1 and 2 have been used by the end of slot t. Thus $|\Delta(t+1) - \Delta(t)| = 1$ for all t. An example of $\Delta(t)$ up to T is shown in Figure 2.1: an edge connecting two adjacent time points represents a particular channel selection i_t of the user, a down edge indicating the selection of a currently under-utilized channel. Let $r(t) = e_{i_t}^\top M_w u^t$, i.e., the mean payoff of choosing i_t at time t. We then have

$$r(t) = \begin{cases} \frac{a^{\Delta(t-1)}}{1+a^{\Delta(t-1)}}, & \text{if } \Delta(t) < \Delta(t-1), \\ \frac{1}{1+a^{\Delta(t-1)}}, & \text{if } \Delta(t) > \Delta(t-1). \end{cases}$$

Suppose along any trajectory of $\Delta(t)$ there exists a point $\Delta(t) = d \geq 2$ such that either of the following cases is true: (C1) $d - 1 = \Delta(t-1) = \Delta(t+1) < \Delta(t)$, $t < T$; or (C2) $\Delta(T-1) < \Delta(T)$. Then consider a change of policy by "folding" the point at t down in (C1) and the point at T in (C2), as shown by the dashed line in the figure. Clearly, we would only change the payoff obtained at time t and $t+1$ for the case (C1) and that at time T for (C2). Let r' denote the mean payoff of this alternate policy. For (C1) we have

$$\begin{aligned} r'(t) + r'(t+1) - r(t) - r(t+1) &= \frac{a^{d-1}}{1+a^{d-1}} + \frac{1}{1+a^{d-2}} - \frac{1}{1+a^{d-1}} - \frac{a^d}{1+a^d} \\ &= \frac{1}{1+a^d} + \frac{1}{1+a^{d-2}} - \frac{2}{1+a^{d-1}} > 0 \end{aligned}$$

as $\frac{1}{1+a^x}$ is strictly convex in x for $x > 0$. It is clear the payoff also increases in (C2) with this change. Thus the payoff can always be increased by folding down such "peaks" if they exist. This eventually leads us to the greedy policy where $\Delta(t) \leq 1$ at all times. ☐

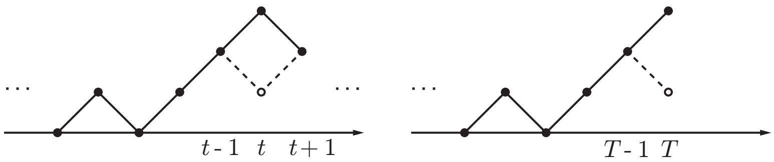


Figure 2.1: The change of policy in the two-channel scenario.

Theorem II.7. *The greedy policy is optimal for the finite-horizon problem (2.1).*

2.2.2 Against Adaptive Attack with Unknown Patterns

In the second case, we consider the more realistic scenario when the user has no such information on the attacker's behavior, and both sides then fight in the dark. Because of the informational symmetry between the two sides, we assume the user also behaves as an online learner with no-regret guarantee, and we are interested in the asymptotic interaction.

Suppose an arbitrary pair of policies g_w and g_u of the user and the attacker that are determined by a pair of no-regret learning algorithms. Utilizing the no-regret property as in the proof of Theorem II.5, it trivially follows that $\bar{r}_w(g_w, g_u, T) \geq \frac{n-1}{n} + o(1)$ and $\bar{r}_u(g_w, g_u, T) \geq \frac{1}{n} + o(1)$, where the $o(1)$ terms are w.r.t. the growth of T. Note that $\bar{r}_w(g_w, g_u) + \bar{r}_u(g_w, g_u) = 1$ for any g_w and g_u. We hence reach the conclusion that any pair of no-regret learning algorithm are *mutually best responses* for the infinite-horizon problem, and up to a diminishing term over a finite horizon. We note that our jamming game is of a zero-sum (constant-sum) nature, and the above result is essentially a restatement of the known convergence of learning limit to Nash equilibrium (NE) in zero-sum games (Chapter 4, [58]). Furthermore, it is clear that in this case there is no loss for the user compared to the baseline when it knows the adaptive pattern of the adversary and responses optimally.

2.3 The Decoy Dilemma

The previous results on the average payoff of the user against a learning attacker are in essence negative. The user cannot effectively escape the curse by the strong performance guarantee blessed to the attacker, even with more information. Therefore, we further consider the situation when there are additional resources available to the user to enhance the defense. In particular, we consider the one when the use of a decoy is viable. A *decoy* by the user is a device capable of performing similar legitimate operations as the user, and indistinguishable to the attacker (i.e., a double). For example, the decoy can be a regular but much cheaper transceiver, one without the ability to receive or perform channel switching. Intuitively, the introduction of a decoy can artificially create the impression of a "most user-active" channel so as to attract a majority of the attacks, thereby allowing the user to perform "under the radar" in a channel less likely to be jammed.

Indeed, this idea can be immediately verified in the infinite-horizon problem for the known case. Define a greedy decoy policy by letting the decoy and the user respectively select the channels with the highest and the lowest probabilities (the worst and the best channels) to be attacked. This policy causes the decoy to persistently transmit in one channel, and the user to use other channels in a round-robin fashion. It is easy to show that $\bar{r}_w(g_w, g_u) = 1$ if the attacker's policy g_u is given by a no-regret learning algorithm, and the user's policy g_w is the above greedy decoy policy. This asymptotic performance is optimal and less careful schemes can result in much inferior gain. For example, if the user and the decoy respectively select the best and the second best channels in each time slot, we have $\bar{r}_w(g_w, g_u) = \frac{n-1}{n}$.

In the unknown case, the lack of prior knowledge on the attacker makes using the decoy as a camouflage more difficult. Meanwhile, it is interesting to observe that if the most user-active channel is unique and *dominantly*, that is, there exists a unique channel k, such that for any subsequence of time slots $\{t_1, t_2, \dots, t_{\tau(T)}\} \subseteq \{1, 2, \dots, T\}$ with $\tau(T) = \Theta(T)^2$,

$$\liminf_{T\to\infty} \frac{1}{\tau(T)} \sum_{i=1}^{\tau(T)} \mathbf{1}_{i_{t_i}=k} > \limsup_{T\to\infty} \frac{1}{\tau(T)} \sum_{i=1}^{\tau(T)} \mathbf{1}_{i_{t_i}=j}$$

for any $j \neq k$, where i_{t_i} is the index of channel chosen by the user at time t_i, then the attacker can guarantee sublinear regret (uniformly or asymptotically) if and only if all suboptimal channel are chosen with time sublinear in T asymptotically. In other words, a strategy that guarantees sublinear regret for the attacker must ultimately identify and aim for the dominantly user-active channel if any. Therefore, the user can always use the decoy to "create" this dominant channel while performing operations in a virtually jamming-free environment, by letting the decoy reside in one channel and using a no-regret algorithm on the rest $n-1$ channels. This will result in the asymptotic optimal average reward, the same as in the case when the adversarial behavior is known.

Embedded in this observation is an interesting dilemma that the attacker faces in the presence of the *possibility* of a decoy that it cannot distinguish. On one hand, if the attacker adopts a no-regret algorithm like Hedge, arguably the best class of algorithms to use under uncertainty, then it is setting itself up for a very effective decoy defense by the evader, so much so that its attacker is rendered useless (asymptotically). This is the point illustrated above. On the other hand, if for this reason the attacker decides not to use such algorithms, then it may face a worse outcome as the alternative algorithm may provide no performance/regret guarantee. In this sense the mere

[2] $\tau(T) = \Theta(T)$ if for any T_0, there exists positive constants C_1 and C_2 such that $C_1 \cdot T \leq \tau(T) \leq C_2 \cdot T$ for any $T > T_0$.

possibility or threat of using a decoy may be viewed as effective defense.

2.4 Concluding Remarks

In the presentation of the previous results, we have implicitly relied on the following assumptions:

1. The attacker can only jam a single channel in each round.
2. The payoff of the user (and accordingly the attacker) is time-invariant and channel-independent.

The first assumption above can be trivially relaxed in light of existing no-regret learning algorithms for multiple plays (see e.g. [54] for an immediate extension of Hedge based on [59]), which provide decision policies to the attacker when it is capable of jamming multiple channels simultaneously, and we refer to [54] for a detailed analysis on jamming defense against multiple attacks. We have seen that a decoy can asymptotically negate the learning ability of the attacker in the last section. Conversely, the effect of the decoy can also be neutralized if multiple attacks are feasible. Hence, a potential arm race may arise between the user and the attacker, and the eventual outcome of this interaction then depends on the relative strength of the two sides.

In contrast to the second implicit assumption in this chapter, a time-varying and channel dependent payoff would allow us to model, for example, the temporal and spectral diversity in general communication systems. We provide in [54] a parallel discussion for channel-dependent but time-invariant payoffs. On the other hand, the discussion on time-varying but channel-independent payoffs can be found in [55], where the time-varying but channel-independent payoffs are results of transmission power control by the user on physically symmetric and stationary channels. In principle, the two aspects in the joint control, namely channel selection and transmission power, are coupled for optimal control w.r.t. properly defined objectives. Interestingly, we show in [55], when confronting an online learning attacker as in this chapter, that the optimal transmission power control can be decoupled from the optimal channel switching, or so-called "one-way" decoupling of the joint control. Also, the overall joint control problem can be reduced to a rate maximization problem with the tuning transmission power only. This decoupling is an immediate result following from the no-regret property of learning algorithm as we have seen in this chapter.

CHAPTER III

Optimal Channel Switching as Jamming Defense - Part II: Against a Resource-Replenishing Attacker with Minimax Optimality

In Chapter II, the regret is considered as the performance metric against an adversary with unknown behavioral pattern, and the optimality is aligned to the no-regret criterion. In this chapter, we revisit the jamming defense problem without prior knowledge on the attacker, and we instead consider the minimax optimality for the worst-case optimal response. Moreover, we assume the attacker is subject to a resource constraint with a replenishing process. That is, each action taken by the attacker consumes an amount of its resource, which may be later replenished. This models the power consumption of the attacker, which typically uses high power to generate jamming signals, while its energy can be recharged via energy harvesting from the environment or wired charging. Note that this model assumes the unconstrained attacker as a trivial case, where the repeated interaction can be reduced to a one-stage zero-sum game.

3.1 Problem Formulation and Preliminaries

Our problem formulation is within the framework as the previous problem in Chapter II, and different details are highlighted as follows. Recall that we denote by $[n] := \{1, 2, \ldots, n\}$ the action space of the user (i.e., the indexed set of channels), and we define $[n]^0 := \{0, 1, \ldots, n\}$ to be the action space of the attacker, where 0 is the null action (i.e., initiating no attack). Assume that the attacker has a finite amount of resource s_t at time t, and any non-null action by the attacker consumes a certain amount of resource; it also obtains replenishment after a round. In particular, given s_t and j_t, the latter of which denotes the attacker's action at t, the resource of the attacker at $t+1$ is given by $s_{t+1} = f_t(s_t, j_t)$, where f_t is a mapping summarizing the consumption and the replenishment process depending on the application scenario. An adversarial action j_t is feasible at time t given s_t if the causality condition $f_t(s_t, j_t) \geq 0$ holds. We denote by $\mathcal{F}_t(s_t) = \{j \in [n]^0 : f_t(s_t, j) \geq 0\}$ the feasible action set of the attacker at t. Let $\mathcal{S}_t := \{s : \mathcal{F}_t(s) = [n]^0\}$, i.e., $\mathcal{S}_t$ is the set of all resource levels such that all actions in $[n]$ are feasible for the attacker. We assume the payoff matrix of the user is given by $M_w = -M := -\begin{bmatrix} \mathbf{0} & \hat{M} \end{bmatrix}$, where $\hat{M} \in \mathbb{R}_+^{n \times n}$ is called the *loss matrix* and M the *augmented loss matrix*, and $\mathbf{0}$ is the zero column vector of length n. We denote by Δ_n and Δ_n^0 the spaces of probability distributions on $[n]$ and $[n]^0$. Given any vector $v = (v_1, v_2, \ldots, v_n)$, define $\text{supp}(v) = \{i : v_i > 0\}$. Define then $\Delta_n^0(s,t) = \{u \in \Delta_n^0 : \text{supp}(u) \subseteq \mathcal{F}_t(s)\}$, which is the set of distributions over feasible actions in $[n]^0$. As in the previous problem, the user chooses a distribution $w^t \in \Delta_n$ in round t, and a choice of channel $i_t \in [n]$ is realized per w^t independently. The attacker chooses a distribution $u^t \in \Delta_n^0(s_t, t)$ based on its resource s_t, and an adversarial action $j_t \in [n]^0$ is realized per u^t independently. After consumption and replenishment, the attacker's available resource for the next round is given by $s_{t+1} = f_t(s_t, j_t)$.

Throughout this chapter, we make the following assumptions.

Assumption III.1.

1. *The user has perfect recall of all its past actions and the observed adversarial actions.*

2. *The user knows the initial budget s_1 of the attacker and the resource dynamics f_t for $t = 1, 2, \ldots, T$.*

The second assumption on the knowledge of the attacker's initial budget is unrealistically strong. However, as we shall show that assuming an arbitrary initial budget of the attacker has no impact on the asymptotic worst-case payoff of the user.The goal of the user is to choose w^t for each round so as to minimize the expected total loss against all distributions over adversarial actions in a certain space, which we shall specify shortly. The strategy of the user can be made either online or offline, and in general it can be summarized by a contingency plan described as follows. At time t, the history of the above game consists of all past actions taken by the user and the attacker before time t, and the resource levels up to time t. Let a realization of the history at t be $h_t = \langle s_1, i_1, j_1, s_2, i_2, j_2, \ldots, s_{t-1}, i_{t-1}, j_{t-1}, s_t \rangle$ with $h_1 = s_1$, and we denote the set of all possible realizations of the history by $\mathcal{H}$. The information state of the user at time t is then exactly given by the history of the game, i.e., $\mathcal{I}_w^t = h_t$. Hence, the user's policy can be given by a mapping $w : \mathcal{H} \to \Delta_n$, where we change the notation from g_w to w for typographical simplicity, and we denote the space of all such mappings/policies as $\mathcal{G}$. We adopt the following notion of a (strong) attacker. It chooses the distribution of adversarial actions following a mapping $u : \mathcal{H} \to \Delta_n^0$ such that $u(h_t) \in \Delta_n^0(s_t, t)$ given the realization of history h_t up to time t; we denote the space of all such mappings as $\mathcal{G}^0$. The user's objective is to minimize the worst-case loss:

$$\min_{w \in \mathcal{G}} \max_{u \in \mathcal{G}^0} \mathbb{E} \left\{ \sum_{t=1}^{T} w(h_t)^\top M u(h_t) \right\}. \tag{3.1}$$

Note that we could also consider a weaker attacker, who chooses adversarial actions according to a mapping $u : \mathbb{R} \times \mathbb{N} \to \Delta_n^0$ such that $u(s_t, t) \in \Delta_n^0(s_t, t)$ given the resource level s_t at time t. As we show later, this weak attacker can be as damaging as a strong attacker in the context of the above decision problem for the user.

While we have not explicitly stated whether the attacker is strategic, the minimax formulation means that the user shall treat the attacker as strategic. More specifically, let $W(w,u) = -\mathbb{E}\left\{\sum_{t=1}^{T} w(h_t)^\top M u(h_t)\right\} = -U(w,u)$, and consider a zero-sum strategic game G where the two players are respectively given the strategy spaces $\mathcal{G}$ and $\mathcal{G}^0$ with the payoff functions W, U. Then, a Nash equilibrium (NE) strategy for player 1 (the user) in G is exactly a minimaximizer of (3.1). The above game theoretical interpretation of (3.1) regards the entire rounds of interaction as a one-shot game. On the other hand, the sequential interaction between the two players results in an extensive game Γ with simultaneous moves, where any realization of the history labels a particular node in the game tree. There exists at least one subgame perfect equilibrium (SPE) for Γ [60], which is also a NE of G. Hence, the minimaximizer of (3.1) exists. We denote by (w^*, u^*) a pair of SPE (or simply equilibrium when there is no ambiguity) strategies for the user and the attacker, and this pair will also be called an optimal solution to (3.1); in particular w^* is a minimaximizer to (3.1) and u^* a corresponding maximizer given w^*. Also, note that the pair (u^*, w^*) is a solution to the problem

$$\max_{u\in\mathcal{G}^0} \min_{w\in\mathcal{G}} \mathbb{E}\left\{\sum_{t=1}^{T} w(h_t)^\top M u(h_t)\right\}. \tag{3.2}$$

For technical reasons, we will consider a slightly perturbed version of problem (3.1) as an intermediate step in our analysis. Let $\epsilon : \mathbb{R} \to \mathbb{R}_+$ be a strictly increasing function parameterized by $\epsilon_{\max}$, such that $\epsilon(s) \le \epsilon_{\max}$ for all $s \in \mathbb{R}_+$, where $\epsilon_{\max}$ is a predetermined constant. The perturbed problem is then given by

$$\min_{w\in\mathcal{G}} \max_{u\in\mathcal{G}^0} \mathbb{E}\left\{\sum_{t=1}^{T} w(h_t)^\top M u(h_t) + \epsilon(s_T)\right\}. \tag{3.3}$$

With a similar argument used for (3.1), we can show an minimax-optimal solution exists for (3.3), which coincides with the SPE of the extensive game induced by (3.3). For the perturbed problem, we will inherit all the notation from (3.1), e.g., w^* is an optimal solution to (3.3). We note that if (w^*, u^*) is a solution to (3.3), the resulting loss in (3.1) is at most $\epsilon_{\max}$ more than the optimal minimax loss, and a similar result holds for (3.2), as shown in the following lemma.

Lemma III.2. *Let*

$$\ell(w) := \max_{u \in \mathcal{G}^0} \mathbb{E}\left\{\sum_{t=1}^{T} w(h_t)^\top M u(h_t)\right\}$$

and

$$g(u) := \min_{w \in \mathcal{G}} \mathbb{E}\left\{\sum_{t=1}^{T} w(h_t)^\top M u(h_t)\right\}.$$

If $\hat{w}^*$ and $\hat{u}^*$ are respective optimal solutions to (3.1) *and* (3.2)*, then $\ell(w^*) \leq \ell(\hat{w}^*) + \epsilon_{\max}$ and $g(u^*) \geq g(\hat{u}^*) - \epsilon_{\max}$.*

Proof. For the first inequality, we have

$$\begin{aligned}
\ell(w^*) &\leq \max_{u \in \mathcal{G}^0} \mathbb{E}\left\{\sum_{t=1}^{T} w^*(h_t)^\top M u(h_t) + \epsilon(s_T)\right\} \\
&= \min_{w \in \mathcal{G}} \max_{u \in \mathcal{G}^0} \mathbb{E}\left\{\sum_{t=1}^{T} w(h_t)^\top M u(h_t) + \epsilon(s_T)\right\} \\
&\leq \max_{u \in \mathcal{G}^0} \mathbb{E}\left\{\sum_{t=1}^{T} \hat{w}^*(h_t)^\top M u(h_t) + \epsilon(s_T)\right\} \\
&\leq \max_{u \in \mathcal{G}^0} \mathbb{E}\left\{\sum_{t=1}^{T} \hat{w}^*(h_t)^\top M u(h_t)\right\} + \epsilon_{\max} = \ell(\hat{w}^*) + \epsilon_{\max}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
g(u^*) &= \max_{u \in \mathcal{G}^0} \min_{w \in \mathcal{G}} \mathbb{E}\left\{\sum_{t=1}^{T} w(h_t)^\top M u(h_t) + \epsilon(s_T) - \epsilon(s_T)\right\} \\
&\geq \min_{w \in \mathcal{G}} \mathbb{E}\left\{\sum_{t=1}^{T} w(h_t)^\top M \hat{u}^*(h_t) + \epsilon(s_T)\right\} - \epsilon_{\max}
\end{aligned}$$

$$\geq \min_{w\in\mathcal{G}} \mathbb{E}\left\{\sum_{t=1}^{T} w(h_t)^\top M\hat{u}^*(h_t)\right\} - \epsilon_{\max} = g(\hat{u}^*) - \epsilon_{\max}.$$

□

We proceed with the following assumptions.

Assumption III.3.

1. *f_t is strictly increasing in the first argument.*

2. *$f_t(s, 0) > f_t(s, i)$ for all $i \in [n]$ and $s \geq 0$.*

The next lemma states in searching for the optimal strategy of the user we can limit our attention to a space smaller than $\mathcal{G}$; similarly, we can reduce the search space for the attacker's strategy. In fact, it can be reduced to that of a weak attacker as defined earlier. Let $\tilde{\mathcal{G}} := \{w \in \mathcal{G} : w(h_t) = w(h'_t),\ \text{if } s_t = s'_t, \forall t\}$ and let $\tilde{\mathcal{G}}_t := \{w \in \mathcal{G} : w(h_\tau) = w(h'_\tau),\ \text{if } s_\tau = s'_\tau, \forall \tau \geq t\}$, hence $\tilde{\mathcal{G}} = \tilde{\mathcal{G}}_1$. Similarly, we define $\tilde{\mathcal{G}}^0$ as a subset of $\mathcal{G}^0$ and $\tilde{\mathcal{G}}^0_t$.

Lemma III.4.

$$\min_{w\in\mathcal{G}} \max_{u\in\mathcal{G}^0} \mathbb{E}\left\{\sum_{t=1}^{T} w(h_t)^\top M u(h_t) + \epsilon(s_T)\right\}$$
$$= \min_{w\in\tilde{\mathcal{G}}} \max_{u\in\tilde{\mathcal{G}}^0} \mathbb{E}\left\{\sum_{t=1}^{T} w(h_t)^\top M u(h_t) + \epsilon(s_T)\right\}.$$

Proof. We show any optimal strategy can be replaced by a strategy in $\tilde{\mathcal{G}}$ without loss of optimality, and we prove by backward induction on the step of replacement. At T, given any h_T, w^* and u^* solve the following problem

$$\min_{w\in\mathcal{G}} \max_{u\in\mathcal{G}^0} w(h_T)^\top M u(h_T) + \epsilon(s_T).$$

Let $w^T \in \Delta_n$ and $u^T \in \Delta_n^0(s_T, T)$ be a NE of the one-stage matrix game with the respective payoff matrix $-M$ and M for the user and the adversary. Clearly, a pair of NE strategies only depend on s_T. Replacing $w^*(h_T)$ and $u^*(h_T)$ with w^T and u^T for all h_T, we obtain an alternative optimal solution in $\tilde{\mathcal{G}}_T$ and $\tilde{\mathcal{G}}_T^0$.

Assume $w^* \in \tilde{\mathcal{G}}_{t+1}$ and $u^* \in \tilde{\mathcal{G}}_{t+1}^0$. Given h_t at t, (w^*, u^*) solves

$$\begin{aligned}
&\min_{w\in\mathcal{G}} \max_{u\in\mathcal{G}^0} \mathbb{E}\left\{\sum_{\tau=t}^{T} w(h_\tau)^\top M u(h_\tau) + \epsilon(s_T) \;\middle|\; h_t\right\} \\
&= \min_{w\in\tilde{\mathcal{G}}_{t+1}} \max_{u\in\tilde{\mathcal{G}}_{t+1}^0} w(h_t)^\top M u(h_t) + \mathbb{E}\left\{\mathbb{E}\left\{\sum_{\tau=t+1}^{T} w(h_\tau)^\top M u(h_\tau) + \epsilon(s_T) \;\middle|\; h_{t+1}\right\} \;\middle|\; h_t\right\} \\
&= \min_{w\in\tilde{\mathcal{G}}_{t+1}} \max_{u\in\tilde{\mathcal{G}}_{t+1}^0} w(h_t)^\top M u(h_t) + \mathbb{E}\{W_{t+1}(w, u, s_{t+1}) | h_t\} \qquad (3.4) \\
&= \min_{w\in\tilde{\mathcal{G}}_{t+1}} \max_{u\in\tilde{\mathcal{G}}_{t+1}^0} w(h_t)^\top M u(h_t) + \sum_{j=0}^{n} u_j(h_t) W_{t+1}(w, u, f_t(s_t, j))
\end{aligned}$$

where in (3.4) we let $W_{t+1}(w, u, s_{t+1}) := \mathbb{E}\{\sum_{\tau=t+1}^{T} w(h_\tau)^\top M u(h_\tau) + \epsilon(s_T) \mid h_{t+1}\}$ since $w(h_\tau)$ and $u(h_\tau)$ only depend on s_τ for all $\tau \geq t+1$ for all $w, u \in \tilde{\mathcal{G}}_{t+1}$. Replacing respectively $w^*(h_t')$ and $u^*(h_t')$ by $w^*(h_t)$ and $u^*(h_t)$ for all $h_t' \in \mathcal{H}$ such that $s_t' = s_t$, we obtain an alternative optimal solution in $\tilde{\mathcal{G}}_t$ and $\tilde{\mathcal{G}}_t^0$, thus completing our proof. □

This result shows that actions in an optimal strategy can be identical for any two nodes in the game tree labeled by h_t and h_t' as long as $s_t = s_t'$ (i.e., Markovian in terms of s_t). Hence, we can reduce the representation of the label of node from the full history h_t to a two-tuple (s_t, t). With slight abuse of notation, we denote $w(h_t)$ as $w(s_t, t)$ for all $w \in \tilde{\mathcal{G}}$, and denote by $(w^*, u^*) \in \tilde{\mathcal{G}} \times \tilde{\mathcal{G}}^0$ an optimal solution to (3.3). We will refer to a subgame rooted at a node labeled by (s_t, t) as a subgame (s_t, t), and we define the payoff of a subgame (s_t, t) for the attacker using u^* provided w^* as

$$U_t^*(s_t) := \mathbb{E}\left\{\sum_{\tau=t}^{T} w^*(s_\tau, \tau)^\top M u^*(s_\tau, \tau) + \epsilon(s_T) \;\middle|\; s_t\right\}$$

$$= \max_{u\in\tilde{\mathcal{G}}^0} \min_{w\in\tilde{\mathcal{G}}} \mathbb{E}\left\{ \sum_{\tau=t}^{T} w(s_\tau,\tau)^\top M u(s_\tau,\tau) + \epsilon(s_T) \;\middle|\; s_t \right\}.$$

Using the perturbation term, we next show the monotonicity of U_t^*.

