Jamming Defense Against a Resource-Replenishing Adversary in Multi-channel Wireless Systems

Qingsi Wang, Shang-Pin Sheng, Jacob Abernethy, Mingyan Liu
University of Michigan, Ann Arbor

***Abstract*—We revisit the jamming defense problem in a multi-channel wireless system, using a general formulation of online learning against an adversary via repeated game-playing. We provide the explicit form of the worst-case optimal channel-hopping strategy of a legitimate user in a multi-stage interaction with a resource-replenishing jamming attacker. Interestingly, we show that the worst imaginary enemy can be given as an adversary who behaves in an i.i.d. manner in this multi-stage interaction, and the optimal strategy of the user 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.**

I. Introduction

In this paper, we revisit the extensively studied jamming defense problem, using a general formulation of online learning against an adversary via repeated game-playing. In particular, we consider the channel-hopping decisions of a legitimate user evading a jamming attacker in a multi-channel system, where the jammer (interchangeably attacker or adversary) 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, and we provide the explicit characterization. This 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.

A big part of the literature on jamming focuses on specific attack and defense mechanisms, see e.g., [1], [2] for a collection of jamming attacks and anti-jamming measures. Examples also include using stronger error detection, correction, and spreading codes at the physical layer [3], [4], [5], [6], exploring the vulnerability in the rate adaptation mechanism of IEEE 802.11 [7], and multi-channel jamming using a single cognitive ratio [8]. Interestingly, jamming can also be used by legitimate users to achieve physical layer security in the presence of an eavesdropper, see e.g., [9], [10], [11].

The interaction between a jammer and a user/defender is often modeled as a strategic game, and this 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 [12], a random access game [13], a differential game between a mobile jammer and mobile users [14], a Stackelberg game [15] and a zero-sum game [16]. 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. [17].

In this work, we directly analyze a multi-stage jamming defense problem, and our formulation is applicable in a general learning context. There has been extensive research in optimal decision in an adversarial environment in the learning theory community, e.g. [18], [19], [20] to name a few. Of these the one closest to our work is [20], where the minimax optimal strategy is constructively obtained against a budget-constrained adversary; this can be reduced to a special case in our formulation as shown in the later sections. Unlike the model in [20], where the game terminates whenever the adversary exhausts its budget, the interaction between the user and the adversary under our formulation continues over an arbitrary horizon with resource replenishment. Moreover, the action space of the adversary can be time-varying given different resource levels over time, which makes the reasoning process of the user potentially more challenging. In summary, our contribution is twofold:

- We present a general online learning framework against a resource-replenishing adversary, which is applicable to jamming defense as well as other competitive game problems.
- We explicitly characterize the minimax optimal strategy for the legitimate user in repeated interactions with the jammer. Interestingly, we show that the worst adversary can be one who behaves in an i.i.d. fashion in this multi-stage interaction, and the optimal strategy of the user is determined by the induced random walk of the adversarial behavior.

Explicit characterization of jamming defense strategies from a dynamic game perspective can also be found in some existing literature, with distinct system models and behavioral constraints as compared to this work (see e.g., [21], [22]). In this regard, our paper can be viewed as a progression in this line of work. The remainder of the paper is organized as follows. Section II formulates the problem, followed by the main results and analysis in Section III. Section IV discusses extension and open problems based on this work, and Section V concludes the paper.

II. Problem Formulation and Preliminaries

We consider a sequential decision problem of a user against a jamming attacker (which will be mainly called the adversary

978-3-901882-63-0/2014 - Copyright is with IFIP

in the rest of this section) in a multi-channel system. We use $[n] := \{1, 2, \ldots, n\}$ to denote the action space of the user (i.e., the indexed set of channels), and $[n]^0 := \{0, 1, \ldots, n\}$ the action space of the adversary, where 0 is the null action (i.e., initiating no attack). Assume that the adversary has a finite amount of resource s_t at time (or round, used interchangeably) t, and any non-null action by the adversary consumes a certain amount of resource; it also obtains replenishment after a round. In particular, given s_t and j_t which denotes the action taken by the adversary at t, the resource of the adversary 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 adversary 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 adversary. Let $\hat{M} \in \mathbb{R}_+^{n \times n}$ be a *loss matrix* and $M = \begin{bmatrix} \mathbf{0} & \hat{M} \end{bmatrix}$ an *augmented loss matrix*, where $\mathbf{0}$ is the zero column vector of length n. These matrices are explained shortly. We denote by Δ_n and Δ_n^0 the spaces of probability distributions on $[n]$ and $[n]^0$. Given any vector $v = (v_1, v_2, \ldots, v_m)$, 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 $[n]^0$. The interaction between the user and the attacker is then given as follows.