Lemma III.5. *$U_t^*(s_t)$ is strictly increasing in s_t for all t.*

Proof. We prove by induction. Given s_T at time T, we have

$$U_T^*(s_T) = w^*(s_T,T) M u^*(s_T,T) + \epsilon(s_T) = v(s_T) + \epsilon(s_T)$$

where $v(s_T)$ is the value of the last stage game, i.e.

$$v(s_T) := \max_{u\in\Delta_n^0(s_T,T)} \min_{w\in\Delta_n} w^\top M u.$$

Since $\Delta_n^0(s,t) \supseteq \Delta_n^0(s',t)$ for all t if $s > s'$, $v(s_T)$ is increasing in s_T, and hence $U_T^*(s_T)$ is strictly increasing in s_T. Assume the monotonicity for $t+1$. Let M_i be the $(i+1)$-th column of M (the j-th column of $\hat{M}$), and let $s_t > s_t'$. We also use w_i or u_i to denote the i-th or the $(i+1)$-th coordinate (function) of $w \in \Delta_n$ and $w \in \tilde{\mathcal{G}}$, or $u \in \Delta_n^0$ or $u \in \tilde{\mathcal{G}}^0$. Then,

$$\begin{aligned}
U_t^*(s_t) &= w^*(s_t,t)^\top M u^*(s_t,t) + \sum_{i=0}^{n} u_i^*(s_t,t) U_{t+1}^*(f_t(s_t,i)) \\
&= \max_{u\in\Delta_n^0(s_t,t)} \min_{w\in\Delta_n} \left\{ w^\top M u + \sum_{i=0}^{n} u_i U_{t+1}^*(f_t(s_t,i)) \right\} \\
&\geq \min_{w\in\Delta_n} \left\{ w^\top M u^*(s_t',t) + \sum_{i=0}^{n} u_i^*(s_t',t) U_{t+1}^*(f_t(s_t,i)) \right\}.
\end{aligned}$$

Note that $F(w) := w^\top M u^*(s_t',t) + \sum_{i=0}^{n} u_i^*(s_t',t) U_{t+1}^*(f_t(s_t,i))$ is a linear function of w, and hence the minimum is attained at some $\tilde{w} \in \Delta_n$. Also, note that for any

$i \in \operatorname{supp}(u^*(s'_t, t))$, we have $f_t(s_t, i) > f_t(s'_t, i) \geq 0$. Hence,

$$\begin{aligned}
U_t^*(s_t) \geq F(\tilde{w}) &> \tilde{w}^\top M u^*(s'_t, t) + \sum_{i=0}^n u_i^*(s'_t, t) U_{t+1}^*(f_t(s'_t, i)) \\
&\geq \min_{w \in \Delta_n} \left\{ w^\top M u^*(s'_t, t) + \sum_{i=0}^n u_i^*(s'_t, t) U_{t+1}^*(f_t(s'_t, i)) \right\} \\
&= w^*(s'_t, t)^\top M u^*(s'_t, t) + \sum_{i=0}^n u_i^*(s'_t, t) U_{t+1}^*(f_t(s'_t, i)) = U_t^*(s'_t),
\end{aligned}$$

which completes the induction. □

With the above preliminary results, we proceed in the next section to analyze the optimal response from the user.

3.2 Channel Switching for Minimax Optimality

In this section, we assume $\hat{M} = \operatorname{diag}(c_1, c_2, \ldots, c_n)$. This corresponds to the loss induced by a binary collision model, and the generalization is discussed in Section 3.3. We present the following main results.

1. The optimal strategy of the user in the perturbed problem (3.3) is to optimally respond to an attacker, who (a) either takes the null action with probability one or takes action i with probability $q_i := \frac{1/c_i}{\sum_{j=1}^n 1/c_j}$ for all $i \in [n]$ when $s_t \in \mathcal{S}_t$, and (b) takes the null action with probability one when $s_t \notin \mathcal{S}_t$.

2. Under additional conditions on the resource dynamics $f_t, t = 1, 2, \ldots, T$, the optimal strategy of the user in problem (3.1) is to optimally respond to an attacker, who (a) randomizes independently and identically at each round and takes action i with probability q_i when $s_t \in \mathcal{S}_t$, and (b) takes the null action with probability one when $s_t \notin \mathcal{S}_t$.

We will refer to the first part of the above claim as the basic characterization, and the second part as the characterization with structure on the replenishment. We also study the asymptotic average worst-case cost of the user applying the optimal strategy at the end of this section.

3.2.1 Basic Characterization

We proceed with a series of characterization on the optimal strategy as shown in the following lemmas.

Lemma III.6. *Any SPE strategy* $u^* \in \tilde{\mathcal{G}}^0$ *for the attacker is such that either* $u_0^*(s_t, t) = 1$ *or* $\operatorname{supp}(u^*(s_t, t)) \supseteq [n]$.

Proof. Let $(w^*, u^*) \in \tilde{\mathcal{G}} \times \tilde{\mathcal{G}}^0$ be a pair of SPE strategies. Assume that $u_0^*(s_t, t) < 1$, and let $\mathcal{N} := [n] - \operatorname{supp}(u^*(s_t, t))$. If $\mathcal{N} \neq \emptyset$, then $\operatorname{supp}(w^*(s_t, t)) \subseteq \mathcal{N}$, i.e., $\operatorname{supp}(w^*(s_t, t)) \cap \operatorname{supp}(u^*(s_t, t)) = \emptyset$. Otherwise, the payoff of any subgame (s_t, t) for the user using w^* provided u^*, which is given by

$$\begin{aligned} W_t^*(s_t) &:= \mathbb{E}\left\{ -\sum_{\tau=t}^{T} w^*(s_\tau, \tau)^\top M u^*(s_\tau, \tau) - \epsilon(s_T) \;\middle|\; s_t \right\} \\ &= \sum_{i=1}^{n} w_i^*(s_t, t)(-u_i^*(s_t, t)c_i + \sum_{j=0}^{n} u_j^*(s_t, t) W_{t+1}^*(f_t(s_t, j))) \\ &= -\sum_{i=1}^{n} w_i^*(s_t, t) u_i^*(s_t, t) c_i + \sum_{j=0}^{n} u_j^*(s_t, t) W_{t+1}^*(f_t(s_t, j)), \end{aligned} \tag{3.5}$$

can be strictly improved by reallocating the probability mass on any action $i \in \operatorname{supp}(u^*(s_t, t))$ to an action $j \in \mathcal{N}$.

Then, we have

$$\begin{aligned} U_t^*(s_t) &= u_0^*(s_t, t) U_{t+1}^*(f_t(s_t, 0)) + \sum_{i=1}^{n} u_i^*(s_t, t)(w_i^*(s_t, t)c_i + U_{t+1}^*(f_t(s_t, i))) \qquad (3.6) \\ &= \sum_{i \in \operatorname{supp}(u^*(s_t,t))} u_i^*(s_t, t) U_{t+1}^*(f_t(s_t, i)) < U_{t+1}^*(f_t(s_t, 0)), \end{aligned}$$

where the last inequality is due to Assumption III.3 and Lemma III.5. Hence, the payoff of the attacker can be strictly improved by choosing the null action with probability one, which contradicts the fact that u^* is a SPE strategy. Therefore, $\text{supp}(u^*(s_t)) \supseteq [n]$. □

Lemma III.7. *For a pair of SPE strategies* (w^*, u^*)*, if* $\text{supp}(u^*(s_t, t)) \supseteq [n]$*, then* $\text{supp}(w^*(s_t, t)) = [n]$.

Proof. Without loss of generality, assume that $f_t(s_t, i) \geq f_t(s_t, j)$ for any $i \geq j > 0$. Assume that there exists $i_1 \in [n]$ such that $i_1 \notin \text{supp}(w^*(s_t, t))$. Since $U^*_{t+1}(f_t(s_t, 0)) > U^*_{t+1}(f_t(s_t, i_1))$, by reallocatng the probability mass to the null action, the attacker can strictly improve its payoff of any subgame (s_t, t), thus resulting in a contradiction. □

Lemma III.8. *Given any pair of SPE strategies* (w^*, u^*)*, then*

$$u^*_i(s_t, t) = q_i(1 - u^*_0(s_t, t))$$

for all $i \in [n]$*, and when* $u^*_0(s_t, t) < 1$*,*

$$w^*_i(s_t, t) = \frac{U^*_t(s_t) - U^*_{t+1}(f_t(s_t, i))}{c_i}.$$

Proof. For u^*, the result is trivial when $u^*_0(s_t, t) = 1$. Assuming $u^*_0(s_t, t) < 1$, we then have $\text{supp}(u^*(s_t, t)) \supseteq [n]$ by Lemma III.6, and thus $\text{supp}(w^*(s_t, t)) = [n]$ by Lemma III.7. Hence, referring to (3.5) by the indifference condition of equilibrium points, we have

$$u^*_i(s_t, t)c_i = u^*_j(s_t, t)c_j$$

for all $i, j \in [n]$. Therefore, $u^*_i(s_t, t) = q_i(1 - u^*_0(s_t, t))$. For w^*, referring to to (3.6), we have

$$w^*_i(s_t, t)c_i + U^*_{t+1}(f_t(s_t, i)) = U^*_t(s_t),$$

for all $i \in [n]$, and the result follows. □

Lemma III.9. *Let* (w^*, u^*) *be a pair of SPE strategies. If* $0 < u_0^*(s,t) < 1$ *for some* $s \in \mathcal{S}_t$ *and* t*, then there exists a strategy* $\tilde{u}$ *such that* $\tilde{u}_0(s,t) = 0$ *for all* $s \in \mathcal{S}_t$ *and* t*, and* $(w^*, \tilde{u})$ *is a pair of SPE strategies. The space of such strategies will be denoted by* $\mathcal{G}^\dagger$.

Proof. Assume that $0 < u_0^*(s_t, t) < 1$ for some $s_t \in \mathcal{S}_t$ and t. Then, by Lemma III.8 we have $\text{supp}(u^*) \supseteq [n]$ and $u_i^*(s_t, t) = \frac{1/c_i}{\sum_{j=1}^n 1/c_j}(1 - u_0^*(s_t, t))$ for all $i \in [n]$. Also,

$$U_t^*(s_t) = w_i^* c_i + U_{t+1}^*(f_t(s_t, i))$$

for all $i \in [n]$, where $U_{t+1}^*(f_t(s_t, i))$ only depends on $u^*(\cdot, \tau)$ and $w^*(\cdot, \tau)$ for all $\tau > t$. Consider an alternative strategy for the attacker such that $\tilde{u} = u^*$ except $\tilde{u}_0(s_t, t) = 0$ and $\tilde{u}_i(s_t, t) = \frac{1/c_i}{\sum_{j=1}^n 1/c_j}$. Referring to (3.5), we note that the continuation part (i.e., the second term) of the user's payoff of the subgame rooted at (s_t, t) is independent of the user's action at t, and note also the values of $\tilde{u}_i(s_t, t)c_i$ are equal among all $i \in [n]$. Hence, given $\tilde{u}$, the user has no incentive to deviate from w^*. On the other hand, the attacker's payoff of the subgame rooted at (s_t, t) using $\tilde{u}$ given w^* is

$$\sum_{i=1}^n \tilde{u}_i(s_t, t)(w_i^* c_i + U_{t+1}^*(f_t(s_t, i))) = U_t^*(s_t).$$

Therefore, $(w^*, \tilde{u})$ is a pair of SPE strategies. Repeating this argument on $\tilde{u}$ whenever necessary, we can obtain a SPE strategy of the attacker as described in the lemma. □

The above results are summarized in the first part of the claim at the beginning of this section, which we reproduce in the following theorem.

Theorem III.10. *Consider a policy of the attacker such that either* $u_i(s_t, t) = q_i := \frac{1/c_i}{\sum_{j=1}^n 1/c_j}$ *for all* $i \in [n]$ *or* $u_0(s_t, t) = 1$ *if* $s_t \in \mathcal{S}_t$ *and* $u_0(s_t, t) = 1$ *if* $s_t \notin \mathcal{S}_t$*. The optimal strategy of the user in problem* (3.3) *is to optimally respond to such an adversarial policy.*

The weakness of the above result is clearly the ambiguity in determining the "either-or" branch of the adversarial policy when $s_t \in \mathcal{S}_t$. In some applied instances, we can reason that the attacker would not use the null action in the perturbed problem and extend this conclusion to the original one, and hence obtain an explicit form of the user's optimal strategy. This illustrated using the following example.

Example. Assume that $f_t(s,i) < s$ for all $i \in [n]$, and $f_t(s,0) = s$, i.e., the attacker only has a finite budget of resource without replenishment. Let $\delta_{\min} := \inf_{s\geq 0, i\in[n]}(s - f_t(s,i))$, and assume that $\delta_{\min} > 0$. Let $T > \frac{s_1}{\delta_{\min}}$, i.e., the horizon is beyond the time that the attacker would exhaust its resource if it always takes a non-null action. Note that whenever $u_0(s_t,t) = 1$, the game is equivalently shortened by one time step. Hence, we can reduce the strategy space of the attacker to the set of strategies such that $u_0(s_t,t) = 0$ for all t whenever $s_t \in \mathcal{S}_t$. Then, we have $u_i^*(s_t,t) = q_i$ for all t whenever $s_t \in \mathcal{S}_t$. Hence, the attacker's equilibrium strategy is to identically and independently randomize before exhausting the resource. Note that u^* we obtained is independent from the perturbation parameter $\epsilon_{\max}$, and moreover, using Lemma III.2 we have $g(u^*) \geq g(\hat{u}^*) - \epsilon_{\max}$ for any $\epsilon_{\max} > 0$, where $g(\hat{u}^*)$ is the optimal value of (3.2). Hence, u^* is an optimal solution to (3.2), and an optimal strategy of the user in (3.1) is to optimally respond to this belief on the adversarial behavior. In particular, it has the structure shown in Lemma III.8 by setting the perturbation term to zero. Let $T(s_t)$ be the minimum time τ such that $s_\tau \notin \mathcal{S}_\tau$ given the resource level s_t at t, when the action i_τ taken by the attacker at each round $\tau \geq t$ is independent and identically distributed (i.i.d.) with the distribution $q = (q_i, i \in [n])$. Note that

$$\mathbb{E}T(s_t) = \mathbb{E}\left\{ \sum_{\tau=t}^{T} \mathbf{1}(s_\tau \in \mathcal{S}_\tau) \;\middle|\; s_t \right\},$$

and

$$\mathbb{E}\{w^*_{i_\tau}(s_\tau,\tau)c_{i_\tau} \mid s_t\} = \mathbb{E}\{w^*_{i_\tau}(s_\tau,\tau)c_{i_\tau} \mid s_\tau \in \mathcal{S}_\tau, s_t\} \cdot \mathbb{P}(s_\tau \in \mathcal{S}_\tau \mid s_t)$$
$$= \frac{1}{\sum_{j=1} 1/c_j}\mathbb{E}\{\mathbf{1}(s_\tau \in \mathcal{S}_\tau) \mid s_t\}.$$

Then,

$$U^*_t(s_t) = \mathbb{E}\left\{\sum_{\tau=t}^{T} w^*_{i_\tau}(s_\tau,\tau)c_{i_\tau} \;\middle|\; s_t\right\} = \frac{1}{\sum_{j=1} 1/c_j}\mathbb{E}\left\{\sum_{\tau=t}^{T}\mathbf{1}(s_\tau \in \mathcal{S}_\tau) \;\middle|\; s_t\right\}$$
$$= \frac{1}{\sum_{j=1} 1/c_j}\mathbb{E}T(s_t),$$

and the optimal strategy of the user is given by

$$w^*_i(s_t,t) = \frac{U^*_t(s_t) - U^*_{t+1}(f_t(s_t,i))}{c_i} = \frac{1}{\sum_{j=1} 1/c_j}\frac{\mathbb{E}T(s_t) - \mathbb{E}T(f_t(s_t,i))}{c_i}$$

before $T(s_t)$. In fact, this is the optimal strategy found by Abernethy and Warmuth constructively in [61].

3.2.2 Characterization with Structure on the Replenishment

As aforementioned, the difficulty of applying Theorem III.10 is that we have to determine whether the attacker chooses the null action with probability one even when all non-null actions are feasible. Intuitively, the only incentive for the attacker to take the null action in such cases is to save resources for a rainy day. However, this incentive goes away if it eventually takes a non-null action and the resource dynamics from that point on is the same had it switched the order of these two actions. This intuitive argument suggests that with more structure imposed on the resource dynamics $f_t, t = 1, 2, \ldots, T$, we may be able to conclude a more explicit form on the user's optimal strategy as shown in the above example. Indeed, we make the following assumption on the structure of the resource dynamics, and justify our previous conjecture in Lemma III.12.

Assumption III.11.

1. *$f_{t+1}(f_t(s,i),j) = f_{t+1}(f_t(s,j),i)$ for any $i, j \in [n]$ and all t .*

2. *For any $s \in \mathcal{S}_t$ and $t < T$, $f_t(s,0) \in \mathcal{S}_{t+1}$ and there exists $i \in [n]$ such that $f_t(s,i) \in \mathcal{S}_{t+1}$.*

Let v be the value of a stage game when all non-null actions are feasbile, i.e., $v := \min_{w\in\Delta_n} \max_{u\in\Delta_n^0} w^\top M u = \frac{1}{\sum_{j=1}^n 1/c_j}$, and let $q_{\min} := \min_{i\in[n]} q_i$. Set $\epsilon_{\max} < q_{\min} v$.

Lemma III.12. *If (w^*, u^*) is a pair of SPE strategies and $u^* \in \mathcal{G}^\dagger$, then $u_0^*(s,t) = 0$ for any $s \in \mathcal{S}_t$ and all t.*

For the sake of readability, the lengthy proof of the above lemma is placed in the appendix, and it proves the second part of our claim at the beginning of this section, which is repeated in the following theorem.

Theorem III.13. *Consider a policy of the attacker such that $u_i(s_t,t) = q_i := \frac{1/c_i}{\sum_{j=1}^n 1/c_j}$ for all $i \in [n]$ if $s_t \in \mathcal{S}_t$ and $u_0(s_t,t) = 1$ if $s_t \notin \mathcal{S}_t$. The optimal strategy of the user in problem* (3.1) *is to optimally respond to such an adversarial policy.*

Proof. Lemma III.12 directly proves the above claim for the perturbed problem (3.3). Using the same argument as shown in the example after Theorem III.10, we conclude that the described adversarial strategy is also an equilibrium strategy in the original problem (3.1), and the result follows. □

The optimal strategy of the user is then given as in Lemma III.8, where U^* can be similarly estimated using Monte-Carlo method as in [61].

3.2.3 Asymptotics

We next consider the average worst-case cost κ using the minimax optimal strategy, which is given by

$$\kappa := \limsup_{T\to\infty} \min_{w\in\mathcal{G}} \max_{u\in\mathcal{G}^0} \mathbb{E}\left\{\frac{1}{T}\sum_{t=1}^{T} w(h_t)^\top M u(h_t)\right\}. \tag{3.7}$$

In this part, we assume a stationary and linear resource replenishment process, that is,

$$f_t(s,i) = f(s,i) = s - d_i + \gamma$$

for all $i \in [n]^0$, where d_i is the resource cost of action i and γ is the resource replenishment rate. We assume $d_0 = 0$ and without loss of generality, suppose $0 = d_0 \leq d_1 \leq \ldots \leq d_n$. We also assume that $\gamma \geq d_1$. Hence, f satisfies Assumption III.3 and III.11. Let $s_{\text{th}} = \min\{d_n - \gamma, 0\}$, and then $\mathcal{S}_t = [s_{\text{th}}, \infty)$. Consequently, using Theorem III.13, we can regard the attacker as behaving randomly and taking action from $[n]$ with the probability distribution q whenever $s \geq s_{\text{th}}$, and choosing the null action with probability one if short of resource. Let S_t be the random process of the attacker's resource level. Let $X_t := \mathbf{1}(S_t \geq s_{\text{th}})$. Let $C_t \in \{c_1, \ldots, c_n\}$ be an i.i.d. process with $\mathbb{P}(C_t = c_i) = q_i$ for all i, and similarly we define a process $D_t \in \{d_1, \ldots, d_n\}$. Moreover, we assume that C_t and D_t are respectively independent from all X_s with $s < t$. Then, the resource dynamics can be written as

$$S_{t+1} = S_t - D_t X_t + \gamma,$$

and the average cost of the user is given by

$$\kappa = \limsup_{T\to\infty} \mathbb{E}\left\{\frac{1}{T}\sum_{t=1}^{T} C_t X_t\right\} = \limsup_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}C_t \mathbb{E}X_t$$

$$= \mathbb{E}C_t \cdot \rho = \frac{n}{\sum_{j=1}^{n} 1/c_i} \cdot \rho,$$

where $\rho := \limsup_{T\to\infty} \mathbb{E}\left\{\frac{1}{T}\sum_{t=1}^{T} X_t\right\}$. Note that S_t admits a stationary distribution (i.e. stable) if and only if $\mathbb{E}D_t > \gamma$. Indeed, consider the two auxiliary queues S'_t and S''_t that are given by $S'_{t+1} = S'_t - D_t + \gamma$, and $S''_{t+1} = \max\{S''_t - D_t, s_{\text{th}}\} + \gamma$. Then, $S'_t \le S_t \le S''_t$ and the two auxiliary queues are positive recurrent if and only if $\mathbb{E}D_t > \gamma$. When S_t is stable, we have

$$\begin{aligned} 0 &= \lim_{T\to\infty} \frac{1}{T}\left(T\gamma - \sum_{i=1}^{T} \mathbb{E}[D_t X_t]\right) \\ &= \lim_{T\to\infty}\left(\gamma - \mathbb{E}D_t \cdot \mathbb{E}\left\{\frac{1}{T}\sum_{t=1}^{T} X_t\right\}\right) = \gamma - \mathbb{E}D_t \cdot \rho. \end{aligned}$$

Hence, when S_t is stable, $\rho = \frac{\gamma}{\mathbb{E}D_t} = \frac{\gamma \sum_{j=1}^{n} 1/c_j}{\sum_{i=1}^{n} d_i/c_i} < 1$ and

$$\kappa = \frac{\gamma}{\frac{1}{n}\sum_{i=1}^{n} d_i/c_i} = \frac{\gamma}{\alpha},$$

where $\alpha := \frac{1}{n}\sum_{i=1}^{n} d_i/c_i$ can be interpreted as the average cost-gain ratio of adversarial actions. When $\mathbb{E}D_t < \gamma$, S_t grows unbounded and we have $\rho = 1$. Thus, $\kappa = \frac{n}{\sum_{j=1}^{n} 1/c_j}$ in this case, the harmonic mean of c_i's.

3.3 Concluding Remarks

In the previous section, we presented the minimax optimal strategy of the user when the cost matrix is assumed to be diagonal, which models binary collision. There are a number of open problems arising from the previous result.

3.3.1 Non-negative $\hat{M}$

Moving from the binary collision model to a more general interference model, we will need to revisit our problem with an arbitrary non-negative cost matrix. We observe that the theory we developed so far for the diagonal $\hat{M}$ applies trivially to the case when $\hat{M} = D + c\mathbf{1}\mathbf{1}^\top$, when D is a diagonal matrix, i.e., $\hat{M}_{ij}$ is a constant c for all off-diagonal entries, by simply noting that $w^\top \hat{M} u = w^\top (D + c\mathbf{1}\mathbf{1}^\top) u = w^\top D u + c$. A more interesting case that can be reduced to a diagonal one is when $\hat{M}$ is a multiple of a doubly-stochastic matrix Q, i.e., $\hat{M} = zQ$ for some $z > 0$. We proceed with the following fact.

Lemma III.14 ([62])**.** *If each row sum of a non-singular matrix is a constant z, then each row sum of its inverse matrix is $1/z$. The same applies to the column sums.*

Hence, $\hat{M}^{-1}\mathbf{1} = \mathbf{1}^\top \hat{M}^{-1} = 1/z$. Consider then the following construction. Let $\hat{D} = \operatorname{diag}^{-1}(\mathbf{1}^\top \hat{M}^{-1}) = zI$, and let $D = \begin{bmatrix} \mathbf{0} & \hat{D} \end{bmatrix}$. For any $\hat{u} \in \Delta_n^0$, let $u = K\hat{u}$ where $K := \begin{bmatrix} 1 & \\ & \hat{M}^{-1}\hat{D} \end{bmatrix}$. Then, $Mu = \begin{bmatrix} \mathbf{0} & \hat{M} \end{bmatrix} \begin{bmatrix} 1 & \\ & \hat{M}^{-1}\hat{D} \end{bmatrix} \hat{u} = D\hat{u}$. Let $\Theta = \{K\hat{u} \mid \hat{u} \in \Delta_n^0\}$. Note that $\hat{u} = K^{-1}u \in \Delta_n^0$ for any $u \in \Delta_n^0$. Hence, $\Theta \supseteq \Delta_n^0$. Consider a mapping $u : \mathcal{H} \to \Theta$, and denote the space of all such mappings as $\mathcal{A}_\Theta$. Let $V(w, u) := \mathbb{E}\left\{\sum_{t=1}^T w(h_t)^\top M u(h_t)\right\}$. Then,

$$\max_{u \in \mathcal{A}} \min_{w \in \mathcal{A}} V(w, u) \le \max_{u \in \mathcal{A}_\Theta} \min_{w \in \mathcal{A}} V(w, u) = \max_{\hat{u} \in \mathcal{A}} \min_{w \in \mathcal{A}} V(w, u).$$

For the problem on the right-hand side, our previous result implies that $\hat{u}_i^*(h_t) = q_i(1 - \hat{u}_0^*(h_t)) = \frac{1}{n}(1 - \hat{u}_0^*(h_t))$. Interestingly, $u(h_t) = K\hat{u}^*(h_t)$ is in fact equal to $\hat{u}^*(h_t)$. Hence, we obtain an optimal solution to the problem on the left-hand side. However, a natural interference model may not be captured by a doubly stochastic structure.

3.3.2 Conversion to a Gain Formulation

In this chapter, we focused on the loss formulation for the user instead of a gain perspective. The problem could be revisited with a gain matrix for the user, and the role of min-max would be exchanged for the user and the attacker. Unlike the loss formulation, we could have developed a theory of the adversarial channel capacity in the presence of a jammer, which would be in parallel to the asymptotic result presented in the previous section. The two formulations are intuitively equivalent in the sense that a gain formulation can always be converted to a loss one by setting the gain matrix as the difference between a multiple of the all-one matrix and a loss matrix. However, the solution technique requires the full characterization of optimal strategies with an arbitrary non-negative loss matrix, as stated in the first open problem. Moreover, we note a fundamental difference can be found in the rationale of decision for the user between the two formulations, which in turn suggests that the user's optimal strategy may be considerably different even for other categories of loss matrices compared to the results for diagonal-related ones. As shown before, the user would strictly prefer a channel that is not in the support of the attacker's strategy in the loss setting, so as to incur no cost. However, the user would prefer to risk using a channel on which the attacker puts positive probability mass, if the gain of this action is much higher than that of a jamming-free one, thus favorable in expectation.

CHAPTER IV

Throughput Optimal Channel Switching in Random Access - Part I: Intuition from Slotted Aloha

In the previous two chapters, we have considered the interplay between a legitimate user and an adversary with conflicting interest. In many applications, the adversary is often a competing peer user rather than a malicious entity, e.g. a jamming attacker. Also, in practical systems, it typically involves a large number of interacting users; the user behavior is typically regulated by the deployed protocol. From this chapter on, we hence will consider a group of users (or nodes interchangeably) competing for spectrum resources, and we consider the decentralized channel switching policies of users in a multi-channel wireless system with medium access regulated by a given protocol, so as to maximize the overall throughput of the network. In particular, we consider the possibility of promoting certain performance measures, e.g. throughput, via individual learning as a way to obtain decentralized implementations of desirable policies.

Motivated by the above, in the next two chapters we ask the question of what type of dynamic channel switching schemes will give the best performance in a multi-channel WLAN. This will be evaluated using the notion of stability region of a scheme.

This is because more effective resource allocation and sharing can achieve a lower overall congestion level, thus expanding the range of sustainable arrival rates and resulting in a larger stability region. The scheme with the largest such region is commonly known as the throughput optimal scheme. With this objective, we set out to study the stability region of a multi-channel WLAN system where users are allowed to dynamically switch between channels, and identify throughput optimal channel-switching schemes for the entire system.

Our ultimate goal is to develop a theory on the throughput optimal switching for networks regulated by the IEEE 802.11 DCF (Distributed Coordination Function), a widely deployed contention-based MAC layer protocol. This chapter aims to provide theoretical preparation and intuition for our later analysis in DCF, using the simpler slotted Aloha scheme. We formally address the question on throughput optimality in DCF in Chapter IV.

4.1 Preliminaries

4.1.1 Slotted Aloha and IEEE 802.11 DCF

Aloha [63, 64] is technically the prototype of a wide range of contention-based random access schemes including 802.11 DCF. An extensive discussion on several variants in the Aloha family of protocols can be found in [65]. We shall focus on the slotted Aloha scheme in this analysis. Unlike reservation-based schemes that utilize, for instance, time-sharing or frequency-sharing to coordinate the medium access, the "free-for-all" approach embedded in slotted Aloha aims to improve the throughput when the traffic is light or moderate, when collision is expected to be rare. Under slotted Aloha, the system operates in discrete time with a uniform length for each time slot. Furthermore, all packets are assumed to require exactly one slot for transmission. Users are assumed to be synchronized, and a user i attempts data transmission at the beginning of a time slot with probability τ_i if its queue of packets is non-empty. Note that this is a simplified version of slotted Aloha in the sense that any node is regarded as backlogged immediately upon packet arrival to an empty queue. If two or more nodes attempt transmissions in the same slot, a collision occurs and packets transmitted are lost (un-decodable); this is assumed to be observable to all nodes involved, and retransmissions take place with the same attempt rates in future slots.

Instead of using coin toss to determine whether a transmission should be attempted, the 802.11 DCF utilizes a process called backoff to further reduce potential collision. The basic idea is for nodes involved in a collision to select and start a random timer that counts down to zero; a node only attempts retransmission when the time expires. Upon each successive collision on the retransmission, a node selects the random timer value from an exponentially increasing range, with the intention that more nodes will back off for longer thereby avoiding further collision. More is detailed in Chapter V.

4.1.2 Stability and Throughput Optimality

The notion of stability throughout this chapter and the next one refers to the existence of limiting distribution of the queueing processes of packets at each user, when the packet queueing is appropriately defined. When the queues are modeled by (irreducible and aperiodic) Markov chains, the stability criterion will specifically be the existence of stationary distribution. The stability region is given by the set of vectors of data/packet arrival rates, such that all queues are stable. A throughput optimal channel switching policy, when the form of policy is appropriately defined, is one that results in a stability region that is the superset of that of any other policy under comparison, if it exists. Formal definitions of these notions will be reviewed in Section 5.2.

4.2 Decentralized Throughput Optimal Switching via Individual Learning

In this section, we consider the interaction in a multi-channel system governed by the simple slotted Aloha scheme, and apply existing results from population game to gain insight on the nature of throughput optimal switching policies. Our results will also shed light on possible decentralized implementations for more complex random access schemes like the DCF. As we shall see, starting from this simpler, slotted Aloha scheme will continue to offer insightful interpretation of our analysis in the setup of DCF later.

We will limit our attention to a homogeneous population of users with identical traffic loads and attempt rates. On the other hand we will model the asymmetry in channel bandwidth by adopting different real-time scales of time slots in different channels. The above setup is summarized as follows:

- The network operates in discrete slots $t = 1, 2, \ldots$. There are N users and K channels. A slot on channel k has a duration of α_k time units; a smaller numeric value of α_k models a physically faster channel in data rate.

- Each node has an infinite queue. Packets arrival according to a Poisson process with rate λ, independent from arrivals of other users and the medium access scheme. Hence, the number of arrivals to a node between two consecutive slot boundaries is a Poisson random variable with mean $\lambda_k = \alpha_k \lambda$ in channel k.

- At the boundary of a slot, whenever its queue is non-empty, a node attempts transmission of the packet at the head of the queue with probability τ. Multiple transmissions from different users result in a collision and all packets involved will remain in their queues.

This simple model captures some essential aspects of a network operating under

DCF, and we aim to gain some insights on the following two questions:

- What type of centralized static allocation policy is throughput optimal?
- Is it possible to reach/approximate the throughput optimal channel allocation via individual learning over repeated interactions?

4.2.1 Centralized throughput optimal policy

Given the homogeneity of the population, an allocation of channels can be given by a vector $(n_1, n_2, \ldots, n_K)$, where n_k denotes the number of users assigned to channel k with $\sum_{i=1}^{K} n_k = N$. Given any allocation of channels, it can be represented by the corresponding population state, which is defined by $x = (x_1, x_2, \ldots, x_K)$ where $x_k = \frac{n_k}{N}$. We denote the space of population states by $X_N = \Delta_K \cap \frac{1}{N}\mathbb{Z}^K$, where Δ_K is the simplex in $\mathbb{R}^K$. A centralized channel allocation policy is hence identified by a mapping $\phi : \mathbb{R}_+ \to X_N$ such that given a symmetric arrival rate λ for all users, it generates a population state.

Given a population state x, consider channel k and relabel the Nx_k users on this channel from 1 to Nx_k. Let Q_i^t be the queue length of node i on channel k at time t, and let $Q^t = (Q_1^t, Q_2^t, \ldots, Q_{Nx_k}^t)$. It is not hard to see that $\{Q^t\}_{t=1}^{\infty}$ is a multidimensional irreducible and aperiodic Markov chain with a countable state space $\mathcal{S} = \mathbb{N}^{Nx_k}$. We first note the following rather intuitive condition for the existence of a stationary distribution of $\{Q^t\}_{t=1}^{\infty}$, which follows from the argument in [66] for a homogeneous system.

Lemma IV.1. *The stationary distribution on channel k exists if and only if $\tau(1-\tau)^{Nx_k-1} > \lambda_k$.*

Hence, a symmetric arrival rate λ is stabilizable if there exists a population state such that the above stability condition holds for all channels. Formally, let $X_k(\lambda) = [0, \frac{1}{N}\log_{1-\tau}\frac{\alpha_k\lambda}{\tau} + \frac{1}{N})$, and let $X(\lambda) = \prod_{k=1}^{K} X_k(\lambda)$. The set of stabilizable symmetric arrival rates is then given by $\Lambda_N = \{\lambda > 0 : X(\lambda) \cap X \neq \emptyset\}$. Recall that a throughput optimal strategy, if it exists, generates a stability region that is the superset of that of any other strategy. In our setup, the throughput optimal strategy trivially exists, and for any stabilizable λ, its allocation can be defined as any population state that stabilizes the network. Λ_N is therefore the stability region of a throughput optimal policy in our setup.

Using the above result, we can immediately obtain a throughput optimal policy in an "almost explicit" form as follows. If $\lambda \in \Lambda$, it is then necessary that

$$\sum_{k=1}^{K} \frac{1}{N} \log_{1-\tau} \frac{\alpha_k \lambda}{\tau} + \frac{K}{N} \leq 1,$$

and $\hat{x} = (\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_K)$ with

$$\hat{x}_k = \frac{\frac{1}{N} \log_{1-\tau} \frac{\alpha_k \lambda}{\tau} + \frac{1}{N}}{\sum_{k=1}^{K} \frac{1}{N} \log_{1-\tau} \frac{\alpha_k \lambda}{\tau} + \frac{K}{N}} = \frac{\log_{1-\tau} \alpha_k + C(\lambda, \tau)}{\sum_{k=1}^{K} \log_{1-\tau} \alpha_k + KC(\lambda, \tau)}$$

where $C(\lambda, \tau) = \log_{1-\tau} \frac{\lambda}{\tau}$ is always in $X(\lambda)$ for a given λ. Then, for a sufficiently large N, a neighboring point of $\hat{x}$ on the grid X can represent the population state of a throughput optimal policy for a stabilizable λ. As $\hat{x}$ suggests, the throughput optimal policy achieves certain form of load balance among different channels. In particular, a channel with higher bandwidth, represented by a smaller α_k, will be assigned more users to leverage the resource. For the special case with physically symmetric channels, α_k is identical for all k, and the allocation policy evenly distributes users on all channels.

Beyond the above observation, Lemma IV.1 also provides a somehow implicit characterization of a throughput optimal policy, which will prove to be more useful for decentralized implementation, via an optimization problem. Specifically, a throughput optimal policy can also be given as one that solves the following optimization problem for any given λ:

$$\min_{x \in X_N} \max_{k \in [K]} \frac{\alpha_k}{(1-\tau)^{N x_k}}, \tag{4.1}$$

and $\lambda \in \Lambda_N$ if and only if the value of the above problem is no less than $\frac{\tau(1-\tau)}{\lambda}$. There are two interesting features of this minimax optimization formulation. We first observe that its solutions exhibit load balancing as the previous explicit solution. That is, there always exists a monotone optimal solution x^* to problem (4.1), such that $x_k^* \geq x_j^*$ whenever $\alpha_k \leq \alpha_j$. Also, note that $(1-\tau)^{N x_k - 1}$ is the probability that a transmission results in a collision when attempted, and the equivalent objective $\frac{\alpha_k}{(1-\tau)^{N x_k - 1}}$ of problem (4.1) physically represents the mean time that a node spends in collision in the entire service process of a packet. In other words, the objective is an indicator of local congestion level.

In the next part, we shall consider the decentralized implementation of a throughput optimal policy in the large population regime, based on the above features given by the formulation in (4.1). We shall consider certain scaling w.r.t. N. To obtain meaningful results, we assume $\tau = \frac{C_1}{N}$ and $\lambda = \frac{C_2}{N}$ for some positive constants C_1 and C_2. That is, the network would maintain a constant total arrival rate and the total number of transmission attempt is bounded on average.

4.2.2 Decentralized implementation via individual learning

The general strategy of our decentralized implementation is to realize a throughput optimal policy via incentive based learning by users, as users dynamically select channels to use. Our formulation and theoretical framework follows that of [67]. We first define a population game among the users, which is identified by a payoff function that specifies the utilities derived by users in each channel, given any population state. We then define the learning algorithm, or the so-called revision protocol [67], which describes a updating procedure followed by the user in making channel selections. The population game together with the learning algorithm then induces a deterministic and dynamic evolution in the mean change of the population state. A recurrent/stable point in this process under a large population regime, if exists, can be closely connected to the game theoretical equilibrium concept of the underlying population game. We then further establish the connection between the equilibrium point and a throughput optimal allocation policy.

An N-user population game with strategy space $[K]$ (the set of all channels)[1] is given by a payoff function $F^N : X_N \to \mathbb{R}^K$, where $F_k^N(x)$ is the payoff to any user selecting channel k when the population state is $x \in X_N$. In light of problem (4.1), we define

$$F_k^N(x) = -\frac{\alpha_k}{(1-\tau)^{Nx_k}}.$$

Recall the scaling $\tau = \frac{C_1}{N}$ and assume $C_1 = 1$ for simplicity in presentation. We define a continuous population game by $F : \Delta_K \to \mathbb{R}^K$, which is given by

$$F_k(x) = -\alpha_k e^{x_k},$$

and is Lipschitz continuous. The sequence of finite population games $\{F^N\}_{N=N_0}^{\infty}$ then converges uniformly to F, where the constant $N_0 > 1$.

We group time slots in each channel into epochs; an epoch on channel k consists of $\frac{L}{\alpha_k}$ slots, where L is a constant and we assume $\frac{L}{\alpha_k}$ to be an integer for all k so that the boundaries of epochs are aligned across channels.[2] At the boundary of an epoch, each user can update its choice of channel with probability $\frac{1}{N}$, and the update opportunities are statistically independent among users. When a public randomization device is available, alternatively one user can be randomly chosen from the population with equal probability to update. The following analysis and results also hold for this setup. The rationale is to have a bounded number of updates on

[1]Recall the notation $[K] := \{1, 2, \ldots, K\}$

[2]One can also assume that L is in addition chosen to be sufficient large, so that the payoff as the *average* time in collision given a population state would be reasonably measured/observed by the user as a mean-field effect. See the next chapter on the mean-field qualifier.

average in each epoch that does not scale w.r.t. N, and the number of updates on a channel is proportional to the local population size. Define the learning algorithm or the revision protocol as a continuous function $\rho : \mathbb{R}^K \times \Delta_K \to \mathbb{R}^{K\times K}$ that specifies the conditional switching probability $\rho_{kj}(F^N(x), x)$ from channel k to channel j when the population state is x and the payoff vector is given by $F^N(x)$. One can show that the images of all F_k^N and F can be contained in some common compact set $\mathcal{K}$ when $N > 1$. In the rest of this section, we will restrict the domain of ρ to $\mathcal{K} \times \Delta$. As a result ρ is uniformly continuous.

Given F^N and ρ, it is not hard to see that they induce a Markov chain $\{X_\ell^N\}_{\ell=0}^\infty$ on the finite state space X^N, where X_ℓ^N is the population state at the beginning of the ℓ-th epoch. Let $\{\xi_{k,i}^{N,x}, k \in [K], i \in [Nx_k]\}$ be a collection of i.i.d. random variables with

$$\mathbb{P}(\xi_{k,i}^{N,x} = z) = \begin{cases} \frac{1}{N} \cdot \rho_{kj}(F^N(x), x), & \text{if } z = \frac{1}{N}(e_j - e_k) \\ 0, & \text{o.w.} \end{cases}$$

where e_k is the natural coordinate vector with a one in the k-th entry, and $\mathbf{0}$ the zero vector of an appropriate size. $\xi_{k,i}^{N,x}$ then denotes the individual contribution to the change in the population state over an epoch by user i on channel k, when the current population state is x. In our setup, each epoch on a channel has duration of L time units. We next consider a conceptual time scaling by letting each time unit be $\frac{1}{LN}$ of a "primitive" time metric, so that an epoch has a duration of $\frac{1}{N}$ in this metric. The primitive time metric can be considered as relative to an outside observer of the system, that is, as the population increases, we shall observe the evolution with accelerated motion. It follows that the expected change per unit of primitive time metric of $\{X_\ell^N\}_{\ell=0}^\infty$ is given by

$$V^N(x) = N\mathbb{E}[X_{\ell+1}^N - X_\ell^N | X_\ell^N = x] = N\mathbb{E}\left[\sum_{k=1}^{K} \sum_{i=1}^{Nx_k} \xi_{k,i}^{N,x}\right]$$

$$= \sum_{k=1}^{K} \sum_{j=1}^{K} x_k \rho_{kj}(F^N(x), x)(e_j - e_k),$$

or for each entry,

$$V_k^N(x) = \sum_{j=1}^{K} x_j \rho_{jk}(F^N(x), x) - x_k.$$

Define the mean dynamics induced by the continuous population game, denoted by $V : \Delta_K \to \mathbb{R}^K$, as

$$V_k(x) = \sum_{j=1}^{K} x_j \rho_{jk}(F(x), x) - x_k.$$

V^N then converges to V uniformly given the uniform convergence of F^N to F and the uniform continuity of ρ. Consider the dynamic system

$$\dot{x} = V(x).$$

Let S_x be the set of all solutions with initial condition $x(0) = x$. Given a solution $\{x(t)\}_{t \geq 0}$, define the limit set of $\{x(t)\}$ by

$$\omega(\{x(t)\}) = \bigcap_{\tau \geq 0} \mathrm{cl}(\{x(t)\}_{t \geq \tau}),$$

where cl denotes the closure. Define the limit of set of a point x aby

$$\omega(x) = \bigcup_{\{x(t)\} \in S_x} \omega(\{x(t)\}),$$

and the Birkhoff center of the dynamic system is then defined by

$$BC(V) = \mathrm{cl}(\{x \in \Delta_K : x \in \omega(x)\}).$$

We next state a result, which is Theorem 3.5 of [68] in our context[3], and it shows

[3]the applicability of this theorem can be verified as in Example 4.1 and 4.2 in [68]. Also see

the Markov chain $\{X_\ell^N\}_{\ell=0}^\infty$ is statistically concentrated on the Birkhoff center of the deterministic dynamic system in the large population regime.

Theorem IV.2. *Let μ^N be an invariant measure of the Markov chain $\{X_\ell^N\}_{\ell=0}^\infty$, and let O be any open set containing $BC(V)$. Then, $\lim_{N\to\infty}\mu^N(O)=1$.*

This result provides a guideline on decentralized implementation of throughput optimal policies: a decentralized implementation would be realized if by appropriately choosing the learning algorithm, thus the system dynamics V, the Birkhoff center $BC(V)$ can be contained in the set of throughput optimal allocations. We shall show in the following that this is indeed feasible. $BC(V)$ can be a singleton given by the unique NE of the population game F, which interestingly is a throughput optimal allocation in the large population regime, as it is a solution to the following limit form of problem (4.1),

$$\min_{x\in\Delta_K}\max_{k\in[K]}\alpha_k e^{x_k}. \tag{4.2}$$

Definition IV.3. For a continuous population game F, x is a NE if $x_k>0$ implies $F_k(x)\geq F_j(x)$ for all $j\in[K]$.

Lemma IV.4. *The NE of the continuous population game F with $F_k(x)=-\alpha_k e^{x_k}$ is given by the solution to*

$$\min_{x\in\Delta_K}\sum_{k=1}^K\alpha_k e^{x_k} \tag{4.3}$$

Proof. This result is immediate by observing that F is a potential game with a concave potential function $f=-\sum_{k=1}^K\alpha_k e^{x_k}$, and hence its NE is unique and is given by the maximizer of the potential function. □

Lemma IV.5. *Let x^* be the optimal solution to problem* (4.3)*. Then it is also an optimal solution to problem* (4.2)*.*

Chapter 12 in [67] for a text on the same topic.

Proof. Without loss of generality, assume that $\alpha_1 \le \alpha_2 \le \cdots \le \alpha_K$. Let $\hat{x}^*$ be an optimal solution to (4.2). We consider two cases.

Case 1. If

$$\min_{x \in \Delta_{K-1}} \max_{k \in [K-1]} \alpha_k e^{x_k} \le \alpha_K,$$

then

$$\min_{x \in \Delta_K} \max_{k \in [K]} \alpha_k e^{x_k} \le \alpha_K.$$

Thus, $\hat{x}_K^* = 0$; otherwise, $\max_{k \in [K]} \alpha_k e^{\hat{x}_k^*} \ge \alpha_K e^{\hat{x}_K^*} > \alpha_K$. We then claim $x_K^* = 0$. Indeed, if $x_K^* > 0$, the KKT conditions for the convex program (4.3) implies $\alpha_k e^{x_k^*}$ is a constant for all k. Hence, $\sum_{k=1}^K \alpha_k e^{x_k^*} > K\alpha_K$. On the other hand, $\sum_{k=1}^K \alpha_k e^{\hat{x}_k^*} \le K\alpha_K$, which is a contradiction. Given $x_K^* = 0$, using again the KKT conditions, we have $\alpha_k e^{x_k^*} \le \alpha_K$, and thus

$$\max_{k \in [K]} \alpha_k e^{x_k^*} = \alpha_K = \max_{k \in [K]} \alpha_k e^{\hat{x}_k^*}.$$

Case 2. If

$$\min_{x \in \Delta_{K-1}} \max_{k \in [K-1]} \alpha_k e^{x_k} > \alpha_K,$$

it is not hard to see that $\hat{x}_k^* > 0$ and $\alpha_k e^{\hat{x}_k^*}$ is a constant for all k. On the other hand, the strong duality holds for (4.3), and $\hat{x}$ satisfies the KKT conditions. Hence, $\hat{x}^* = x^*$. □

Given the above results, the remaining question is on the choice of the learning algorithm or the revision protocol ρ, so as to satisfy our assumption on ρ and induce a dynamic system identified by V with $BC(V)$ being the unique NE.

Toward that end, we note there are a number of ways to choose ρ depending on the information available to each user on the population state after each epoch, and the same class of system dynamics can be induce by a family of learning algorithms. Also, typical classes of system dynamics also tend to share a number of similar features, thus greatly expanding the domain of applications. A detailed exhibition of this topic can be found in [67]. Below we show one example. Assume the users can observe the population state at the end of each epoch. Consider then the pairwise comparison rule, which is given by

$$\rho_{kj}(y,x) = \begin{cases} \frac{[y_k - y_j]_+}{R}, & \text{if } j \neq k \\ 1 - \sum_{\ell=1}^{K} \frac{[y_k - y_\ell]_+}{R}, & \text{if } j = k \end{cases}$$

where R is a constant such that $R \geq \sup_{y \in \mathcal{K}} \sum_{k \in [K]} y_k$. That is, the probability of switching to a different channel is proportional to the excess payoff that can be derived from switching. Since F is a potential game, following Theorem 7.1.2 in [67] we conclude that $BC(V)$ is given by the unique NE of F.

4.3 Concluding Remarks: Hope and Challenge

The preceding analysis provides us with a few interesting conjectures on what we should expect when we perform a similar analysis on DCF.

1. We have analyzed the decentralized implementation of the *static* throughput optimal policy in the simplified setup, whereby via individual learning and adjustment, the population state converges to the static load balance in the long run, which is also throughput optimal in one shot. The nature of this implementation as a multi-stage interaction among nodes in fact gives arise to dynamic channel switching policies, and in that context throughput optimality can and should be considered in the bigger space of dynamic policies. As measured by the individual/total throughput, dynamic load balancing outperforms its static counterpart; this is easily seen from the convexity of the individual/total throughput in the population state. One would then expect that a dynamic throughput optimal policy should be such that it achieves load balance on av-

erage while avoiding completely idle channels.

2. Decentralized implementation of throughput optimal polices was shown to be feasible in this simple model. This implementation features the collision rate or in general the congestion level as an individual measure of utility from using a channel, which seems both intuitive and promising in terms of engineering practice. Indeed, local congestion level on a channel usually can be easily estimated by users residing in that channel. However, a key component of this type of implementation is the learning algorithm or revision protocol (including the pairwise comparison rule shown as an example), which typically requires certain global information[4] such as the population size and the population state, to infer the global congestion level or other payoff related quantities. Such global information is typically incomplete or entirely unavailable in real time to individual nodes without centralized knowledge sharing or excessive overhead due to massive message exchange. On the other hand, it is possible that load balance may be achieved using symmetric randomization among nodes without such global information. This motivates a decentralized implementation in an even simpler form that we shall discuss in the next Chapter.

[4]The family of imitative protocol [67] can solely rely on local information, but they suffer from the problem of extinction and the induced dynamics would have a much bigger Birkhoff center than the singleton NE. See Chapters 4, 5 and 7 in [67] for a detailed discussion.

CHAPTER V

Throughput Optimal Channel Switching in Random Access - Part II: IEEE 802.11 WLANs

In this chapter, we shall build a complete model to analyze the DCF protocol incorporating the heterogeneity in traffic loads among users. Using the Aloha based simplification in the last chapter, we observed the throughput optimal channel allocation policy exhibits the rule-of-thumb principle of load balancing and that decentralized implementation may be feasible using the local congestion level as an indicator of payoff of using a channel. These lessons will turn out to be pivotal in interpreting key results we shall derive for DCF, which inherits the essence from the slotted Aloha. A summary of notation used throughout this chapter can be found in Appendix D.

5.1 802.11 DCF Backoff Mechanism

Compared to the slotted Aloha, instead of using coin toss to determine whether a transmission should be attempted, the 802.11 DCF utilizes a process called backoff to further reduce potential collision. The underlying system is still slotted, however with heterogeneous slot lengths, which will be discussed at length in Section 5.2. Also, all transmission attempts are aligned with slot edges, as we have seen in the model of slotted Aloha. The channel is modeled as a collision channel with feedback, and all nodes can observe the state of the channel that can either be busy or idle. The medium is idle if no node is transmitting and busy otherwise. Each idle slot is of a fixed length, say $20\mu s$ in the 802.11b specification, and the length of a busy slot depends on the packet length and whether it is a successful transmission or a collision. To initiate a transmission, a node initializes a counter, called backoff counter, with the value randomly generated from 0 to $W_0 - 1$. The backoff counter decrements once during each idle slot, and this process freezes if the channel is busy and resumes when it is idle again. When the counter reaches zero, the node starts transmission, and a collision occurs if there are two or more simultaneous attempts. Retransmissions are then performed with the same backoff process but a larger window, say W_1, to set the backoff counter, hoping to reduce the likelihood of a second collision. Particularly, $W_j = 2^{\min(j,m)} W_0$ where m is the maximum number of backoff stages, and this scheme is commonly known as the *binary exponential backoff*. After M unsuccessful retransmissions where $M \geq m$, the packet is discarded. Here we introduced an even simplified version of 802.11 DCF in the sense that all packets are assumed to experience the backoff process and the so-called post-backoff feature is omitted due to negligible impact on performance measures.

5.2 Problem Formulation

Consider a multiple access system using the IEEE 802.11 DCF (see Section 4.1 for a brief account on DCF. There are N users, indexed by the set $[N] = \{1, 2, \ldots, N\}$, each with an infinite buffer, one transceiver (i.e., a single wireless interface) and uses the same parameterization. We assume the channel is ideal and there is no MAC-level packet discard, i.e., there is no retransmission limit of a packet after collision. Throughout the analysis we also adopt a few other simplifying assumptions to make the problem tractable; these will be stated in the context to which they apply. We later show that these simplifications do not impact the accuracy of the model under normal operating parameter values.

The key to our method is to model the queue at each node with a service process defined by 802.11 DCF as a *slotted mean field Markov chain* [69]. A "mean-field" model approximates the effect of all the other players on any given one by a single averaged effect, when studying a large number of interacting players, and each individual has only a marginal impact on the overall population. Note that this idea is consistent with the continuous population game in the slotted Aloha simplification that we have seen in the last chapter. We first define the notion of slot.

Definition V.1. A *virtual backoff timer* of the system (or of a virtual node) is a universal timer for all nodes in the system: it counts down indefinitely, alternating between the count-down mode (when nodes in the system are counting down) and the freezing mode (when some node in the system is transmitting). A *slot* is the time period between two successive decrements of the virtual backoff timer.

Remark V.2. The above definition provides a universal slot time for all nodes in the system, and we shall assume that the real backoff timer at each node is synchronized to this virtual timer on slot boundaries. The motivation behind such a construction originates from the principal difficulty in modeling a non-saturated system: the service process at each node runs in embedded time in terms of a slot, which is in general a random variable, whereas the packet arrival process is more naturally described in real time [69]. This difficulty does not exist in saturated analysis, see e.g., [45], where arrival processes do not play a role.

We further introduce three key assumptions in our model, followed by a discussion on their implications and limitations.

Assumption V.3.

(A1) *The MAC layer arrival process at node i is Poisson with rate λ_i bits per second.*

(A2) *(i) The service time of a packet, i.e. the time from the initial backoff to* *successful transmission, is exponential with service rate μ_i at node i, and independent of all arrival processes.*

(ii) Given the vector of arrival rates $(\lambda_1, \lambda_2, \ldots, \lambda_N)$ at all nodes, the vector of service rates $(\mu_1, \mu_2, \ldots, \mu_N)$ takes values from a set *given by a* correspondence *mapping from the arrival rate vector, i.e., there exists a correspondence $\boldsymbol{\mu}$ such that $(\lambda_1, \lambda_2, \ldots, \lambda_N) \overset{\boldsymbol{\mu}}{\mapsto} \{(\mu_1, \mu_2, \ldots, \mu_N)\}$. Each service rate vector will represent a state of the system given the same incoming traffic load.*

(A3) *Let $S(t)$ be the counting process of the number of slots accumulated up to time t and let $Q_i(t)$ be the number of packets in the MAC queue of node i at time t.*[1] *$S(t)$ is assumed to be independent of $Q_i(t)$ and renewal.*

The above simplifying assumptions are not entirely realistic. Typically, due to congestion control by upper-layer protocols, e.g., TCP, the arrival process to the MAC layer is neither Poisson nor independent of the service process. However, as our objective is to explore the inherent properties of 802.11 DCF, the independence assumption is adopted to decouple the MAC layer from upper layers, while the Poisson and exponential assumptions are adopted to avoid technicalities that can obscure the main insight. Note that under the mean field methodology, each node is analyzed in isolation from the activities of all other nodes which are collectively regarded as an aggregate stationary process. Within such a framework the packet service time is taken to be stationary (see e.g., Bianchi's well-known mean field Markovian model of the service process [45]).

With **A1** and **A2**, each $Q_i(t)$ is then a well-defined $M/M/1$ queue for any given pair of $(\lambda_1, \lambda_2, \ldots, \lambda_N)$ and $(\mu_1, \mu_2, \ldots, \mu_N)$, and each queue is stable if and only if $Q_i(t)$ is positive recurrent. Equivalently we may consider the utilization factor ρ_i

[1] As to our notation convention in this chapter, we shall use letters in boldface to denote vector-valued quantities explicitly, and we place time indices as arguments for random processes.

at node i, given by $\rho_i = \min\{\frac{\lambda_i}{\mu_i}, 1\}$: the queue is stable if and only if $\rho_i < 1$. Let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N)$, and we then formally define the stability region of system as follows.

Definition V.4. The system is *stable* given $\boldsymbol{\lambda}$ if all user queues are stable for all $(\mu_1, \mu_2, \ldots, \mu_N) \in \boldsymbol{\mu}(\boldsymbol{\lambda})$ determined by the DCF scheme. The *stability region* Λ is the set of all $\boldsymbol{\lambda} \in \mathbb{R}_+^N$ such that the system is stable.

If $Q_i(t)$ is positive recurrent, then it is ergodic and we have $\lim_{t\to\infty} P(Q_i(t) > 0) = 1 - \pi_i(0) = \rho_i$, where $\{\pi_i(u)\}_{u=0}^{\infty}$ is the stationary distribution of $Q_i(t)$. If $Q_i(t)$ is transient or null recurrent, in which case $\rho_i = 1$, we have $\lim_{t\to\infty} P(Q_i(t) = 0) = 0 = 1 - \rho_i$. Therefore, ρ_i is asymptotically given by $\lim_{t\to\infty} P(Q_i(t) > 0)$ in all cases in our model.

For technical reasons we shall also consider the *embedded* queueing process $\hat{Q}_i(n)$, $n = 1, 2, \cdots$, defined by $\hat{Q}_i(n) := Q_i(T_n)$, where T_n is the time of the n-th slot boundary. $\hat{Q}_i(n)$ is thus a discrete-time process constructed by observing $Q_i(t)$ at slot boundaries. For an arbitrary process $S(t)$, $\hat{Q}_i(n)$ is not necessarily Markovian. However, given assumption **A3**, durations between slot boundaries are i.i.d., constituting sampling periods that are independent of $Q_i(t)$. Hence $\hat{Q}_i(t)$ is a discrete-time Markov chain under our assumption. It's worth noting that **A3** does not exactly hold in reality because the slot length is a function of a node's activity, and thus the state of its queue, even with the mean field simplification of other nodes' behavior (this is more precisely shown in the appendix). However, this dependence weakens when the number of nodes or the backoff window size is sufficiently large. We empirically show that this assumption does not impact the accuracy of prediction even with a small node population and backoff window size.

Let $\hat{\rho}_i$ denote the utilization factor under the discrete-time system $\hat{Q}_i(n)$. In general $\hat{\rho}_i \neq \rho_i$. Indeed we show in Appendix C that $\hat{\rho}_i \leq \rho_i$ where equality holds if and only if $\rho_i = 1$ or $\rho_i = 0$, i.e., node i is either saturated or idle. Similar to ρ_i, $\hat{\rho}_i$ is asymptotically given by $\lim_{n\to\infty} P(\hat{Q}_i(n) > 0)$.

We shall adopt Bianchi's decoupling approximation [45] as another key assumption, stated as follows. Define $C_i(j) := 1$ if the j-th attempt by node i results in a collision, and $C_i(j) := 0$ if it results in a success.

Assumption V.5.

(A4) *[Bianchi's Decoupling Approximation] For each node $i \in [N]$, the collision sequence $\{C_i(j)\}$ is i.i.d. with $P(C_i(j) = 1) = p_i$ for some constant p_i.*

In reality successive attempts by the same node may occur if it repeatedly selects timer value 0 while other nodes' timers remain frozen. In such cases the above assumption ceases to hold. This phenomenon can be prominent when the window size is small, and has been taken into account in some recent work [70]. In this study we shall ignore the possibility of successive attempts for simplicity of presentation and adopt **A4**. (A more precise model is possible by imposing independence not on all attempts but only the first attempt in each such sequence.) This is reasonable when the initial window size is sufficiently large. Our empirical results are fairly close between with and without consideration of successive attempts for large backoff windows. For small backoff windows, the discrepancy between the two will be illustrated in the numerical results.

Define respectively N_i^s and N_i^{tx} as the numbers of slots and transmission attempts that node i takes in serving one packet. $\overline{W}_i := \frac{\mathbb{E}N_i^s}{\mathbb{E}N_i^{tx}}$ is referred to as the average size of backoff window of node i.

Using Bianchi's approximation, we have

$$\begin{aligned}\mathbb{E}N_i^s &= \sum_{k=0}^{\infty}\sum_{j=0}^{k}\frac{2^{\min\{j,m\}}W+1}{2}(p_i)^k(1-p_i)\\ &= \sum_{j=0}^{\infty}\frac{2^{\min\{j,m\}}W+1}{2}\left(\sum_{k=j}^{\infty}(p_i)^k(1-p_i)\right)\end{aligned}$$

$$= \sum_{j=0}^{\infty} \frac{2^{\min\{j,m\}}W + 1}{2} (p_i)^j$$

where W is the size of the initial backoff window and m is the value of the maximum backoff stage. Also note $\mathbb{E}N_i^{tx} = \frac{1}{1-p_i}$. Therefore, $\overline{W}_i$ is given by

$$\overline{W}_i = \frac{1}{2}\left[W\left((1-p_i)\sum_{j=0}^{m-1}(2p_i)^j + (2p_i)^m\right) + 1\right].$$

We next derive a relationship between the transmission attempt probability and $\hat{\rho}_i$. Let $\tau_i(n)$ be the probability that node i initiates a transmission attempt in the n-th slot.

Lemma V.6. *$\tau_i := \lim_{n\to\infty} \tau_i(n)$ exists and is given by $\tau_i = \hat{\rho}_i/\overline{W}_i$.*

Proof. Denote by $Tx(n)$ the event that node i initiates an attempt in the n-th slot. Then

$$\tau_i(n) = P(Tx(n)|\hat{Q}_i(n) > 0) \cdot P(\hat{Q}_i(n) > 0) +$$
$$+ P(Tx(n)|\hat{Q}_i(n) = 0) \cdot P(\hat{Q}_i(n) = 0).$$

Consider now the sequence of slots in which node i has a packet in service. Given the decoupling among nodes, the occurrences of slots in which node i starts the service for a packet thus form renewal events. Regarding each transmission attempt as one unit of reward and using the renewal reward theory, we then obtain

$$\lim_{n\to\infty} P(Tx(n)|\hat{Q}_i(n) > 0) = \frac{\mathbb{E}N_i^{tx}}{\mathbb{E}N_i^s} = \frac{1}{\overline{W}_i}.$$

Since $\lim_{n\to\infty} P(\hat{Q}_i(n) > 0) = \hat{\rho}_i$, and $P(Tx(n)|\hat{Q}_i(n) = 0) = 0$, the result follows. ☐

To put the above result in context, one easily verifies that in the extreme case where all nodes are saturated and identical, we have $\hat{\rho}_i = \rho_i = \rho = 1$ and $p_i = p$ for all i. Consequently,

$$\begin{aligned}\tau_i = \tau &= \frac{2}{W\left((1-p)\sum_{j=0}^{m-1}(2p)^j + (2p)^m\right) + 1} \\ &= \frac{2(1-2p)}{(1-2p)(W+1) + pW(1-(2p)^m)},\end{aligned}$$

which is exactly the same as obtained in [45] equation (7).

5.3 Single Channel Stability Region

5.3.1 The Stability Region Equation Σ

Our first main result is the following theorem on the quantitative description of Λ. Let $\mathbb{E}S_{i,Q,\overline{Tx}}$ denote the conditional average length of a slot given that the queue at node i is non-empty but i does not transmit in this slot. T_s and T_c denote the time duration of a successful transmission and a collision, respectively.

Theorem V.7. $\boldsymbol{\lambda} \in \Lambda$ *if and only if for any solution* $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_N) \in [0,1]^N$ *to the following system of equations* Σ *given* $\boldsymbol{\lambda}$*,*

$$\Sigma : \begin{cases} \tau_i = \dfrac{\hat{\rho}_i}{\overline{W}_i}, \forall i & \text{(a)} \\ p_i = 1 - \prod_{j \neq i}(1-\tau_j), \forall i & \text{(b)} \\ \rho_i = \min\left\{\dfrac{\lambda_i}{P}\left(\dfrac{\overline{W}_i - 1}{1-p_i}\mathbb{E}S_{i,Q,\overline{Tx}} + T_c\dfrac{p_i}{1-p_i} + T_s\right), 1\right\}, \forall i & \text{(c)} \end{cases}$$

where P *is the packet payload size, we have* $0 \leq \rho_i < 1$ *for all* $i \in [N]$*. Let*

$$\mathrm{C} : \begin{cases} 0 \leq \tau_i \leq 1, \forall i & \text{(i)} \\ 0 \leq \rho_i < 1, \forall i & \text{(ii)} \end{cases}$$

We shall denote this constrained system of equations by $(\Sigma, \mathrm{C}, \boldsymbol{\lambda})$*.*

Proof. Σ(a) is the result of Lemma V.6, and Σ(b) is an immediate consequence of the definition of p_i. Let the average service time at node i be $\overline{X}_i$ seconds per bit; the average service time per packet is thus $P\overline{X}_i$. Define $\overline{Y}_i(j)$ as

$$\overline{Y}_i(j) = T_c + \left(\frac{2^{\min\{j,m\}}W+1}{2} - 1\right)\mathbb{E}S_{i,Q,\overline{Tx}}.$$

Physically, $\overline{Y}_i(j)$ is the average time between the beginning of the j-th transmission attempt, which results in a collision, and the beginning of the $(j+1)$-th attempt, given that node i encounters at least j collisions before completing the service of some packet. Since the collision sequence is geometric, we have

$$\begin{aligned}
P\overline{X}_i &= \sum_{k=0}^{\infty}\left[\left(\frac{W+1}{2}-1\right)\mathbb{E}S_{i,Q,\overline{Tx}} + \sum_{j=1}^{k}\overline{Y}_i(j) + T_s\right] \times (p_i)^k(1-p_i) \\
&= \sum_{j=1}^{\infty}\sum_{k=j}^{\infty}\overline{Y}_i(j)(p_i)^k(1-p_i) + \left(\frac{W+1}{2}-1\right) \times \mathbb{E}S_{i,Q,\overline{Tx}} + T_s \\
&= \sum_{j=1}^{\infty}(p_i)^j\overline{Y}_i(j) + \left(\frac{W+1}{2}-1\right)\mathbb{E}S_{i,Q,\overline{Tx}} + T_s.
\end{aligned}$$

Therefore,

$$\begin{aligned}
P\overline{X}_i &= \sum_{j=1}^{\infty}\left[(p_i)^j\left(T_c + \left(\frac{2^{\min\{j,m\}}W+1}{2}-1\right) \times \mathbb{E}S_{i,Q,\overline{Tx}}\right)\right] + \\
&\quad + \left(\frac{W+1}{2}-1\right)\mathbb{E}S_{i,Q,\overline{Tx}} + T_s \\
&= \sum_{j=0}^{\infty}\left[\frac{2^{\min\{j,m\}}W-1}{2}(p_i)^j\right]\mathbb{E}S_{i,Q,\overline{Tx}} + T_c\sum_{j=1}^{\infty}(p_i)^j + T_s \\
&= \frac{\overline{W}_i-1}{1-p_i}\mathbb{E}S_{i,Q,\overline{Tx}} + T_c\frac{p_i}{1-p_i} + T_s.
\end{aligned}$$

Note that $\tau_i < 1$ for all i, and we have $p_i < 1$ for all i as a result. In addition, $\mathbb{E}S_{i,Q,\overline{Tx}}$ is finite (computed in the appendix). Hence we conclude that the packet service time is finite. Thus, the utilization factor of node i is given by $\rho_i = \min\{\lambda_i\overline{X}_i, 1\}$ and

Σ(c) follows, and C(i) is for the validity of τ as a probability measure. (Σ, C(i), $\boldsymbol{\lambda}$) then constitutes a full description on the system utilization. C(ii) is the necessary and sufficient condition for stability as commented in the previous section. $\square$

For a given set of system parameter values, two sets of quantities are needed to compute Σ: $\mathbb{E}S_{i,Q,\overline{Tx}}$ and $\hat{\rho}_i$, $\forall i \in \mathcal{N}$. These are computed in Appendix C, respectively. In particular, in Appendix C we show that though it is analytically intractable, $\hat{\rho}_i$ is well approximated by

$$\hat{\rho}_i \approx \frac{\rho_i \mathbb{E}S_{i,\overline{Q}}}{\rho_i \mathbb{E}S_{i,\overline{Q}} + (1-\rho_i)\mathbb{E}[S_{i,Q}]},$$

where $\mathbb{E}S_{i,Q}$ (resp. $\mathbb{E}S_{i,\overline{Q}}$) is the conditional average length of a slot given that the queue at node i is non-empty (resp. empty) at the beginning of this slot.

5.3.2 Characterizing the Solutions to Σ

Without the stability constraint C(ii), (Σ, C(i), $\boldsymbol{\lambda}$) can be rewritten as a vector-valued fixed point equation $\boldsymbol{\tau} = \boldsymbol{\Gamma}(\boldsymbol{\tau})$ over $[0,1]^N$, where $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_N)$, and the existence of solutions can be shown by Brouwer's fixed point theorem. Hence, the condition in Theore V.7 is not vacuous. However, the uniqueness of its solution is in general difficult to prove; nevertheless, under the condition of a sufficiently large initial backoff window W, we have the following result on the uniqueness of its solution.

With a large initial backoff window W, the probability of collision is small, so we have $\overline{W}_i \approx \frac{W+1}{2}$. We also observe that $\mathbb{E}S_{i,Q} \approx \mathbb{E}S_{i,\overline{Q}}$ when W is large (cf. Appendix C). Consequently, we can approximate $\hat{\rho}_i$ by ρ_i. Also, using the first-order Taylor approximation, we have $\prod_{j\neq i} \frac{1}{1-\tau_j} \approx 1 + \sum_{j\neq i} \tau_j$ for small τ. Note that the minimization operator in Σ is redundant when combined with C(ii). Hence, let $T_s = T_c = T$ for simplicity of presentation, and $(\Sigma, \mathrm{C}, \boldsymbol{\lambda})$ can be then approximated by the following constrained system of equations,

$$\widetilde{\Sigma}: \begin{cases} \tau_i = \frac{\rho_i}{\frac{W+1}{2}}, \forall i & \text{(a)} \\ \rho_i = \frac{\lambda_i}{P}\left[\frac{W-1}{2}\left(\sigma + T\sum_{j\neq i}\tau_i\right) + T\left(1+\sum_{j\neq i}\tau_i\right)\right], \forall i & \text{(b)} \end{cases}$$

subject to the same set of constraints.

Proposition V.8. $(\widetilde{\Sigma}, \boldsymbol{\lambda})$ *admits a unique solution.*

Proof. See Appendix C. □

Remark V.9. 1) The above result suggests that Σ has a unique solution when W, the initial window size, is sufficient large. As an approximation we shall take this condition to be equivalent to a large average backoff window. This is because the probability of a (first-attempt) collision decays inverse-linearly in W, and thus $\overline{W}_i$ is dominated by W when W is sufficiently large. In this case, the correspondence (μ), which is implicitly given by Σ, reduces to a vector-valued function.

2) As we shall see numerically in the next section, multiple fixed point solutions may arise when W is small; this will be referred to as multi-equilibrium (as opposed to "multistable" or "metastable" [69] to avoid confusion). As to the nature of multiplicity of solutions, consider the two-user scenario with symmetric load and attempt rate. The number of packets that can delivered in a slot is then given by $2\tau(1-\tau) = 2\frac{\hat{\rho}}{W}(1-\frac{\hat{\rho}}{W})$, which can be mapped to the deliverable amount of data per unit of time given the average scaling of a slot. Since $\hat{\rho} \in [0,1]$, we have $\tau \in [0, \frac{1}{W}]$, i.e., loosely speaking, the size of backoff window controls the effective range of τ. When W is large, the function $2\tau(1-\tau)$ exhibits monotonicity, while for small values of W, for given $2\tau(1-\tau)$ measured by an outside observer, there can exist two different states of the system utilization, thus the multi-equilibrium phenomenon. As we shall see, this upper-bounding effect of backoff window size on the attempt rate will continue to play a key role in interpreting other results.

In the proof of Proposition V.8, we in fact obtained the approximated unique solution to $(\Sigma, \boldsymbol{\lambda})$. Therefore, by imposing feasibility constraints C, we can induce a simplified version of $(\Sigma, \mathrm{C}, \boldsymbol{\lambda})$, which is an approximation to Λ and is easier to compute.

Corollary V.10. *When W is sufficiently large, Λ is approximated by*

$$\tilde{\Lambda} = \left\{ \boldsymbol{\lambda} \in \mathbb{R}_+^N \;\middle|\; 0 < \frac{\gamma_i^1(\lambda_i) \sum_j \gamma_j^2(\lambda_i)}{1 - \sum_i \gamma_j^1(\lambda_i)} + \gamma_i^2(\lambda_i) < \frac{2}{W+1}, \forall i \right\}$$

where $\gamma_i^1(\lambda_i) = \frac{\lambda_i T}{P} \Big/ \left(1 + \frac{\lambda_i T}{P}\right)$, *and* $\gamma_i^2(\lambda_i) = \frac{\lambda_i((W-1)\sigma + 2T)}{P(W+1)} \Big/ \left(1 + \frac{\lambda_i T}{P}\right)$.

5.4 Numerical Results: Single Channel

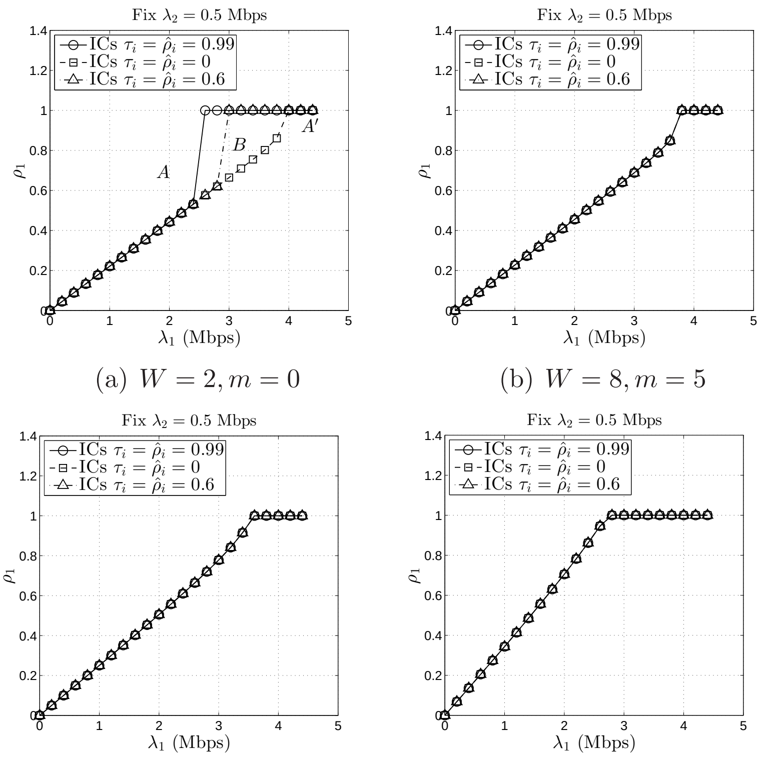
Using $(\Sigma, \mathrm{C}, \boldsymbol{\lambda})$, we can quantitatively describe the stability region of a single channel system, and some numerical results for the two-user case are illustrated in this section. The parameters used in both the numerical computation and the simulation are reported in Table C.1 in Appendix C. Under the basic access mechanism of DCF we have

$$\begin{cases} T_s = \frac{P}{\text{Tx. Rate}} + \text{Header} + \text{ACK} + \text{DIFS} + \text{SIFS} + 2\delta \\ T_c = \frac{P}{\text{Tx. Rate}} + \text{Header} + \text{DIFS} + \delta \end{cases}$$

where δ is the propagation delay.

5.4.1 Multi-equilibrium and Discontinuity in ρ

We first illustrate the existence of multi-equilibrium solutions and discontinuity of $\rho_i(\boldsymbol{\lambda})$ in $\boldsymbol{\lambda}$; this is shown in Figure 5.1. We fix the value of λ_2 and increase λ_1 from 0 to 4.5 Mbps. For each pair $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$, we solve for the fixed point(s) of $\boldsymbol{\Sigma}$ with the same set of initial values of τ_i and $\hat{\rho}_i$ for $i \in [N]$ to which we refer as a set of initial conditions (ICs). We then convert the results to $\boldsymbol{\rho} = (\rho_i, i \in [N])$ using Eqn. Σ(c).



(a) $W = 2, m = 0$

(b) $W = 8, m = 5$

(c) $W = 32, m = 5$

(d) $W = 128, m = 5$

Figure 5.1: Solution components for various scenarios: an illustration.

The collection of the pairs $(\boldsymbol{\lambda}, \boldsymbol{\rho}(\boldsymbol{\lambda}))$ then constitutes a *solution component* for this set of ICs. Note that this is obtained by solving (Σ, C(i), $\boldsymbol{\lambda}$) without considering the stability constraint C(ii). We repeat the above computation for different sets of ICs under the same system parameters including W and m. The entire process is then repeated for different pairs (W, m). For each pair (W, m), the resulting solution components constitute an overall correspondence between the vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\rho}(\boldsymbol{\lambda})$, and this is plotted for ρ_1 vs. λ_1 in Figure 5.1.

In the first scenario as shown in Figure 5.1(a), where the initial window is of the smallest possible size for two users and window expansion is disallowed ($m = 0$), three different zones of the correspondence $\rho_1(\lambda_1)$ are present, labeled as A, A' and B in the figure. In zones A and A', a single fixed point is admitted and $\rho_1(\lambda_1)$ reduces to a function, while in zone B we see two solutions. Along each solution component, there is a jump in ρ_1 in zone B as λ_1 increases; this is essentially a phase transition from stable to unstable regions. What this result illustrates is that depending on the initial condition, certain input rates may or may not lead to a feasible solution (a point that corresponds to stable queues). Thus when such multi-equilibrium exists, we may have a collection of "stability regions" given different initial conditions, and this phenomenon is illustrated in Figure 5.3 and discussed in the next subsection in detail. Recall that under our definition of stability region and Theorem V.7, an arrival rate vector is considered within the stability region if and only if any initial condition induces so; the stability region thus defined is therefore the infimum of this collection when multiple equilibria exist.

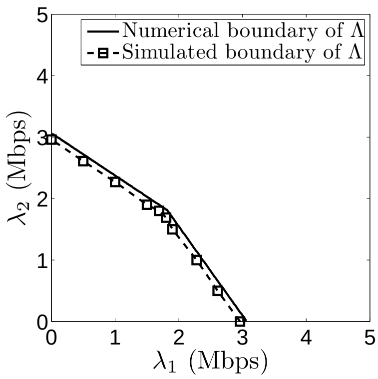
Intuitively, initial conditions with large values suggest a pessimistic prediction on the system stability under $\boldsymbol{\lambda}$, and it may thus result in a small Λ; by contrast, ICs with small values render an optimistic one and a larger Λ. Empirically, we find that the set of ICs with $\tau_i = \rho_i \approx 1$ for $i \in [N]$ results in the earliest jump in ρ_1 and the one with $\tau_i = \rho_i = 0$ for $i \in [N]$ gives the latest. Consequently, solution components resulting from these two sets of ICs define the boundary of zone B and the corresponding "stability regions", forming the empirical supremum and infimum of the collection of "stability regions".

Inspecting the set of figures Fig. 5.1(a)-5.1(d), we see that as the initial window increases, the multi-equilibrium gradually vanishes and the gap in ρ_1 caused by the jump discontinuity closes.

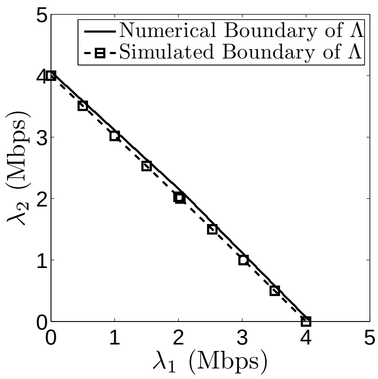
5.4.2 Numerical and Empirical Stability Regions

We numerically solve $(\Sigma, \mathrm{C}, \boldsymbol{\lambda})$ with two nodes to obtain the corresponding Λ, and then compare it with the simulated boundary. In simulation, for each fixed λ_2, we increase λ_1 with a step size $\Delta\lambda$, and compute the empirical throughput of node i obtained under $\boldsymbol{\lambda}$, denoted as $S_i^{\boldsymbol{\lambda}}$, and the number of backlogged packets at node i by the end of simulation, denoted as $B_i^{\boldsymbol{\lambda}}$. The simulator declares a point $\boldsymbol{\lambda}$ unstable if there exists at least one i such that $S_i^{\boldsymbol{\lambda}} < \lambda_i$ and $B_i^{\boldsymbol{\lambda}} P/(\lambda_i T_f) > \beta_i$, by the simulation time T_f, where β_i is an instability threshold, $0 < \beta_i < 1$. In the experiment we set $\Delta\lambda = 0.1$ Mbps (100 Kbps), $T_f = 10$ sec and $\beta_i = \beta = 1\%$. The stable point (λ_1, λ_2) such that $(\lambda_1 + \Delta\lambda, \lambda_2)$ is unstable is declared a point on the simulated boundary; the experiment is repeated for each λ_2 and the empirical mean value of λ_1 is recorded. Due to symmetry, only half of the boundary points are evaluated. The results are shown in Figure 5.2.

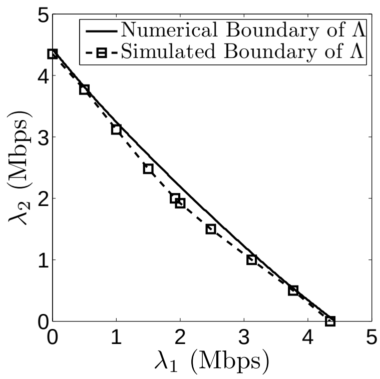
Our main observation is that when the initial (or average) backoff window is large, the stability region is convex (Figure 5.2(a)). The convexity gradually disappears as the window size decreases and the region is given by a near-linear boundary in Figure 5.2(b). It becomes clearly concave when the window size is small (Figure 5.2(c)). Interestingly, the case of $W = 32$ is the most frequently studied in the literature, and a linear boundary of the capacity region has been observed in [49]. As shown here, this linear boundary is only a special case in a spectrum of convex-concave boundaries.



(a) $W = 128, m = 5$



(b) $W = 32, m = 5$



(c) $W = 8, m = 5$

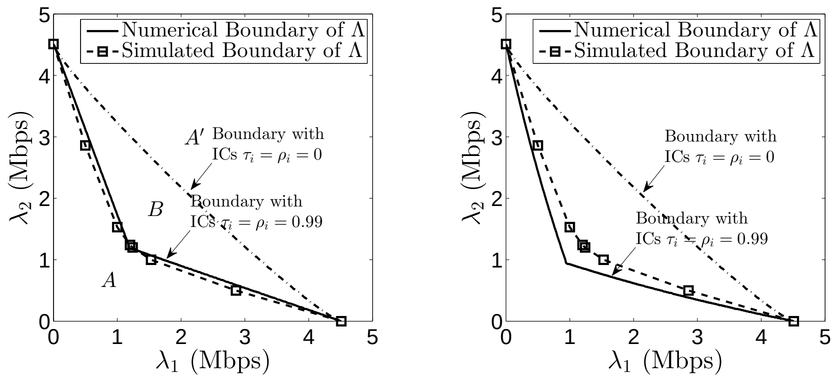
Figure 5.2: The stability regions in various scenarios - part I.

It is worth noting that in [50], Leith *et al.* established the general log-convexity of the rate region of 802.11 WLANs. This implies that the rate region could be either convex or concave, though [50] did not associate this with the window size as we have explicitly done here. It also suggests that the rate region and the stability region may be quite similar in nature; this however is not a formally proven statement, nor are we aware of such in the case of 802.11.

The change in the shape of the stability region as W changes may be explained as follows. Small W represents a highly aggressive configuration. This is much more beneficial when there is a high degree of asymmetry between the users' arrival rates. This is reflected in the concave shape of the region. When W is large, users are non-aggressive, which is more beneficial when arrival rates are similar, resulting in the convex shape. Numerically, the $W = 8$ case gives the largest stability region. This seems to suggest that the largest stability region is given by the smallest choice of W such that a unique feasible solution to $(\Sigma, \mathrm{C}, \boldsymbol{\lambda})$ exists. It would be very interesting to see if this could be established rigorously.

In Figure 5.3, we compute the stability regions of the case where $W = 2$ and $m = 0$ for two different sets of ICs. As discussed earlier, when multi-equilibrium exists we may have a collection of "stability regions". This is clearly seen in Figure 5.3: three different zones A, A' and B in the correspondence $\rho_1(\lambda_1)$ are mapped accordingly onto Λ. The boundary of stability region in our definition corresponds to the innermost one. As noted in [69], the simulated boundary reflects time-averages of multiple equilibria. Hence, our notion of stability region provides an estimation of the inner bound of the empirical stability region in this case.

As mentioned earlier, for small backoff windows the occurrence of successive attempts is non-trivial, which our model has ignored. The first-attempt decoupling approximation mentioned after **A4** captures the nodal behavior more accurately, and the adaptation of Σ using this alternative assumption is detailed in [71]. In Fig-



(a) Using Bianchi's decoupling approximation, $W = 2$ and $m = 0$

(b) Using the first-attempt decoupling approximation, $W = 2$ and $m = 0$

Figure 5.3: The stability regions in various scenarios - part II.

ure 5.3(b), we plot the counterpart of Figure 5.3(a) using the first-attempt decoupling approximation, and the discrepancy between results obtained using these two assumptions does exists. This is most notably shown in the numerical boundary A. The fact that the simulated boundary is now in between the two numerical boundaries verifies that this alternative assumption is more accurate. We do note however that for large windows this gap diminishes judging from numerical observation, which is to be expected.

5.4.3 Discussion: From 802.11 DCF Back to Aloha

We next recall results on the stability region of slotted Aloha, the natural prototype of modern 802.11 DCF, and provide an intuitive argument on why the qualitative properties of the stability region shown in the previous section are to be expected.

In [72], Massey and Mathys studied an information theoretical model of multi-access channel which shares several fundamental features with slotted Aloha. They investigated the Shannon capacity region of this channel with n users, which is shown to be the following subset of $\mathbb{R}^n_+$:

$$C = \left\{ \text{vect}\left(\tau_i \prod_{j \neq i} (1 - \tau_j) \right) \;\middle|\; 0 \leq \tau_i \leq 1, 1 \leq i \leq N \right\},$$

where $\text{vect}(v_i) = (v_1, v_2, \ldots, v_N)$, and τ_i is the transmission attempt rate of user i. In [53], Anantharam showed that the closure of the stability region of slotted Aloha is also given by C, under a geometrically distributed aggregate arrival process with parameter $1/(\sum_i \lambda_i)$ and probability $\lambda_i / \sum_j \lambda_j$ that such an arrival is at node i.

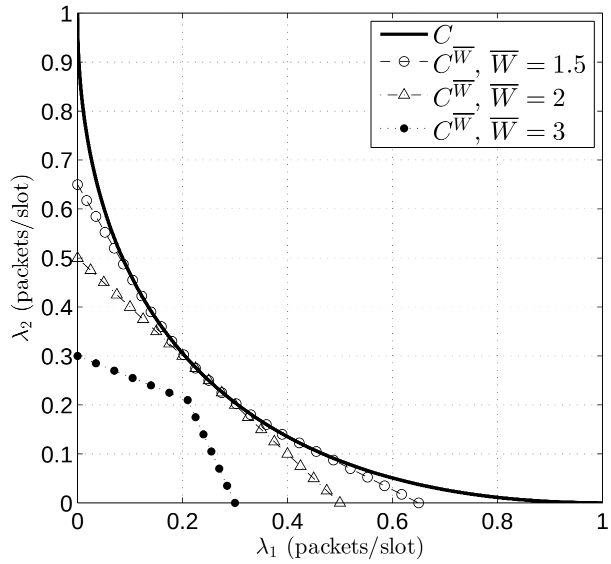


Figure 5.4: The stability region of slotted ALOHA and induced subsets.

The above result on slotted Aloha can be used to explain the stability region of 802.11 DCF. Note that the main difference between the two lies in the collision avoidance mechanism. Instead of attempting transmission with probability $0 \leq \tau \leq 1$ in a slot under slotted Aloha, under DCF each user randomly chooses a backoff timer value within a window. The effect the average backoff length $\overline{W}$ has on transmission under DCF is akin to that of restricting the attempt rate τ within an upper bound $\frac{1}{\overline{W}}$ under slotted Aloha. Hence, the stability region of 802.11 DCF may be viewed as a subset of C provided that we properly scale a slot to real time.

To verify this intuition, let $C^{\overline{W}}$ be the subset of C when $0 \leq \tau_i \leq \frac{1}{\overline{W}}$ for all i. In Figure 5.4, we plot C and $C^{\overline{W}}$ with different values of $\overline{W}$. We see that as $\overline{W}$ grows, $C^{\overline{W}}$ evolves from a concave set to a convex set, consistent with what we observed of 802.11 DCF in the previous subsection. It must be pointed out that this connection, while intuitive, is not a precise one technically. For instance, this connection might suggest that the stability region of 802.11 DCF will reduce to C when the average backoff length is 1. This is however not true. In this trivial case, the stability region of 802.11 DCF is reduced to one dimension, i.e., the system is unstable for $N \geq 2$. This is because the retransmission probability of DCF is also lower bounded by the reciprocal of the window size at its backoff stage, and in the case when the backoff length is one another collision occurs with certainty.

5.5 Multi-channel Analysis

Using a similar, mean-field Markovian model as we did in the single channel case, we can show that the stability region of a multi-channel system under a certain switching policy $\boldsymbol{g}$ is given by another system of equations denoted as $(\Sigma^{\boldsymbol{g}}, \mathrm{C}, \boldsymbol{\lambda})$, under the arrival rates $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N)$, and subject to the feasibility constraints C; this is given later in the section. In addition to the same set of assumptions made in the single channel model, we assume that the system has K channels, indexed by the set $[K] = \{1, 2, \ldots, K\}$.

The fundamental conceptual issue accompanying channelization is the notion of a channel switching policy, either centralized or distributed, that introduces channel occupancy and packet assignment distributions for each node. An additional technical issue induced by channelization is the heterogeneity of embedded time units among different channels. Since the slot length in a channel is by nature a random variable that depends on random packet arrivals, channels are in general strongly asynchronous in the embedded time units. Thus, as nodes switch among channels, we may need to switch the corresponding reference of embedded time in the slot based analysis. We therefore define the notion of a slot in different contexts as follows.

Definition V.11. Consider the virtual backoff timer defined earlier separately for a *single channel*. A *channel-slot (c-slot)* is defined as the time interval between two consecutive decrements on this virtual timer for a given channel.

Definition V.12. Consider a virtual backoff timer *at each node* that counts down indefinitely according to the node's backoff state, and is synchronized to the virtual timer of the channel in which the node resides and is done upon switching. A *node-slot (n-slot)* is defined as the time interval between two consecutive decrements on a given node's virtual backoff timer.

Remark V.13. There is no inherent difference between the two types of slots. However, this differentiation of time references becomes crucial when we define quantities based on the random embedded time. This observation will be made more concrete in the analysis. We shall also omit the explicit association of a channel (node) index with a slot whenever it does not cause ambiguity.

A channel switching or scheduling policy $\boldsymbol{g}$ induces a number of distributions related to $\Sigma^{\boldsymbol{g}}$. Denote by $\mathcal{Q}_i^n(j) = \{q_i^{(k)}(j), k \in [K]$, where $q_i^{(k)}(j)$ is the probability that node i is on channel k at the beginning of its j-th n-slot. $\mathcal{Q}_i^n(j)$ is referred to as the the channel occupancy distribution *in n-slots* of node i in the j-th n-slot.

Denote by $\mathcal{Q}_i^c(j) = \{\hat{q}_i^{(k)}(j), k \in [K]\}$, where $\hat{q}_i^{(k)}(j)$ is the probability that node i is on channel k at the beginning of its j-th c-slot. $\mathcal{Q}_i^c(j)$ is referred to as the channel occupancy *profile* of node i at the j-th c-slot. Note that $\mathcal{Q}_i^c(j)$ is not necessarily a distribution and $\sum_{k\in[K]} \hat{q}_i^{(k)}(j)$ need not be 1 for a given j.

Denote by $\mathcal{Q}_i^p(\ell) = \{\tilde{q}_i^{(k)}(\ell), k \in [K]\}$, where $\tilde{q}_i^{(k)}(\ell)$ is the probability that the ℓ-th packet of node i is served on channel k, and $\mathcal{Q}_i^p(\ell)$ is referred to as the packet assignment distribution of node i.

We have the following assumptions on policy $\boldsymbol{g}$.

Assumption V.14.

(A5) *Under $\boldsymbol{g}$, Bianchi's approximation is still satisfied.*

*(**A6**)* $\boldsymbol{g}$ *is independent of the binary state of the queue at any node (empty vs. non-empty).*

*(**A7**)* $\boldsymbol{g}$ *is* nonpreemptive *in a channel for the entire service process of a packet; that is, a channel-switching decision is only made before or after the service process of a packet.*

*(**A8**) The corresponding steady state distribution of $\mathcal{Q}_i^n(j)$, $\mathcal{Q}_i^c(j)$ and $\mathcal{Q}_i^p(\ell)$ exist under $\boldsymbol{g}$, denoted by $\mathcal{Q}_i^n$, $\mathcal{Q}_i^c$ and $\mathcal{Q}_i^p$, and they agree with the respective limits as their respective arguments tend to infinity*[2]

Similar as in single channel analysis, we impose the Markovian assumption on the discrete-time queueing process $\hat{Q}_i^{(k)}(n)$, which is the embedded process of $Q_i(t)$ (queue state of node i) sampled at the boundaries of c-slots of channel k, and define $\hat{\rho}_i^{(k)} = \lim_{n\to\infty} P(\hat{Q}_i^{(k)}(n) > 0)$. Also, let $\tau_i^{(k)}(n)$ be the probability that node i initiates a transmission attempt in the n-th c-slot of channel k. We then have the following lemma; its proof is similar to that of Lemma V.6 (based on **A6** and **A8**) and omitted.

Lemma V.15. *$\tau_i^{(k)} := \lim_{n\to\infty} \tau_i^{(k)}(n)$ exists and is given by $\tau_i^{(k)} = \hat{q}_i^{(k)}\hat{\rho}_i^{(k)}/\overline{W}_i^{(k)}$, where $\overline{W}_i^{(k)} := \frac{\mathbb{E}[N_i^{s,(k)}]}{\mathbb{E}[N_i^{tx,(k)}]}$ is the average backoff window size of node i on channel k, with $N_i^{s,(k)}$ and $N_i^{tx,(k)}$ defined in parallel as in the single channel case.*

Remark V.16. Under **A7**, $\overline{W}_i^{(k)}$ is given by

$$\overline{W}_i^{(k)} = \frac{1}{2}\left[W\left((1-p_i^{(k)})\sum_{j=0}^{m-1}(2p_i^{(k)})^j + (2p_i^{(k)})^m\right) + 1\right],$$

where $p_i^{(k)}$ is the probability of collision on channel k given a transmission attempt and W is the initial backoff window size.

[2] These limiting quantities are related by well-define correspondences, which are detailed in [71], and those relations are used to numerically evaluate the stability region equation for a multi-channel system presented in this section.

Given any scheduling policy $\boldsymbol{g}$, let $\Lambda^{\boldsymbol{g}}$ be the corresponding stability region, and we have the following theorem characterizing $\Lambda^{\boldsymbol{g}}$.

Theorem V.17. *$\boldsymbol{\lambda} \in \Lambda^{\boldsymbol{g}}$ if and only if for any solution $\boldsymbol{\tau} = (\boldsymbol{\tau}^{(k)}, k \in [K])$ where $\boldsymbol{\tau}^{(k)} = (\tau_i^{(k)}, i \in [N]) \in [0,1]^N$ to the following system of equations $\Sigma^{\boldsymbol{g}}$ given $\boldsymbol{\lambda}$,*

$$\Sigma^{\boldsymbol{g}} : \begin{cases} \tau_i^{(k)} = \dfrac{\hat{q}_i^{(k)} \hat{\rho}_i^{(k)}}{\overline{W}_i^{(k)}}, \ \forall i, k & \text{(a)} \\ p_i^{(k)} = 1 - \prod_{j \neq i} (1 - \tau_j^{(k)}), \ \forall i, k & \text{(b)} \\ \rho_i = \min \left\{ \dfrac{\lambda_i}{P} \sum_{k \in [K]} \left[\tilde{q}_i^{(k)} \left(\dfrac{\overline{W}_i^{(k)} - 1}{1 - p_i^{(k)}} \mathbb{E} S_{i,Q,\overline{Tx}}^{(k)} + T_c^{(k)} \dfrac{p_i^{(k)}}{1 - p_i^{(k)}} + T_s^{(k)} \right) \right], 1 \right\}, & \\ \forall i, k & \text{(c)} \end{cases}$$

we have $0 \leq \rho_i < 1$ for all $i \in [N]$, where P is the packet payload size, and $\mathbb{E} S_{i,Q,\overline{Tx}}^{(k)}$ is the conditional average length of a c-slot on channel k given that the queue at node i is non-empty but i does not transmit in this slot. Let

$$\mathrm{C} : \begin{cases} 0 \leq \tau_i^{(k)} \leq 1, \ \forall i, k & \text{(i)} \\ 0 \leq \rho_i < 1, \ \forall i & \text{(ii)} \end{cases}$$

We shall denote this constrained system of equations by $(\Sigma^{\boldsymbol{g}}, \mathrm{C}, \boldsymbol{\lambda})$.

Proof. The proof is an immediate extension of the proof of Theorem V.7, given assumptions on $\boldsymbol{g}$. □

The existence of a solution to $\Sigma^{\boldsymbol{g}}$ can be similarly established using Brouwer's fixed point theorem. We next study its uniqueness and the throughput optimality of a switching policy by resorting to an approximation given below, due to the complexity of $\Sigma^{\boldsymbol{g}}$. For the rest of this section, we shall limit our discussion to the symmetric case where the channels have the same bandwidth and the system uses the same parameterization in all channels. We extend our discussion to more generic settings in the next section.

Definition V.18. A scheduling policy is *unbiased* if the stationary channel occupancy distribution induced by such a policy is identical for every node, i.e., $q_i^{(k)} = q^{(k)}$ for all $i \in [N]$ and $k \in [K]$. It is denoted by $\boldsymbol{g}^U$, and the space of unbiased policies is $\mathcal{G}^U$.

We can obtain an approximation to $(\Sigma^{\boldsymbol{g}^U}, \mathrm{C}, \boldsymbol{\lambda})$ similarly as we did for Σ, using $\hat{q}^{(k)} \approx \tilde{q}^{(k)} \approx q^{(k)}$:

$$\widetilde{\Sigma}^{\boldsymbol{g}^U} : \begin{cases} \tau_i^{(k)} = \dfrac{q^{(k)}\rho_i}{\frac{W+1}{2}} & \text{(a)} \\ \rho_i = \dfrac{\lambda_i}{P} \displaystyle\sum_{k\in[K]} \left\{ q^{(k)} \left[\frac{W-1}{2} \Big(\sigma + T \sum_{j\neq i} \tau_j^{(k)} \Big) + T \Big(1 + \sum_{j \neq i} \tau_j^{(k)} \Big) \right] \right\} & \text{(b)} \end{cases}$$

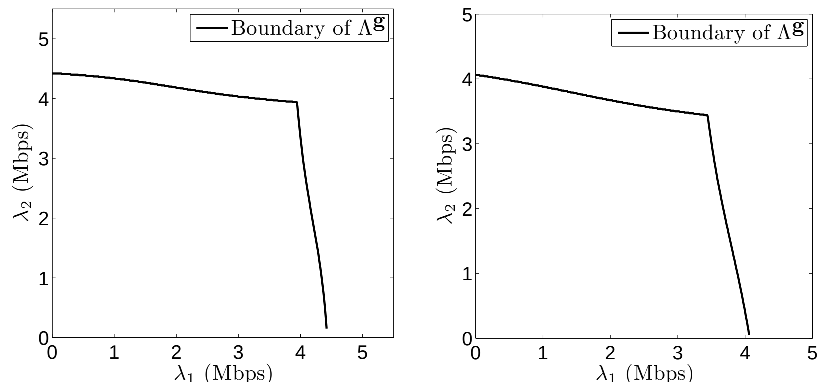
and we have the following result.

Theorem V.19. *Consider a system modeled by* $\widetilde{\Sigma}^{\boldsymbol{g}^U}$ *and the associated stability region* $\Lambda^{\boldsymbol{g}^U}$. *For all sufficiently large initial window sizes* W, *(i) the system of equations* $(\Sigma^{\boldsymbol{g}^U}, \boldsymbol{\lambda})$ *admits a unique solution, and (ii)* $\boldsymbol{g}^U$ *is throughput optimal within the class* $\mathcal{G}^U$ *if* $q^{(k)} = \frac{1}{K}$ *for all* k. *These are referred to as* equi-occupancy *policies.*

Proof. We omit the proof on uniqueness, which is similar to the single-channel case; see Appendix C for the proof on throughput optimality. □

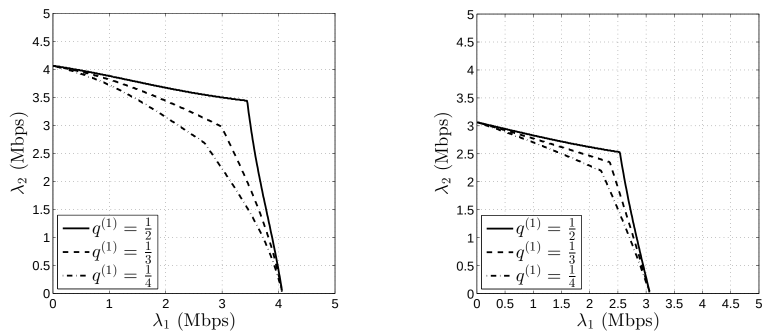
The above results provide the following insights in addition to what we have observed in the single-channel case. Firstly, it's worth noting that $\Sigma^{\boldsymbol{g}}$ reduces to Σ in the single-channel case by properly configuring related parameters, and $\Sigma^{\boldsymbol{g}}$ thus constitutes a unified framework in describing the stability region of 802.11 DCF.

Secondly, the uniqueness of the solution to $(\Sigma^{\boldsymbol{g}^U}, \boldsymbol{\lambda})$ is in fact true for even small windows. As an example, in Figure 5.5, we plot the numerical boundaries of stability regions for various window settings with equal channel occupancy. Compared to results in the single-channel case, convexity of the stability region is observed even with small backoff windows in the two-channel case. Also, the numerical multi-equilibrium phenomenon disappears in this case. One way to explain this is by considering the discounting effect of channelization on the attempt rate. The attempt rate of each node in a channel is discounted by the occupancy probability in that channel. As discussed in the single-channel case, the attempt rate is roughly upper bounded by the reciprocal of the average backoff window size. Hence channelization has the effect of window expansion. The same explanation also applies to the observation that the stability region in a multi-channel system is nearly always convex.



(a) $W = 8, m = 5$ (b) $W = 32, m = 5$

Figure 5.5: The stability region of two-channel 802.11 DCF under the equi-occupancy policy.



(a) $W = 32, m = 5$ (b) $W = 128, m = 5$

Figure 5.6: Throughput optimality of equi-occupancy distribution.

Thirdly, given symmetric channelization, equal occupancy time is equivalent to equal packet assignment in each channel. The optimality of equi-occupancy policies therefore confirms the intuitive notion that load balancing (either in the number of active nodes or in the amount of date flow) optimizes the system performance in terms of expanding the stability region. In Figure 5.6, we plot the analytical boundaries of stability regions corresponding to different unbiased policies in two scenarios. As can been seen, the equi-occupancy policy results in a stability region that is the superset of those of the other unbiased policies. It is also worth noting that as the backoff window increases, the gap between the superset region and other inferior regions decreases, as the reciprocal of the window size becomes the dominant factor in upper bounding the attempt rate.

5.6 Applicability and Implementation of Unbiased Policies in Both Symmetric and Asymmetric Systems

In this section we discuss the applicability of the class of unbiased policies. We then present a number of practical implementations and their use in both symmetric and asymmetric systems.

5.6.1 Unbiased Policies

We have so far restricted our policy space to unbiased policies that induce a node-independent channel occupancy or packet assignment distribution. Note that while nodes in the same system are typically programmed with the same protocol stack, the same protocol may not necessarily yield the same statistical behavior among different nodes. Nevertheless, there are a number of circumstances in which node-independent behaviors are induced, which justifies our focus on unbiased policies. Firstly, if the protocol explicitly prescribes packet allocation to each channel, the resulting packet assignment distributions are identical for all nodes. Secondly, if nodes have identical arrival processes, they then have unbiased behavior as well. Unbiasedness can also be observed in a saturated network (however, such a system is unstable).

More generally, we note that when a node is active (i.e., its queue is non-empty and it is in the service process), from a mean-field point of view the channel conditions observed by this node is fully characterized by $p_i^{(k)}$ for each k (as a result of the decoupling assumption), which is a function of $\tau_j^{(k)}$ for all $j \neq i$. Therefore, the set of attempt rates $\{\tau_i^{(k)}; \forall i, \forall k\}$ characterizes the contention condition in the system. If nodes are asymptotically symmetric, that is, $\lim_{N\to\infty} \tau_i^{(k)}/\tau_j^{(k)} = 1$, for all $i \neq j$ and k, then we have

$$\lim_{N\to\infty} \frac{p_i^{(k)}}{p_j^{(k)}} = \lim_{N\to\infty} \frac{1-\prod_{l\neq i}(1-\tau_\ell^{(k)})}{1-\prod_{l\neq j}(1-\tau_\ell^{(k)})} = 1 + \lim_{N\to\infty} \frac{A(\tau_j^{(k)}-\tau_i^{(k)})}{A\tau_i^{(k)}+(1-A)} = 1,$$

where $A = \prod_{l\neq i,j}(1-\tau_\ell^{(k)})$. In this case we may consider the behavior induced by the underlying protocol on each node identical, and the corresponding policy unbiased. Note that the decoupling assumption is regarded as asymptotically true for a large number of nodes, so we may consider the asymptotic symmetry as an adjoint condition if we impose the decoupling approximation in modeling.

5.6.2 Practical Implementation of Throughput Optimal Unbiased Policies: Symmetric Channels

We have shown that when channels are symmetric the optimal switching policy within the class of unbiased policies is the equi-occupancy policy that balances load precisely. When channels are asymmetric, i.e., have different bandwidths, it is natural to expect that a load balancing policy yields throughput optimal performance, and to interpret a balanced load as having a packet assignment distribution proportional to the channel bandwidths. We shall see that this interpretation is reasonable though not precise.

We begin by commenting on how such policies may be realized in a symmetric system. We describe two very simple heuristics that implement an unbiased policy, and in particular, the equi-occupancy policy when channels are symmetric. The description is given in the two-channel case for simplicity. The first is called SAS (switching after success), and the second SAC (switching after collision). In both schemes, a switching probability is assigned to each backoff stage. Under SAS (resp. SAC), a node switches to the other channel with probability $\alpha_\ell^{(k)}$ upon a successful transmission (resp. collision) if it is at the ℓ-th backoff stage on channel k when this success (resp. collision) occurs. In addition, in SAC, after switching to the other channel, a node does not reset its backoff stage; instead, it continues the exponential backoff due to the last collision. Note that SAS can be used to implement any arbitrary packet assignment distribution (and thus load distribution), which is a useful feature when we proceed to the implementation under asymmetric channels. This is because with the assumption of nonpreemptiveness of the policy, i.e., **A7**, switching after each successful transmission is equivalent to assigning packets.

These two schemes heuristically implement the equi-occupancy policy in the following sense, when the switching probability profiles are identical in all channels and the channels are symmetric. Consider the two-dimensional Markov chains for a two-channel system in the form of Bianchi's model [45], where each state in one channel has a mirror state in the other. Since for both SAS and SAC, the corresponding Markov chain is irreducible with a finite number of states, using the argument of symmetry, the symmetric solution is the unique stationary distribution that reflects equi-occupancy. It should be noted however that neither of the above is a perfect solution and the key may be a proper combination of the two. The problem with SAS is that it can result in empty channels (the node that succeeded in the transmission happens to be the only node in that channel). When this happens nodes can tend to cluster in the non-empty channel for significant periods of time due to collision and backoff, while our mean field Markov analysis implicitly assumes no channels are empty for long. On the other hand, the problem with SAC (SAC rarely results in empty channels and avoids clustering in one channel) is that it interrupts the service process of a packet in a given channel, thus violating the nonpreemptive assumption about the policy.

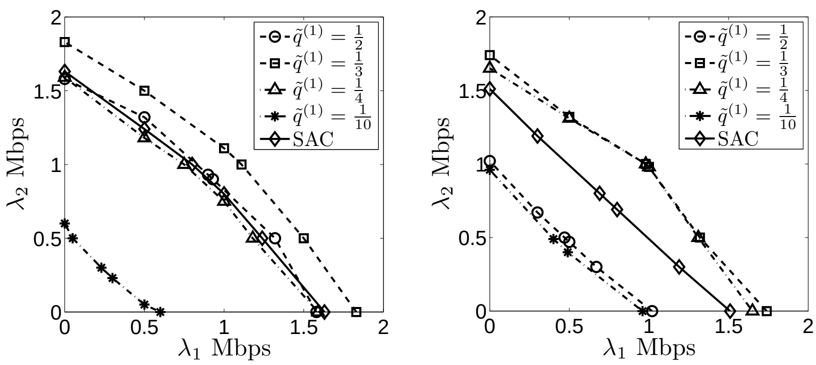
Compared to the decentralized implementation in the slotted Aloha simplification, we note both schemes similarly use the congestion level as payoff of the residing channel, though in a vague way using the backoff stage as an indicator, which agrees with our conjecture then. Also, both schemes rely solely on the real-time information of the residing channel, conceptually a further reduced version of the simple learning algorithm we presented in Chapter IV. Furthermore, SAC is consistent with our setup on the update opportunities, where a channel with more nodes statistically has a greater number of updates, and SAS has a bounded number of updates in each slot. Meanwhile, we are also aware of discrepancies between the two heuristics and the implementation in Chapter IV. The number of updates in a slot scales w.r.t. the population size in SAC and the number of updates in a channel is roughly inversely proportional to the local population size in SAS.

It is also worth noting that when SAS or SAC implements the equi-occupancy policy, or more generally known occupancy (or packet assignment) distributions, our model and assumptions admit an $M/M/1$ type of delay analysis. For instance, the average packet delay of a stable node i is given by $\frac{\rho_i}{\lambda_i(1-\rho_i)}$ and can be numerically evaluated through the stability equations.

5.6.3 Practical Implementation of Throughput Optimal Unbiased Policies: Asymmetric Channels

We next proceed to asymmetric channels and examine how these two heuristics perform in this setting, and in doing so also empirically examine when the stability region is maximized. In particular, we focus on the performance of a policy when the majority of the nodes have similar arrival rates, and we examine the advantage of load balancing in improving stability. In our experiment, we fix 10 nodes with an arrival rate 0.5Mbps that creates a mean-field background in a two-channel network while inspecting the stability region of another two nodes, which is the intersection of the aggregate stability region with the plane of these two nodes' arrival rates. All nodes use the same policy in a single experiment.

In Figure 5.7, we plot the empirical boundary of stability regions under different packet assignment distributions (implemented using SAS). As shown, policies with packet assignment ratio close to the bandwidth ratio indeed result in larger stability regions. However, while it seems safe to claim that properly balancing active time among channels according to their bandwidths improves the system performance, it remains unknown whether an exact match in load assignment is the optimal policy due to the nonlinearity of slot length in each channel w.r.t. active nodes. In addition, in practice we may not even know the effective bandwidth of each channel when channel conditions are imperfect.

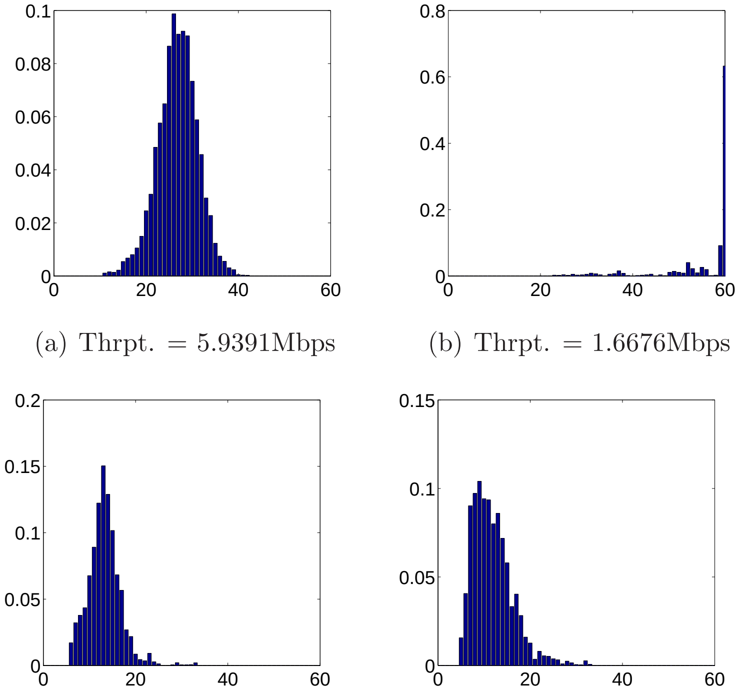


(a) The bandwidth ratio = 1:2 (b) The bandwidth ratio = 1:3

Figure 5.7: The intersection of simulated stability region with the plane of arrival rates of the two nodes under inspection.

It is therefore highly desirable to have an adaptive mechanism that dynamically adjusts the load distribution in practical implementation. Below we show that SAC to a large extent can achieve this goal, with the reason being that collision rate reflects the contention level and bandwidth information. Figure 5.7 also shows the empirical stability region obtained using SAC with a switching probability at the ℓ-th backoff stage $\alpha_\ell^{(k)} = \frac{\ell}{m}$ for all k, where m is the maximum backoff stage. SAC is clearly not optimal, but it maintains good performance under different bandwidth ratios.

We further highlight the adaptiveness of SAC in comparison to SAS. Assume that the active node population in each channel is the same and static, given then the same period of time, faster channels experience more transmission successes than slower ones. Therefore, if a SAS-like switching policy is adopted for a relatively congested network, nodes would cluster in the slower channels and the throughput performance degrades significantly. However, if the congestion is due to bandwidth asymmetry, then this is reflected in the collision rate of transmission, which in turns triggers channel reallocation under SAC. We illustrate this point using the following experiment. Consider a two-channel system with strongly asymmetric channels, where the bandwidth of channel 1 (2) is 1Mbps (10Mbps). The system consists of 60 nodes each with an arrival rate 0.1Mbps, and this aggregate arrival rate (6Mbps) is slightly below the empirical saturation throughput under this setting. In the first test, we compare the resulting distribution of number of nodes on channel 1 between SAC and SAS with the switching probability $\alpha_\ell^{(k)} = 0.5$ for all stages in both channels, and we repeat the inspection with the switching probability $\alpha_\ell^{(k)} = \frac{\ell}{m}$ at stage ℓ in the second test; the duration of simulation is 180 seconds. The switching probability profile in the first test can be regarded as a blind configuration, while the second profile can be taken as an adaptive configuration that partially incorporates collision history into switching decisions. In Figure 5.8, we plot the histograms of the number of nodes on channel 1, as well as the empirical throughput obtained. As can be seen, the blindly configured SAS drives nodes to cluster in the slower channel, while SAC avoids this problem. Interestingly, SAS has comparable performance as SAC if we adjust the switching probabilities as we did in the second test, which reflects the congestion level in the residing channel, and both distributions "match" the bandwidth ratio. It suggests that while SAS is not as adaptive as SAC, it remains a valid alternative implementation and could achieve comparable performance when configured appropriately, as did above.



(a) Thrpt. = 5.9391Mbps (b) Thrpt. = 1.6676Mbps

(c) Thrpt. = 5.9720Mbps (d) Thrpt. = 5.9811Mbps

Figure 5.8: Histogram of node population in the slower channel: (a) SAC ((b) SAS) with $\alpha_\ell = 0.5$; (c) SAC ((d) SAS) with $\alpha_\ell = \frac{\ell}{m}$.

5.6.4 Fairness under Throughput Optimal Policies

The general philosophy of SAS is that a node immediately vacates a channel in which it just had a success so other nodes can have a chance, while that of SAC is to keep using that channel until it gets inferior. While at opposite ends of the spectrum, this altruism and egoism respectively achieves the same system level fairness when universally adopted by all nodes in the network due to symmetry[3].

To illustrate further, consider a possibly asymmetric two-channel system with a mixture of saturated and unsaturated nodes, and consider two notions of fairness.

[3] *Strategic* behavior could lead to unfair advantage if users deviate from the preset rule. Consider for instance a two-channel example where all but one node adopt SAS thus clustering in an inferior channel, while one node persists in the good channel using SAC.

Under the first notion, fairness is measured by the individual throughput achieved by a node, compared to other similarly loaded nodes. For stable nodes, their throughput is simply their arrival rates. For saturated nodes, their attempt rates become essentially the same after queues have built up. This together with the fact that the implementation of SAS and SAC are not user-specific suggests the individual throughput is identical among saturated nodes.

Under the second notion, we measure fairness by the portion of a user's packets served in the better channel. Recall that SAS can be used to implement any arbitrary packet assignment distribution by tuning the conditional switching probabilities at each backoff stage after a successful transmission. For instance, if the switching probabilities are set to 1 in the worse channel at all stages while 1/2 in the better one, each node should then have on average 2/3 of its packets served in the better channel in the long term. This is independent of the arrival process or attempt rate of any node, and hence this type of fairness is also achieved.

5.7 Signal Quality plus Congestion Level in Channel Selection

Our primary intention is to study how congestion should be factored into switching decisions in a multi-channel system, and have so far assumed a perfect channel condition in terms of signal quality. In this section we consider the impact of considering congestion *in addition* to signal quality in making channel switching decisions. Below we first consider extending the current model to include packet loss due to poor channel/signal quality, and then empirically study how SAS and SAC perform under imperfect channel conditions compared to a switching policy that solely relies on signal quality estimates.

Different signal quality can be captured by a probability of packet failure loss for each transmission attempt, independent from losses due to collision, denoted by $\pi^{(k)}$ for channel k. We consider two cases depending on whether we shall assume that a node can distinguish a collision loss from a packet failure loss due to poor signal quality. In the first case when a node is able to distinguish the two, then Automatic Repeat reQuest (ARQ) can be applied upon a failed transmission within the same channel reservation (i.e., a node does not release the channel upon a packet failure but will continue to retransmit). For simplicity we shall assume there is no re-try limit, and thus the introduction of packet failure losses only affects the duration of a data session after a successful channel reservation, which was denoted by $T_s^{(k)}$ in the origin model for a successful transmission. This effectively leads to asymmetric channels even if they have the same amount of bandwidth. Since the duration of a single data session is generally much greater than the channel coherence time, we shall assume that packet failures occur independently in each re-transmission attempt with probability $\pi^{(k)}$. The number of retransmissions then follows a geometric distribution, and the expected duration of a data session after a successful reservation of channel k is given by $\frac{T_q^{(k)}}{1-\pi^{(k)}} + T_s^{(k)}$, where $T_q^{(k)}$ is the duration of a transmission that resulted in packet failure.

In the second case, when a node is not able to distinguish a packet failure loss from collision, it will simply regard each unsuccessful transmission attempt as being involved in a collision. As a result the conditional collision probability given a transmission attempt in $\Sigma^{\mathbf{g}}(\mathrm{b})$ is updated as

$$p_i^{(k)} = 1 - (1 - \pi^{(k)}) \prod_{j \neq i} (1 - \tau_j^{(k)}).$$

In both cases, the original model can be extended to compute the corresponding stability regions.

We now numerically compare the proposed congestion-aware switching algorithm

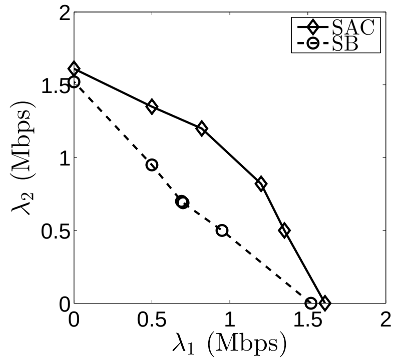


Figure 5.9: Congestion-based vs. signal-based: stability region.

to a method that uses only signal quality. Consider three channels with equal bandwidth (a third of 11Mbps) but different signal qualities modeled as packet loss probabilities for a given transmission attempt (0.1, 0.2 and 0.3 for the three channels, respectively). Assume nodes can tell collision loss from failure loss. We fix 20 nodes each with an arrival rate 0.1Mbps that creates a mean-field background as in the previous section, while tuning the arrival rates of two additional nodes. We then inspect the stability region projected onto the plane where these two nodes' arrival rates reside.

In one scenario, all nodes use SAC together with ARQ within each data session until success. In the other scenario, all nodes use a signal-based (SB) switching method that essentially performs an online estimate of the packet failure loss rate in each channel, by tracking the total number of successful transmissions and the total number of transmission attempts within each data session (after successful channel reservation), and switches to (or remains in) the channel with the lowest current estimate upon each successful packet transmission. In the long run one expects nodes to cluster in the best channel even while it gets more congested. This is indeed observed in our simulation; the resulting stability regions are depicted in Figure 5.9[4]. We also report the average number of nodes in each channel at near-saturated points

[4]Note that only a limited number of boundary points are identified to sketch the stability regions; the connecting lines are hence not necessarily the exact boundary.

during a simulation of 30 seconds in Table 5.1, which confirms our intuition.

Table 5.1: Congestion-based vs. signal-based: node distribution.

Channel		1	2	3	
Node	SAC	6.69	7.21	8.10	$(\lambda_1, \lambda_2) = (1, 1)$
Distr'n	SB	14.22	5.30	2.48	$(\lambda_1, \lambda_2) = (.6, .6)$

In this study we have used a rather simple signal-based algorithm. Nevertheless it validates our observation that considering *only* signal quality can be a very detrimental thing to do when there is significant congestion in the system.

CHAPTER VI

Conclusion and Future Work

6.1 Summary of Main Contributions

In this dissertation, we studied channel switching decision problems in multi-channel wireless networks, where channel switching is viewed as a defense strategy against a jamming attacker in adversarial environment, and as a spectrum resource allocation technique in multi-user systems. In addition to the application of classic decision and optimization theory in our solution framework, we aimed to formulate and address these problems from a learning perspective, which in other words is on the adaptive and evolving responses of individuals based on limited but accumulating information in repeated interaction. In the context of system engineering, interacting players are typically given by non-strategic and programmable devices. Hence, unlike the study of learning theory in the economic or the psychological discipline, the justification of rationality on the individual choice of learning algorithms, which is in essence the information fusion and strategy update law, can be much relaxed in system engineering, which offers a new approach to decentralized implementations of optimal control.

In our first part of study, we presented two problems in Chapter II and III on using channel hopping as jamming defense. We investigated the optimal responses from the user with respect to different levels of knowledge on the adversary and criteria of optimality. The attacker-user model we used is also often known as the pursuit-evasion model in the literature, and the framework and the results we developed in this part is applicable to two-player interaction in problems in this general context. When the interference to a given user is not malicious but rather the results of contention among a group of legitimate users, channelization can become an effective resource sharing mechanism other than defense strategy. From this point of view, we studied the decentralized channel switching of users for throughput enhancement in random access systems. Our main focus was to develop a theory on throughput optimal switching strategies in the industrial standard 802.11 DCF scheme (Chapter V), and we started from an Aloha simplification with homogeneous users (Chapter IV), which allowed an elegant learning formulation and provided plentiful insights for the technically more complex DCF. As our later study on DCF showed, the learning technique does provide a guideline to implementing decentralized optimal control.

6.2 Future Work

Future work remains in the following areas.

- Rendezvous problem in networks. Seeking rendezvous between a pair of transmitter and receiver in a multi-channel system can be a critical issue in opportunistic access schemes, when a control channel or centralized message exchange is infeasible. The formulation we had for the jamming defense problem can be naturally extended to the rendezvous problem. However, the jamming defense problem has the particular zero-sum nature, and this was the key to enable basic convergence of learning to establish optimality results. In contrast, players in the rendezvous problem seek to coordinate, and the underlying game is general-sum. While the action pairs in pure Nash equilibria are socially optimal, no-regret learning algorithms are no longer guaranteed to converge to such profiles of actions. It is therefore an open question if any general or ad hoc decision-making/learning protocol can be devised.

- Channel allocation in multihop networks. The problems we have studied so far assume the one-hop or start-shaped topology. While this simple type of networks characterizes a number of application scenarios, specifically for networks with infrastructure, multihop networks have been proposed and widely studied as an option for infrastructure-free wireless networks. With channelization, the spatial reuse and diversity can be further enhanced in multi-hop networks and we may expect a considerable performance boost with intelligent and dynamic spectrum sharing, namely transmission scheduling and channel allocation. Throughput optimal transmission scheduling in multihop random access networks has been extensively studied in the literature based on Lyapunov analysis. One step beyond the study we presented in Chapter IV and V could be a revisit of the scheduling problem with the additional degree of freedom in channel allocation. Our interest lies in the nature of the centralized solution as well as practical decentralized implementation.

APPENDICES

APPENDIX A

Supplements to Chapter II

Proof of Theorem II.7

Proof. Define $\Delta_{ij}(t) := G_i(t) - G_j(t)$. Then,

$$\alpha_k(t) = \frac{1}{\sum_{j=1}^{m} a^{\Delta_{jk}(t-1)}},$$

and

$$r^\pi(t) = \frac{\sum_{j\neq\pi(t)} a^{\Delta_{j\pi(t)}(t-1)}}{1 + \sum_{j\neq\pi(t)} a^{\Delta_{j\pi(t)}(t-1)}}.$$

Let $\mathcal{K}(t) = \arg\min_{k\in[K]} G_k(t)$, and define $\mathcal{T} = \{t \leq T : \max_{k\notin\mathcal{K}(t)} \Delta_{k,j}(t) \geq 2, j \in \mathcal{K}(t)\}$. Suppose that $\mathcal{T} \neq \emptyset$, and let $t_0 = \min \mathcal{T}$. Then, either (C1) there exists some time t_1 with $t_0 < t_1 \leq T$ when some location $j \in \mathcal{K}(t_0)$ is selected for the first time after t_0 by the evader or (C2) any location $j \in \mathcal{K}(t_0)$ is never selected by the horizon T.

Consider first the case (C1). Without loss of generality, assume that the location selected at $t_1 - 1$ is 2 and 1 is chosen at t_1. Let $\Delta_{ij}(t_1 - 1) = d_{ij}$. Then,

- $\Delta_{ij}(t_1) = \Delta_{ij}(t_1 + 1) = d_{ij}$ for all $i, j \geq 3$;

- $\Delta_{1j}(t_1) = d_{1j}$ for all $j \geq 3$, $\Delta_{12}(t_1) = d_{12} - 1$, $\Delta_{1j}(t_1+1) = d_{1j}+1$ for all $j \geq 3$, and $\Delta_{12}(t_1) = d_{12}$;
- $\Delta_{2j}(t_1) = d_{2j} + 1$ for all $j \neq 2$, $\Delta_{2j}(t_1 + 1) = d_{2j} + 1$ for all $j \geq 3$, and $\Delta_{21}(t_1 + 1) = d_{21}$.

Consider now a change of policy by selecting location 1 at $t_1 - 1$ and location 2 at t_1. Denote Δ under this new policy by Δ'. Then,

- $\Delta'_{ij}(t_1) = \Delta'_{ij}(t_1 + 1) = d_{ij}$ for all $i, j \geq 3$.
- $\Delta'_{1j}(t_1) = d_{1j} + 1$ for all $j \geq 2$, $\Delta'_{1j}(t_1 + 1) = d_{1j} + 1$ for all $j \geq 3$, and $\Delta'_{12}(t_1) = d_{12}$;
- $\Delta'_{2j}(t_1) = d_{2j}$ for all $j \geq 3$, $\Delta'_{21}(t_1) = d_{21} - 1$, $\Delta'_{2j}(t_1+1) = d_{2j}+1$ for all $j \geq 3$, and $\Delta'_{21}(t_1 + 1) = d_{21}$.

Hence, this change of policy only affects the reward of the evader collected at $t_1 - 1$ and t_1. Denote by r' the reward under this alternative policy, and we have

$$
\begin{aligned}
& r'(t_1 - 1) + r'(t_1) - r(t_1 - 1) - r(t_1) \\
&= \frac{\sum_{k\geq 3} a^{d_{k1}} + a^{d_{21}}}{1 + \sum_{k\geq 3} a^{d_{k1}} + a^{d_{21}}} + \frac{\sum_{k\geq 3} a^{d_{k2}} + a^{d_{12}+1}}{1 + \sum_{k\geq 3} a^{d_{k2}} + a^{d_{12}+1}} \\
&\quad - \frac{\sum_{k\geq 3} a^{d_{k2}} + a^{d_{12}}}{1 + \sum_{k\geq 3} a^{d_{k2}} + a^{d_{12}}} - \frac{\sum_{k\geq 3} a^{d_{k1}} + a^{d_{21}+1}}{1 + \sum_{k\geq 3} a^{d_{k1}} + a^{d_{21}+1}} \\
&= \frac{1}{1 + C + a^{d_{21}+1}} + \frac{1}{1 + D + a^{d_{12}}} - \frac{1}{1 + C + a^{d_{21}}} - \frac{1}{1 + D + a^{d_{12}+1}},
\end{aligned}
$$

where $C = \sum_{k\geq 3} a^{d_{k1}}$ and $D = \sum_{k\geq 3} a^{d_{k2}}$. Note that $C = Da^{d_{21}}$ and $d_{12} = -d_{21}$. Set $d = d_{21}$, and we obtain

$$
\begin{aligned}
& r'(t_1 - 1) + r'(t_1) - r(t_1 - 1) - r(t_1) \\
&= \frac{1}{1 + Da^d + a^{d+1}} + \frac{1}{1 + D + a^{-d}} - \frac{1}{1 + Da^d + a^d} - \frac{1}{1 + D + a^{-d+1}},
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{a^d - a^{d+1}}{(1 + Da^d + a^{d+1})(1 + Da^d + a^d)} + \frac{a^{-d+1} - a^{-d}}{(1 + D + a^{-d})(1 + D + a^{-d+1})} \\
&= \frac{(a^{2d-1} - a^{d-1})(a-1)^2}{(1 + Da^d + a^{d+1})(1 + Da^d + a^d)(1 + Da^{d-1} + a^{d-1})} \\
&> 0.
\end{aligned}
$$

For (C2), it is clear that alternatively selecting location 1 at T results in a higher reward.

Therefore, the optimal policy would never allow the difference between the times that any two locations are selected to be greater than 2. In other word, the optimal policy always selects the most under-utilized location. When there are multiple locations with the same lowest number of times of the evader's presence, the evader would be indifferent in selecting any location between/among them, since locations are symmetric (and the reward is only related to the the relative difference between the numbers of location usage). $\square$

APPENDIX B

Supplements to Chapter III

Proof of Lemma III.12

Proof. Assume that there exist s and t such that $s \in \mathcal{S}_t$ and $u_0^*(s,t) = 1$. Then, there exists σ and τ such that $\sigma \in \mathcal{S}_\tau$, $t \leq \tau < T$, $u_0^*(\sigma, \tau) = 1$ and $u_0^*(s', t') = 0$ for all $s' \in \mathcal{S}_{t'}$ for all $t' > \tau$; otherwise, $u_0^*(s'', T) = 1$ for some $s'' \in \mathcal{S}_T$, which is clearly not an equilibrium strategy for the attacker. If $\tau = T - 1$, then

$$\begin{aligned} U_{T-1}^*(\sigma) &= U_T^*(f_{T-1}(\sigma, 0)) \\ &= \sum_{i=1}^{n} q_i w_i^*(f_{T-1}(\sigma, 0), T) c_i + \epsilon(f_{T-1}(\sigma, 0)) \leq v + \epsilon_{\max}. \end{aligned}$$

Consider an alternative strategy $\tilde{u}$ such that $\tilde{u} = u^*$ except that $\tilde{u}_i(\sigma, T-1) = q_i$. Then,

$$\tilde{U}_{T-1}(\sigma) := \sum_{i=1}^{n} q_i (w_i^*(\sigma, T-1) c_i + U_T^*(f_{T-1}(\sigma, i))) = v + \sum_{i=1}^{n} q_i U_T^*(f_{T-1}(\sigma, i)).$$

Let $k \in [n]$ be such that $f_{T-1}(\sigma, k) \in \mathcal{S}_T$. Then $\tilde{U}_{T-1}(\sigma) \geq v + q_k v \geq v + q_{\min} v$. Hence $\tilde{U}_{T-1}(\sigma) > U_{T-1}^*(\sigma)$, which contradicts the fact that u^* is a SPE strategy.

$$
\begin{aligned}
U^{(k)}_{\tau+k}(h^{\sigma}_{\tau+k}) &= \sum_{i=1}^{n} w^*_i(\sigma^{(k)}, \tau+k) U^{(k)}_{\tau+k+1}\Big(\langle h^{\sigma}_{\tau+k}, i, 0, f_{\tau+k}(\sigma^{(k)}, 0)\rangle\Big) \qquad \text{(B.1)}\\
&= v + \sum_{i=1}^{n} w^*_i(\sigma^{(k)}, \tau+k) \cdot \sum_{\ell=1}^{n}\sum_{j=1}^{n} w^*_\ell(f_{\tau+k}(\sigma^{(k)}, 0), \tau+k+1) q_j \cdot \\
&\quad \cdot U^{(k)}_{\tau+k+2}\Big(\langle h^{\sigma}_{\tau+k}, i, 0, f_{\tau+k}(\sigma, 0), l, j, f_{\tau+k+1}(f_{\tau+k}(\sigma^{(k)}, 0), j)\rangle\Big)\\
&= v + \sum_{i=1}^{n} w^*_i(\sigma^{(k)}, \tau+k) \sum_{j=1}^{n} q_j V_{\tau+k+2}(f_{\tau+k+1}(f_{\tau+k}(\sigma^{(k)}, 0), j)),
\end{aligned}
$$

$$
\begin{aligned}
U^{(k+1)}_{\tau+k}(h^{\sigma}_{\tau+k}) &= \sum_{i=1}^{n} q_i w^*_i(\sigma^{(k)}, \tau+k) c_i + \\
&\quad + \sum_{i=1}^{n}\sum_{j=1}^{n} w^*_i(\sigma^{(k)}, \tau+k) q_j U^{(k+1)}_{\tau+k+1}\Big(\langle h^{\sigma}_{\tau+k}, i, j, f_{\tau+k}(\sigma^{(k)}, j)\rangle\Big) \qquad \text{(B.2)}
\end{aligned}
$$

$$
\begin{aligned}
&= v + \sum_{i=1}^{n} w^*_i(\sigma^{(k)}, \tau+k) \sum_{j=1}^{n} q_j \sum_{\ell=1}^{n} w^*_\ell(f_{\tau+k}(\sigma^{(k)}, j), \tau+k+1) \cdot \\
&\quad \cdot U^{(k+1)}_{\tau+k+2}\Big(\langle h^{\sigma}_{\tau+k}, i, j, f_{\tau+k}(\sigma^{(k)}, j), l, 0, f_{\tau+k+1}(f_{\tau+k}(\sigma^{(k)}, j), 0)\rangle\Big)\\
&= v + \sum_{i=1}^{n} w^*_i(\sigma^{(k)}, \tau+k) \sum_{j=1}^{n} q_j \sum_{\ell=1}^{n} w^*(f_{\tau+k+1}(\sigma^{(k)}, 0), \tau+k+1) \cdot \\
&\quad \cdot U^{(k+1)}_{\tau+k+2}\Big(\langle h^{\sigma}_{\tau+k}, i, j, f_{\tau+k}(\sigma^{(k)}, j), l, 0, f_{\tau+k+1}(f_{\tau+k}(\sigma^{(k)}, 0), j)\rangle\Big)\\
&= v + \sum_{i=1}^{n} w^*_i(\sigma^{(k)}, \tau+k) \sum_{j=1}^{n} q_j V_{\tau+k+2}(f_{\tau+k+1}(f_{\tau+k}(\sigma^{(k)}, 0), j)) = U^{(k)}_{\tau+k}(h^{\sigma}_{\tau+k}).
\end{aligned}
$$

Thus, $\tau < T-1$. Now consider a particular subgame with the full label h^{σ}_{τ} such that $s_\tau = \sigma$. We shall alternate u^* and construct inductively a sequence of strategies that only differ from u^* within this subgame. These alternative strategies will be in $\mathcal{G}^0 - \tilde{\mathcal{G}}^0$, i.e., it can depend on the past actions instead of only the resource level, and we shall show that the last strategy of this sequence strictly improves the payoff of the attacker. To make the dependency on the full history explicit, we use the notation

$$
U_t(w, u, h_t) := \mathbb{E}\left\{ \sum_{r=t}^{T} w(h_r)^\top M u(h_r) + \epsilon(s_T) \;\middle|\; h_t \right\}
$$

for the value of the subgame labeled by h_t, and denote $u(h_t)$ as the strategy of the attacker at the node h_t of the game tree. To simplify our notation, since $w^* \in \mathcal{G}^\dagger \subseteq \tilde{\mathcal{G}}$, we shall keep write $w^*(s,t)$ as the strategy of the user at some node h_t such that $s_t = s$. Note that

$$\begin{aligned}
U^*_\tau(h^\sigma_\tau) &:= U_\tau(w^*, u^*, h^\sigma_\tau) \\
&= \sum_{i=1}^n w^*_i(\sigma,\tau) U^*_{\tau+1}\Big(\langle h^\sigma_\tau, i, 0, f_\tau(\sigma,0)\rangle\Big) \\
&= v + \sum_{i=1}^n w^*_i(\sigma,\tau) \sum_{\ell=1}^n \sum_{j=1}^n w^*_\ell(f_\tau(\sigma,0),\tau+1) q_j \cdot \\
&\quad \cdot U^*_{\tau+2}\Big(\langle h^\sigma_\tau, i, 0, f_\tau(\sigma,0), l, j, f_{\tau+1}(f_\tau(\sigma,0), j)\rangle\Big).
\end{aligned}$$

and $U^*_{\tau+2}\big(\langle h^\sigma_\tau, i, 0, f_\tau(\sigma,0), l, j, f_{\tau+1}(f_\tau(\sigma,0), j)\rangle\big)$ only depends on $f_{\tau+1}(f_\tau(\sigma,0), j)$ since $w^*, u^* \in \tilde{\mathcal{G}}_{\tau+2}$. Denote then this number by $V_{\tau+2}(f_{\tau+1}(f_\tau(\sigma,0), j))$. Hence,

$$U^*_\tau(h^\sigma_\tau) = v + \sum_{i=1}^n w^*_i(\sigma,\tau) \sum_{j=1}^n q_j V_{\tau+2}(f_{\tau+1}(f_\tau(\sigma,0), j)).$$

Let i_1 and j_1 be such that $i_1 \in \operatorname{supp}(w^*(\sigma,\tau))$ and $f_\tau(\sigma, j_1) \in \mathcal{S}_{\tau+1}$, where j_1 exists due to our assumption. Consider an alternative strategy $u^{(1)}$ such that $u^{(1)} = u^*$ except that $u^{(1)}_i(h^\sigma_\tau) = q_i$ for all $i \in [n]$ and $u^{(1)}_0(\langle h^\sigma_\tau, i, j, f_\tau(\sigma,j)\rangle) = 1$ for all $i, j \in [n]$. Then,

$$\begin{aligned}
U^{(1)}_\tau(h^\sigma_\tau) &:= U_\tau(w^*, u^{(1)}, h^\sigma_\tau) \\
&= \sum_{i=1}^n q_i w^*_i(\sigma,\tau) c_i + \sum_{i=1}^n \sum_{j=1}^n w^*_i(\sigma,\tau) q_j U^{(1)}_{\tau+1}\Big(\langle h^\sigma_\tau, i, j, f_\tau(\sigma,j)\rangle\Big) \\
&= v + \sum_{i=1}^n w^*_i(\sigma,\tau) \sum_{j=1}^n q_j U^{(1)}_{\tau+1}\Big(\langle h^\sigma_\tau, i, j, f_\tau(\sigma,j)\rangle\Big) \qquad \text{(B.3)} \\
&= v + \sum_{i=1}^n w^*_i(\sigma,\tau) \sum_{j=1}^n q_j \sum_{\ell=1}^n w^*_\ell(f_\tau(\sigma,j), \tau+1) \cdot
\end{aligned}$$

$$
\begin{aligned}
&\cdot U^{(1)}_{\tau+2}\Big(\langle h^\sigma_\tau, i, j, f_\tau(\sigma, j), l, 0, f_{\tau+1}(f_\tau(\sigma, j), 0)\rangle\Big)\\
&= v + \sum_{i=1}^{n} w^*_i(\sigma,\tau) \sum_{j=1}^{n} q_j \sum_{\ell=1}^{n} w^*_\ell(f_\tau(\sigma, j), \tau+1)\cdot\\
&\cdot U^{(1)}_{\tau+2}\Big(\langle h^\sigma_\tau, i, j, f_\tau(\sigma, j), l, 0, f_{\tau+1}(f_\tau(\sigma, 0), j)\rangle\Big)
\end{aligned}
$$

$U^{(1)}_{\tau+2}\big(\langle h^\sigma_\tau, i, j, f_\tau(\sigma, j), l, 0, f_{\tau+1}(f_\tau(\sigma, 0), j)\rangle\big)$ only depends on $f_{\tau+1}(f_\tau(\sigma, 0), j)$ and is equal to $V_{\tau+2}(f_{\tau+1}(f_\tau(\sigma, 0), j))$ by noting that $u^{(1)} \in \tilde{\mathcal{G}}_{\tau+2}$ and $u^{(1)} = u^*$ at any node h_t with $t \geq \tau + 2$ by construction, so we have

$$
U^{(1)}_\tau(h^\sigma_\tau) = v + \sum_{i=1}^{n} w^*_i(\sigma,\tau) \sum_{j=1}^{n} q_j V_{\tau+2}(f_{\tau+1}(f_\tau(\sigma, 0), j)) = U^*_\tau(h^\sigma_\tau),
$$

i.e., $u^{(1)}$ does not change the value of the subgame labeled by h^σ_τ, and also by (B.3), for each $i \in \mathrm{supp}(w^*(\sigma,\tau))$, the subgame labeled by $\langle h^\sigma_\tau, i, j, f_\tau(\sigma, j)\rangle$ can be reached with positive probability under the strategy w^* and $u^{(1)}$, and hence $U^{(1)}_{\tau+1}\big(\langle h^\sigma_\tau, i, j, f_\tau(\sigma, j)\rangle\big)$ has a positive weight in the evaluation of $U^{(1)}_\tau(h^\sigma_\tau)$ as well $U^*_\tau(h^\sigma_\tau)$ for all $i \in [n]$. Let

$$
f_{\tau,\tau+k-1}(\sigma, j_1, \ldots, j_k) := f_{\tau+k-1}(f_{\tau,\tau+k-2}(\sigma, j_1, \ldots, j_{k-1}), j_k),
$$

where $f_{\tau,\tau}(\sigma, j_1) := f_\tau(\sigma, j_1)$, and $f_{\tau,\tau-1}(\sigma) := \sigma$, and let $\sigma^{(k)} := f_{\tau,\tau+k-1}(\sigma, j_1, \ldots, j_k)$. That is, $\sigma^{(k)}$ is the resource level at $\tau + k$ when the resource level at τ is σ and the actions taken by the attacker from τ to $\tau + k - 1$ are given by $j_1, j_2, \ldots, j_k$. Also, let $h^\sigma_{\tau+k} := \langle h^\sigma_{\tau+k-1}, i_k, j_k, \sigma^{(k)}\rangle$, where i_r and j_r are chosen such that $i_r \in \mathrm{supp}(w^*(\sigma^{(r-1)}, \tau + r - 1))$ and $\sigma^{(r)} \in \mathcal{S}_{\tau+r}$ for all $r = 1, 2, \ldots, k-1$, which is feasible by our assumption.

Suppose that we have constructed a sequence of strategies $u^{(r)}$ based on $u^{(r-1)}$ for $r = 1, 2, \ldots, k$, such that $u^{(r)} = u^{(r-1)}$ except that in the subgame labeled by $h^\sigma_{\tau+r-1}$

we set $u_i^{(r)}(h^\sigma_{\tau+r-1}) = q_i$ for all $i \in [n]$ and

$$u_0^{(r)}(\langle h^\sigma_{\tau+r-1}, i, j, f_{\tau+r-1}(\sigma^{(r-1)}, j)\rangle) = 1$$

for all $i, j \in [n]$, which implies that $u^{(r)} \in \tilde{\mathcal{G}}_{\tau+r+1}$ and $u^{(r)} = u^{(r-1)}$ at all nodes h_t with $t \geq \tau + r + 1$. Suppose the constructed strategies satisfy that $U^{(r)}_{\tau+r-1}(h^\sigma_{\tau+r-1}) = U^{(r-1)}_{\tau+r-1}(h^\sigma_{\tau+r-1})$ where $U^{(r)}_{\tau+r-1}(h^\sigma_{\tau+r-1}) := U_{\tau+r-1}(w^*, u^{(r)}, h^\sigma_{\tau+r-1})$ and $U^{(0)}_\tau := U^*_\tau$ for all r, which implies $U^{(k)}_\tau(h^\sigma_\tau) = U^*_\tau(h^\sigma_\tau)$. Also, suppose each subgame labeled by $h^\sigma_{\tau+r}$ can be reached with positive probability under w^* and $u^{(r)}$ for all $r = 1, 2, \ldots, k$, which implies that $U^{(k)}_{\tau+k}(h^\sigma_{\tau+k})$ has a positive weight in in the evaluation of $U^{(k)}_\tau(h^\sigma_\tau)$.

Consider then a strategy $u^{(k+1)}$ such that $u^{(k+1)} = u^{(k)}$ except that in the subgame labeled by $h^\sigma_{\tau+k}$ we set $u_i^{(k+1)}(h^\sigma_{\tau+k}) = q_i$ for all $i \in [n]$ and

$$u_0^{(k+1)}(\langle h^\sigma_{\tau+k}, i, j, f_{\tau+k}(\sigma^{(\tau+k)})\rangle) = 1$$

for all $i, j \in [n]$. Then, $u^{(k+1)} \in \tilde{\mathcal{G}}_{\tau+k+2}$ and $u^{(k+1)} = u^{(k)}$ at all nodes h_t with $t \geq \tau + k + 2$. Consequently, we have (B.1), where $V_{\tau+k+2}(f_{\tau+k+1}(f_{\tau+k}(\sigma^{(k)}, 0), j))$ is some number that only depends on $f_{\tau+k+1}(f_{\tau+k}(\sigma^{(k)}, 0), j)$; on the other hand, we have (B.2). Hence, $U^{(k)}_\tau(h^\sigma_\tau) = U^*_\tau(h^\sigma_\tau)$. Also, by (B.2), $U^{(k+1)}_{\tau+k+1}(h^\sigma_{\tau+k+1})$ has a positive weight in in the evaluation of $U^{(k)}_\tau(h^\sigma_\tau)$, which completes our induction.

This inductive construction can proceed until $\tau + k = T - 1$, and we have $u_0^{(T-1-\tau)}(h^\sigma_{T-1}) = 1$ where $\sigma^{T-1-\tau} \in \mathcal{S}_{T-1}$. However, by further modifying $u^{(T-1-\tau)}$ as shown in the beginning of this proof, we can strictly improve $U^{(T-1-\tau)}_{T-1}(h^\sigma_{T-1})$, thus increasing $U^{(T-1-\tau)}_\tau(h^\sigma_\tau)$ so as to be greater than $U^*_\tau(h^\sigma_\tau)$, which is a contradiction to the fact that u^* is a SPE strategy. □

APPENDIX C

Supplements to Chapter V

Proof of $\rho \geq \hat{\rho}$

Proof. We first define the following stochastic processes generated by the queueing process at node i. Let $T_{i,Q}(t)$ (resp. $T_{i,\overline{Q}}(t)$) be the total length of real time periods up to time t that the queue at node i is non-empty (resp. empty) (or i is busy (resp. (idle)). Let $N_{i,Q}(t)$ (resp. $N_{i,\overline{Q}}(t)$) be the total number of slots up to time t that the queue at node i is non-empty (resp. empty) at the beginning of slots. These processes are well-defined on the same sample space. Conditional on that the queue is stable, due to ergodicity, ρ_i and $\hat{\rho}_i$ can then be expressed respectively as

$$\rho_i = \lim_{t\to\infty} \frac{T_{i,Q}(t)}{t} = \lim_{t\to\infty} \frac{T_{i,Q}(t)}{T_{i,Q}(t) + T_{i,\overline{Q}}(t)},$$

and

$$\hat{\rho}_i = \lim_{t\to\infty} \frac{N_{i,Q}(t)}{N_{i,Q}(t) + N_{i,\overline{Q}}(t)},$$

almost surely. Let $\Delta_i(t)$ be the total time fragments of busy periods in idle slots of node i up to time t, and let $S_{i,Q}(k)$ ($S_{i,\overline{Q}}(k)$) be the length of the k-th busy (resp.

idle) slot. Quantities described above are illustrated in Figure C.1. Then, we have

$$T_{i,Q}(t) - \Delta_i(t) = \sum_{k=1}^{N_{i,Q}(t)} S_{i,Q}(k),$$

and

$$t = \sum_{k=1}^{N_{i,Q}(t)} S_{i,Q}(k) + \sum_{k=1}^{N_{i,\overline{Q}}(t)} S_{i,\overline{Q}}(k).$$

Therefore,

$$\begin{aligned}
\rho_i \geq \lim_{t\to\infty} \frac{T_{i,Q}(t) - \Delta_i(t)}{t} &= \lim_{t\to\infty} \frac{\sum_{k=1}^{N_{i,Q}(t)} S_{i,Q}(k)}{\sum_{k=1}^{N_{i,Q}(t)} S_{i,Q}(k) + \sum_{k=1}^{N_{i,\overline{Q}}(t)} S_{i,\overline{Q}}(k)} \\
&= \lim_{t\to\infty} \left[\frac{\sum_{k=1}^{N_{i,Q}(t)} S_{i,Q}(k)}{N_{i,Q}(t)} N_{i,Q}(t) \middle/ \left(\frac{\sum_{k=1}^{N_{i,Q}(t)} S_{i,Q}(k)}{N_{i,Q}(t)} N_{i,Q}(t) + \right.\right. \\
&\qquad \left.\left. + \frac{\sum_{k=1}^{N_{i,\overline{Q}}(t)} S_{i,\overline{Q}}(k)}{N_{i,\overline{Q}}(t)} N_{i,\overline{Q}}(t) \right) \right],
\end{aligned}$$

almost surely. Let $\mathbb{E}S_{i,Q}$ and $\mathbb{E}S_{i,\overline{Q}}$ be the conditional average lengths of an arbitrary slot, given that the queue at node i is non-empty or empty at the beginning of slot, respectively. We claim that $\mathbb{E}S_{i,Q} > \mathbb{E}S_{i,\overline{Q}}$. Note also that $N_{i,Q}(t) \to \infty$ and $N_{i,\overline{Q}}(t) \to \infty$ as $t \to \infty$ conditional on stability. Consequently, following ergodicity, we obtain

$$\rho_i \geq \lim_{t\to\infty} \frac{N_{i,Q}(t)\mathbb{E}S_{i,Q}}{N_{i,Q}(t)\mathbb{E}S_{i,Q} + N_{i,\overline{Q}}(t)\mathbb{E}S_{i,\overline{Q}}} \geq \lim_{t\to\infty} \frac{N_{i,Q}(t)}{N_{i,Q}(t) + N_{i,\overline{Q}}(t)} = \hat{\rho}_i.$$

When the queue is unstable, we have $\rho_i = \hat{\rho}_i = 1$. In either case, we have $\rho_i \geq \hat{\rho}_i$. It remains to justify the claim made above, which appears in the next part.

Computation of $\mathbb{E}S_{\{\cdot\}}$ and Related Quantities Given an event $\{\cdot\}$, let $P_{\text{idle};\{\cdot\}}$, $P_{\text{succ};\{\cdot\}}$ and $P_{\text{coll};\{\cdot\}}$ be the conditional probabilities that a slot is idle, that the transmission attempt in the slot is a success, and that the attempt results in a collision,

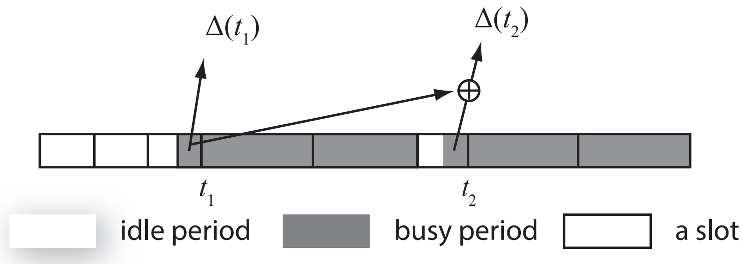


Figure C.1: Slotted time dynamics.

respectively. Notice that $P_{\text{coll};\{\cdot\}} = 1 - P_{\text{idle};\{\cdot\}} - P_{\text{succ};\{\cdot\}}$. Therefore,

$$\mathbb{E}S_{\{\cdot\}} = \sigma \cdot P_{\text{idle};\{\cdot\}} + T_s \cdot P_{\text{succ};\{\cdot\}} + T_c \cdot P_{\text{coll};\{\cdot\}}.$$

where σ, T_s and T_c are the lengths of an empty system slot, a successful transmission, and a collision, respectively. Define then by $\tau_{i,Q}$ the conditional probability that node i transmits in an arbitrary slot, given its queue is non-empty at the beginning of this slot, and hence we have $\tau_{i,Q} = \frac{1}{W_i}$. Consequently,

$$\begin{aligned}
P_{\text{idle};i,\overline{Q}} &= \prod_{j\neq i}(1-\tau_j),\\
P_{\text{succ};i,\overline{Q}} &= \sum_{j\neq i}\tau_j \prod_{l\neq i,j}(1-\tau_\ell),\\
P_{\text{idle};i,Q} &= (1-\tau_{i,Q})\prod_{j\neq i}(1-\tau_j),\\
P_{\text{succ};i,Q} &= \sum_{\ell}\tilde{\tau}_\ell \prod_{j\neq l}(1-\tilde{\tau}_\ell),
\end{aligned}$$

where $\tilde{\tau}_j = \tau_{i,Q}\mathbf{1}_{j=i} + \tau_j\mathbf{1}_{j\neq i}$. Since $P_{\text{idle};i,Q} < P_{\text{idle};i,\overline{Q}}$ and $\sigma < \min\{T_s, T_c\}$, we have $\mathbb{E}S_{i,Q} > \mathbb{E}S_{i,\overline{Q}}$ and they are both finite. Explicit expressions for other variations of $\mathbb{E}S_{\{\cdot\}}$ can be derived in a similar way, and are thus omitted. □

Approximation of $\hat{\rho}_i$ Due to the analytical intractability of $\Delta_i(t)$, we are interested in proper approximations of $\hat{\rho}_i$ that can lead to good estimate of Λ; a good estimate in the context of stability study means a tight underestimation. Recall that $\hat{\rho}_i \le \rho_i$ and equality holds if and only if $\rho_i = 1$ or $\rho_i = 0$; therefore by replacing $\hat{\rho}_i$ by ρ_i in Σ(c), solutions to the resulting system of equations form an underestimation of Λ but accurate when $\rho_i = 1$ or 0 for all i. Moreover, when $0 < \hat{\rho}_i < 1$, we have

$$\hat{\rho}_i = \lim_{t\to\infty} \frac{\frac{T_{i,Q}(t)-\Delta_i(t)}{S^{\text{av}}_{i,Q}(t)}}{\frac{T_{i,Q}(t)-\Delta_i(t)}{S^{\text{av}}_{i,Q}(t)} + \frac{T_{i,\overline{Q}}(t)+\Delta_i(t)}{S^{\text{av}}_{i,\overline{Q}}(t)}}$$
$$\le \lim_{t\to\infty} \frac{\frac{T_{i,Q}(t)}{T_{i,Q}(t)+T_{i,\overline{Q}}(t)} S^{\text{av}}_{i,\overline{Q}}(t)}{\frac{T_{i,Q}(t)}{T_{i,Q}(t)+T_{i,\overline{Q}}(t)} S^{\text{av}}_{i,\overline{Q}}(t) + \frac{T_{i,\overline{Q}}(\omega,t)}{T_{i,Q}(t)+T_{i,\overline{Q}}(t)} S^{\text{av}}_{i,Q}(t)} = \frac{\rho_i \mathbb{E}S_{i,\overline{Q}}}{\rho_i \mathbb{E}S_{i,\overline{Q}} + (1-\rho_i)\mathbb{E}S_{i,Q}} \le \rho_i,$$

where

$$S^{\text{av}}_{i,Q}(t) = \frac{1}{N_{i,Q}(t)} \sum_{k=1}^{N_{i,Q}(t)} S_{i,Q}(k)$$

and defining

$$\hat{\hat{\rho}}_i = \frac{\rho_i \mathbb{E}S_{i,\overline{Q}}}{\rho_i \mathbb{E}S_{i,\overline{Q}} + (1-\rho_i)\mathbb{E}S_{i,Q}},$$

we have $\hat{\rho}_i \le \hat{\hat{\rho}}_i \le \rho_i$. Hence, substituting $\hat{\rho}_i$ with $\hat{\hat{\rho}}_i$ in Σ(c), we can obtain a tighter underestimation of Λ than with ρ_i, thus trading off computational complexity for higher accuracy. Empirical results suggest that $\hat{\hat{\rho}}$ is sufficiently close to $\hat{\rho}$, and we use $\hat{\hat{\rho}}$ as $\hat{\rho}$ throughout our computation.

Proof of Proposition V.8 Substituting $\widetilde{\Sigma}$(b) in (a), we obtain

$$\begin{aligned}
\tau_i &= \frac{2\lambda_i}{P(W+1)}\left[\frac{W-1}{2}\left(\sigma + T\sum_{j\ne i}\tau_j\right) + T\left(1+\sum_{j\ne i}\tau_j\right)\right] \\
&= \frac{2\lambda_i}{P(W+1)}\left[\frac{W+1}{2}T\sum_{j\ne i}\tau_j + \frac{W-1}{2}\sigma + T\right] \\
&= \frac{\lambda_i T}{P}\sum_{j\ne i}\tau_j + \frac{\lambda_i((W-1)\sigma+2T)}{P(W+1)},
\end{aligned}$$

which can be rewritten as

$$\tau_i = \Big(\frac{\lambda_i T}{P}\sum_j \tau_j + \frac{\lambda_i((W-1)\sigma+2T)}{P(W+1)}\Big)\Big/\Big(1+\frac{\lambda_i T}{P}\Big).$$

Therefore, let $y = \sum_j \tau_j$, $\gamma_i^1 = \frac{\lambda_i T}{P}\Big/\Big(1+\frac{\lambda_i T}{P}\Big)$ and $\gamma_i^2 = \frac{\lambda_i((W-1)\sigma+2T)}{P(W+1)}\Big/\Big(1+\frac{\lambda_i T}{P}\Big)$, and we have $\tau_i = \gamma_i^1 y + \gamma_i^2$. Then, $\widetilde{\Sigma}$ is equivalent to

$$\widetilde{\Sigma}: \begin{cases} \tau_i = \gamma_i^1 y + \gamma_i^2 & \text{(a}'\text{)} \\ y = \sum_i \left(\gamma_i^1 y + \gamma_i^2\right) & \text{(b}'\text{)} \end{cases}$$

which admits only one solution, namely

$$\tau_i = \frac{\gamma_i^1 \sum_j \gamma_j^2}{1-\sum_i \gamma_j^1} + \gamma_i^2.$$

Proof of Theorem V.19

Proof. Using $\widetilde{\Sigma}^{\boldsymbol{g}^U}$(a), we can rewrite $\widetilde{\Sigma}^{\boldsymbol{g}^U}$(b) as follows:

$$\begin{aligned}
\rho_i &= \frac{\lambda_i}{P}\sum_{k\in[K]}\left\{q^{(k)}\left[\frac{W-1}{2}\Big(\sigma + T\sum_{j\neq i}\tau_j^{(k)}\Big) + T\Big(1+\sum_{j\neq i}\tau_j^{(k)}\Big)\right]\right\} \\
&= \theta_i^1 \sum_{k\in[K]}\Big(q^{(k)}\sum_{j\neq i}\tau_j^{(k)}\Big) + \theta_i^2 \\
&= \theta_i^1 \sum_{k\in[K]} \phi_i(q^{(k)};\rho_j, j\neq i) + \theta_i^2,
\end{aligned}$$

where $\theta_i^1 = \frac{\lambda_i(W+1)T}{2P}$, $\theta_i^2 = \frac{\lambda_i(W-1)\sigma+2T}{2P}$, and $\phi_i(q^{(k)};\rho_j, j\neq i) = q^{(k)}\sum_{j\neq i}\tau_j^{(k)} = \sum_{j\neq i}\alpha_j\big[q^{(k)}\big]^2$ with $\alpha_j = \frac{2\rho_j}{W+1} > 0$ for all j. Notice that $\phi_i(q^{(k)};\rho_j, j\neq i)$ is a convex function of $q^{(k)}$ given any fixed ρ_j where $j\neq i$, and it is also an increasing function of ρ_j's given any fixed $q^{(k)}$. We then have

$$\begin{aligned}\rho_i &= \theta_i^1 \sum_{k\in[K]} \phi_i(q^{(k)}) + \theta_i^2 \\ &= \theta_i^1 \cdot K \sum_{k\in[K]} \left(\frac{1}{K}\phi_i(q^{(k)})\right) + \theta_i^2 \\ &\geq \theta_i^1 \cdot K\phi_i \left(\sum_{k\in[K]} \left(\frac{1}{K}q^{(k)}\right)\right) + \theta_i^2 \\ &= \theta_i^1 \cdot K\phi_i \left(\frac{1}{K}\right) + \theta_i^2,\end{aligned}$$

where the equality holds when $q_i^{(k)} = \frac{1}{K}$. Therefore, when switching to the equi-occupancy policy from any arbitrary unbiased policy, the utilization factor of each node is always non-increasing. Hence, we conclude that the equi-occupancy scheduling policy is throughput optimal in $\mathcal{G}^U$. □

Specifications of the implementation of test bench

Total bandwidth	11 Mbps
Data packet length P	1500 Bytes
DIFS	50 μs
SIFS	10 μs
ACK packet length (in time units)	203 μs
Header length (in time units)	192 μs
Empty system slot time σ	20 μs
Propagation delay δ	1 μs
Initial backoff window size W	32
Maximum backoff stage m	5
Data rate granularity $\Delta\lambda$	100 Kbps
Instability threshold constant	1%
Total simulated time T_f	10 seconds

Table C.1: Specifications of the implementation of test bench.

APPENDIX D

Glossary of Notation

Chapter II

w^t	mixed strategy of the user at time t
$\mathcal{A}_w$	action space of the user
M_w	payoff matrix of the user
$\mathcal{I}_w^t$	information state of the user at time t
g_w	decision policy of of the user
$u^t, \mathcal{A}_u, M_u, \mathcal{I}_u^t, g_u$	counterparts of the attacker
T	time horizon
n	number of channels, $[n] = \{1, 2, \ldots, n\}$

Chapter III

$[n]^0$	{0, 1, . . . , n}
s_t	resource level of the attacker at time t
f_t	resource dynamics at time t
$\mathcal{F}_t(s)$	attacker's feasible action set given resource level s at time t
$\mathcal{S}_t$	set of resource levels such that all actions are feasible at time t
h_t	history of the game at time t
Δ_n	space of distributions over $[n]$

Δ_n^0	space of distributions over $[n]^0$
w	decision policy of of the user
$\mathcal{G}$	space of policies as complete contingency plans for the user
$\tilde{\mathcal{G}}$	reduced space of policies depending only on resource level and time for the user
$u, \mathcal{G}^0, \tilde{\mathcal{G}}^0$	counterparts of the attacker
c_k	loss incurred by being attacked on channel k

Chapter IV

N	number of users
K	number of channels
α_k	time duration of a slot on channel k
λ	symmetric traffic arrival rate to each user
τ	symmetric transmission attempt rate of each user
x	population state
x_k	portion of the user population on channel k $x = (x_1, x_2, \ldots, x_K)$
X_N	finite grid space of population states with N users
Δ_K	simplex in $\mathbb{R}^K$
F^N	payoff function of N-user population game
F	payoff function of continuous population game
ρ	revision protocol

Chapter V

subscript $_i$	index of user
superscript $^{(k)}$	index of channel
λ_i	data bit arrival rate at user i
$Q_i(t)$	continuous-time queueing process of user i, i.e., number of packets queued by user i at time t

$\hat{Q}_i(n)$	embedded queueing process of user i, i.e., number of packets queued by user i at the beginning of slot n
W	initial backoff window size
$\overline{W}_i$	average backoff window size of user i
m	maximum number of backoff stages
τ_i	transmission attempt rate of user i
p_i	collision rate of user i
$S_{\{\cdot\}}$	time duration of a slot given event $\{\cdot\}$
ρ_i	utilization factor of the continuous-time queueing process
$\hat{\rho}_i$	utilization factor of the embedded queueing process
$\mathcal{Q}_i^n$	steady state occupancy distribution of user i w.r.t. n-slots
$q_i^{(k)}$	steady state probability that user i is in channel k at the beginning of an n-slot of user i
$\mathcal{Q}_i^c$	steady state occupancy distribution of user i w.r.t. c-slots
$\hat{q}_i^{(k)}$	steady state probability that user i is in channel k at the beginning of a c-slot of channel k
$\mathcal{Q}_i^p$	steady state packet assignment distribution of user i
$\tilde{q}_i^{(k)}$	steady state probability that a packet of user i is served in channel k

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] K. Tan, J. Zhang, J. Fang, H. Liu, Y. Ye, S. Wang, Y. Zhang, H. Wu, W. Wang, and G. Voelker, "Sora: High Performance Software Radio using General Purpose Multi-core Processors," *USENIX NSDI 2009*, 2009.

[2] Y. Li, J. Fang, K. Tan, J. Zhang, Q. Cui, and X. Tao, "Soft-LTE: A Software Radio Implementation of 3GPP Long Term Evolution Based on Sora Platform," *Demo in ACM MobiCom 2009*, 2009.

[3] F. K. Jondral, "Software-Defined Radio: Basics and Evolution to Cognitive Radio," *EURASIP Journal Wireless Communications and Networking*, vol. 2005, pp. 275–283, August 2005.

[4] D. Matula, "A Periodic Optimal Search," *The American Mathematical Monthly*, vol. 71, no. 1, pp. 15–21, 1964.

[5] W. Black, "Discrete Sequential Search," *Information and Control*, vol. 8, pp. 159–162, 1965.

[6] J. Milton C. Chew, "A Sequential Search Procedure," *The Annals of Mathematical Statistics*, vol. 38, no. 2, pp. 494–502, 1967.

[7] R. Ahlswede and I. Wegener, *Search Problems*. John Wiley & Sons, 1987.

[8] D. Assaf and S. Zamir, "Optimal Sequential Search: A Bayesian Approach," *The Annals of Statistics*, vol. 13, no. 3, pp. 1213–1221, 1985.

[9] F. Kelly, "On Optimal Search with Unknown Detection Probabilities," *Journal of Mathematical Analysis and Applications*, vol. 88, no. 2, pp. 422–432, 1982.

[10] S. M. Pollock, "A Simple Model of Search for a Moving Target," *Operations Research*, vol. 18, no. 5, pp. 883–903, 1970.

[11] R. R. Weber, "Optimal Search for a Randomly Moving Object," *Journal of Applied Probability*, vol. 23, no. 3, pp. 708–717, 1986.

[12] W. Xu, K. Ma, W. Trappe, and Y. Zhang, "Jamming Sensor Networks: Attack and Defense Strategies," *Network, IEEE*, vol. 20, pp. 41–47, may-june 2006.

[13] A. Wood, J. Stankovic, and G. Zhou, “DEEJAM: Defeating Energy-Efficient Jamming in IEEE 802.15.4-based Wireless Networks,” in *SECON ’07*, pp. 60 –69, 2007.

[14] G. Noubir and G. Lin, “Low-power DoS Attacks in Data Wireless LANs and Countermeasures,” *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 7, no. 3, pp. 29–30, 2003.

[15] E. Kehdi and B. Li, “Null Keys: Limiting Malicious Attacks Via Null Space Properties of Network Coding,” in *INFOCOM ’09*, pp. 1224 –1232, april 2009.

[16] J. Chiang and Y.-C. Hu, “Cross-Layer Jamming Detection and Mitigation in Wireless Broadcast Networks,” *Networking, IEEE/ACM Transactions on*, vol. 19, no. 1, pp. 286–298, 2011.

[17] C. Popper, M. Strasser, and S. Capkun, “Anti-jamming Broadcast Communication Using Uncoordinated Spread Spectrum Techniques,” *Selected Areas in Communications, IEEE Journal on*, vol. 28, no. 5, pp. 703–715, 2010.

[18] G. Noubir, R. Rajaraman, B. Sheng, and B. Thapa, “On the Robustness of IEEE 802.11 Rate Adaptation Algorithms Against Smart Jamming,” in *WiSec ’11*, WiSec ’11, (New York, NY, USA), pp. 97–108, ACM, 2011.

[19] A. Sampath, H. Dai, H. Zheng, and B. Zhao, “Multi-channel Jamming Attacks using Cognitive Radios,” in *ICCCN ’07*, pp. 352–357, 2007.

[20] R. Negi and S. Goel, “Secret Communication Using Artificial Noise,” in *Vehicular Technology Conference*, vol. 3, pp. 1906–1910, 2005.

[21] L. Dong, Z. Han, A. Petropulu, and H. Poor, “Cooperative Jamming for Wireless Physical Layer Security,” in *SSP ’09*, pp. 417 –420, 31 2009-sept. 3 2009.

[22] S. Gollakota and D. Katabi, “Physical Layer Wireless Security Made Fast and Channel Independent,” in *INFOCOM ’11*, pp. 1125–1133, 2011.

[23] R. Isaacs, *Differential Games*. Wiley, 1965.

[24] J. D. Grote, ed., *The Theory and Application of Differential Games*. D. Reidel Publishing Company, 1975.

[25] Y. Yavin and M. Pachter, eds., *Pursuit-Evasion Differential Games*. Pergamon Press, 1987.

[26] T. Başar and G. Olsder, *Dynamic Noncooperative Game Theory*. Society for Industrial and Applied Mathematics, 2nd edition ed., 1998.

[27] E. Altman, K. Avrachenkov, and A. Garnaev, “A Jamming Game in Wireless Networks with Transmission Cost,” in *Network Control and Optimization*, Springer Berlin Heidelberg, 2007.

[28] Y. Sagduyu and A. Ephremides, "A Game-Theoretic Analysis of Denial of Service Attacks in Wireless Random Access," in *WiOpt '07*, pp. 1 –10, april 2007.

[29] S. Bhattacharya and T. Başar, "Game-theoretic Analysis of an Aerial Jamming Attack on a UAV Communication Network," in *ACC '10*, pp. 818–823, 2010.

[30] V. Navda, A. Bohra, S. Ganguly, and D. Rubenstein, "Using Channel Hopping to Increase 802.11 Resilience to Jamming Attacks," in *INFOCOM '07, Mini-Conference*, pp. 2526–2530, 2007.

[31] K. Pelechrinis, C. Koufogiannakis, and S. Krishnamurthy, "On the Efficacy of Frequency Hopping in Coping with Jamming Attacks in 802.11 Networks," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 10, pp. 3258 –3271, 2010.

[32] H. Li and Z. Han, "Dogfight in Spectrum: Combating Primary User Emulation Attacks in Cognitive Radio Systems, Part I: Known Channel Statistics," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 11, pp. 3566–3577, 2010.

[33] N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games.* New York, NY, USA: Cambridge University Press, 2006.

[34] N. Littlestone and M. K. Warmuth, "The Weighted Majority Algorithm," *Information and Computation*, vol. 108, no. 2, pp. 212–261, 1994.

[35] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119 – 139, 1997.

[36] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. Schapire, "Gambling in a Rigged Casino: The Adversarial Multi-armed Bandit Problem," in *Foundations of Computer Science, 1995. Proceedings., 36th Annual Symposium on*, pp. 322–331, 1995.

[37] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire, "The Nonstochastic Multiarmed Bandit Problem," *SIAM J. Comput.*, vol. 32, no. 1, pp. 48–77, 2003.

[38] S. Arora, E. Hazan, and S. Kale, "The Multiplicative Weights Update Method: a Meta-Algorithm and Applications," *Theory of Computing*, vol. 8, no. 6, pp. 121–164, 2012.

[39] H. P. Young, *Strategic Learning and its Limits.* Oxford University Press, 2004.

[40] A. Blum and Y. Mansour, "From External to Internal Regret," *The Journal of Machine Learning Research*, vol. 8, pp. 1307–1324, 2007.

[41] C. Daskalakis, R. Frongillo, C. Papadimitriou, G. Pierrakos, and G. Valiant, "On Learning Algorithms for Nash Equilibria," in *Algorithmic Game Theory* (S. Kontogiannis, E. Koutsoupias, and P. Spirakis, eds.), vol. 6386 of *Lecture Notes in Computer Science*, pp. 114–125, Springer Berlin Heidelberg, 2010.

[42] R. Kleinberg, G. Piliouras, and E. Tardos, “Multiplicative Updates Outperform Generic No-regret Learning in Congestion Games: Extended Abstract,” in *STOC ’09*, 2009.

[43] G. Kasbekar and A. Proutiere, “Opportunistic Medium Access in Multi-channel Wireless Systems: A Learning Approach,” in *Allerton ’10*, pp. 1288–1294, 2010.

[44] *IEEE Standard 802.11, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.* IEEE Standards Association, 2012.

[45] G. Bianchi, “Performance Analysis of the IEEE 802.11 Distributed Coordination Function,” *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 535–547, 2000.

[46] A. Kumar, E. Altman, D. Miorandi, and M. Goyal, “New Insights from a Fixed Point Analysis of Single Cell IEEE 802.11 WLANs,” in *Proceedings of IEEE INFOCOM*, 2005.

[47] G. R. Cantieni, Q. Ni, C. Barakat, and T. Turletti, “Performance Analysis under Finite Load and Improvements for Multirate 802.11,” *Elsivier Computer Communications*, vol. 28(10), pp. 1095–1109, 2005.

[48] D. Malone, K. Duffy, and D. J. Leith, “Modeling the 802.11 Distributed Coordination Function in Non-saturated Heterogeneous Conditions,” *IEEE/ACM Transactions on Networking*, vol. 15(1), pp. 159–172, 2007.

[49] A. Jindal and K. Psounis, “The Achievable Rate Region of 802.11-Scheduled Multi-hop Networks,” *IEEE/ACM Transactions on Networking*, vol. 17(4), pp. 1118–1131, 2009.

[50] D. Leith, V. Subramanian, and K. Duffy, “Log-convexity of Rate Region in 802.11e WLANs,” *IEEE Communications Letters*, vol. 14(1), pp. 57–59, 2010.

[51] A. Raniwala and T. Chiueh, “Architecture and Algorithms for an IEEE 802.11-based Multi-Channel Wireless Mesh Network,” in *Proceedings of IEEE INFOCOM*, 2005.

[52] A. Mohsenian-Rad and V. Wong, “Distributed Multi-Interface Multichannel Random Access Using Convex Optimization,” *Mobile Computing, IEEE Transactions on*, vol. 10, pp. 67–80, Jan. 2011.

[53] V. Anantharam, “The Stability Region of the Finite-User Slotted ALOHA Protocol,” *IEEE Transactions on Information Theory*, vol. 37, pp. 535–540, 1991.

[54] Q. Wang and M. Liu, “Learning in Hide-and-Seek,” in *IEEE INFOCOM ’14*, 2014.

[55] Q. Wang and M. Liu, “Joint Control of Transmission Power and Channel Switching against Adaptive Jamming,” in *Allerton ’13*, 2013.

[56] Q. Wang and M. Liu, "Jamming Defense Against a Resource-Replenishing Adversary in Multi-channel Wireless Systems," in *IEEE WiOpt '14*, 2014.

[57] Q. Wang and M. Liu, "Throughput Optimal Switching in Multichannel WLANs," *IEEE Transactions on Mobile Computing*, vol. 12, pp. 2470–2482, Dec. 2013.

[58] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, eds., *Algorithmic Game Theory*. Cambridge University Press, 2007.

[59] T. Uchiya, A. Nakamura, and M. Kudo, "Algorithms for Adversarial Bandit Problems with Multiple Plays," in *Proceedings of the 21st international conference on Algorithmic learning theory*, pp. 375–389, Springer-Verlag, 2010.

[60] R. Selten, "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," 1975.

[61] J. D. Abernethy and M. K. Warmuth, "Repeated Games against Budgeted Adversaries," in *NIPS '10*, 2010.

[62] A. Wilansky, "The Row-Sums of the Inverse Matrix," *The American Mathematical Monthly*, vol. 58, no. 9, pp. 614–615, 1951.

[63] N. Abramson, "THE ALOHA SYSTEM: Another Alternative for Computer Communications," in *Proceedings of the November 17-19, 1970, Fall Joint Computer Conference*, AFIPS '70 (Fall), (New York, NY, USA), pp. 281–285, ACM, 1970.

[64] L. G. Roberts, "ALOHA Packet System with and Without Slots and Capture," *SIGCOMM Comput. Commun. Rev.*, vol. 5, pp. 28–42, Apr. 1975.

[65] D. Bertsekas and R. Gallager, *Data Networks (2nd edition)* . Prentice Hall, 1992.

[66] T. Bonald, S. Borst, N. Hegde, and A. Proutiére, "Wireless data performance in multi-cell scenarios," in *Proceedings of the Joint International Conference on Measurement and Modeling of Computer Systems*, SIGMETRICS '04/Performance '04, (New York, NY, USA), pp. 378–380, ACM, 2004.

[67] W. H. Sandholm, *Population Games and Evolutionary Dynamics*. The MIT Press, 2010.

[68] G. Roth and W. Sandholm, "Stochastic Approximations with Constant Step Size and Differential Inclusions," *SIAM Journal on Control and Optimization*, vol. 51, no. 1, pp. 525–555, 2013.

[69] K. R. Duffy, "Mean Field Markov Models of Wireless Local Area Networks," *Markov Processes and Related Fields*, vol. 16(2), pp. 295–328, 2010.

[70] E. Felemban and E. Ekici, "Single Hop IEEE 802.11 DCF Analysis Revisited: Accurate Modeling of Channel Access Delay and Throughput for Saturated and Unsaturated Traffic Cases," *IEEE Transactions on Wireless Communications*, vol. 10, no. 10, pp. 3256–3266, 2011.

[71] Q. Wang and M. Liu, "Throughput Optimal Switching in Multichannel WLANs," http://arxiv.org/pdf/1201.6065v1.pdf.

[72] J. L. Massey and P. Mathys, "The Collision Channel Without Feedback," *IEEE Transactions on Information Theory*, vol. 31, pp. 192–204, 1985.