Initialization: The adversary has a finite amount of initial resource s_1. T is a finite time horizon.

For $t = 1, 2, \ldots, T$**:**

1) The user chooses a distribution $w^t \in \Delta_n$, and an action $i_t \in [n]$ is realized per w^t independently in this round.
2) The adversary 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 in this round. After consumption and replenishment, its available resource for the next round is given by $s_{t+1} = f_t(s_t, j_t)$.
3) The user observes j_t and suffers a loss given by $M_{i_t j_t}$, which is the (i_t, j_t)-th entry of M.

Throughout this paper, we make the following assumption.

Assumption 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 adversary and the resource dynamics f_t for $t = 1, 2, \ldots, T$.

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 $\mathcal{H}$. Then, the user's strategy can be given by a mapping $w : \mathcal{H} \to \Delta_n$, and we denote the space of all such mappings as $\mathcal{A}$. We adopt the following notion of a (strong) adversary. 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 $\mathcal{A}^0$. The user's objective is to minimize the worst-case loss:

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

Note that we could also consider a weaker adversary, 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 adversary can be as capable as a strong adversary in the context of the above decision problem for the user.

While we have not explicitly stated whether the adversary is strategic, the minimax formulation means that the user shall treat the adversary 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 $\mathcal{A}$ and $\mathcal{A}^0$ with the payoff functions W, U. Then, a Nash equilibrium (NE) strategy for player 1 (the user) in G is exactly a minimaximizer of (1). The above game theoretical interpretation of (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 [23], which is also a NE of G. Hence, the minimaximizer of (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 (1); in particular w^* is a minimaximizer to (1) and u^* a corresponding maximizer given w^*. Also, note that the pair (u^*, w^*) is a solution to the problem

$$\max_{u \in \mathcal{A}^0} \min_{w \in \mathcal{A}} \mathbb{E}\left\{ \sum_{t=1}^{T} w(h_t)^\top M u(h_t) \right\}. \tag{2}$$

For technical reasons, we will consider a slightly perturbed version of problem (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 $\epsilon(s) \leq \epsilon_{\max}$ for all $s \in \mathbb{R}_+$, where $\epsilon_{\max}$ is a predetermined constant. The perturbed problem is then given by

$$\min_{w \in \mathcal{A}} \max_{u \in \mathcal{A}^0} \mathbb{E}\left\{ \sum_{t=1}^{T} w(h_t)^\top M u(h_t) + \epsilon(s_T) \right\}. \tag{3}$$

With a similar argument used for (1), we can show an minimax-optimal solution exists for (3), which coincides with

the SPE of the extensive game induced by (3). For the perturbed problem, we will inherit all the notation from (1), e.g., w^* is an optimal solution to (3). We note that if (w^*, u^*) is a solution to (3), the resulting loss in (1) is at most $\epsilon_{\max}$ more than the optimal minimax loss, and a similar result holds for (2), as shown in the following lemma. Due to the limit of space, proofs for the preliminary results in this section are omitted.

Lemma 1: Let

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

and

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

If $\hat{w}^*$ and $\hat{u}^*$ are respective optimal solutions to (1) and (2), then $\ell(w^*) \le \ell(\hat{w}^*) + \epsilon_{\max}$ and $g(u^*) \ge g(\hat{u}^*) - \epsilon_{\max}$.

We proceed with the following assumptions.

Assumption 2:

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 $s \ge 0$.

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

Lemma 2:

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

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 $w \in \tilde{\mathcal{A}}$, and denote by $(w^*, u^*) \in \tilde{\mathcal{A}} \times \tilde{\mathcal{A}}^0$ an optimal solution to (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 adversary using u^* provided w^* as

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

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

Lemma 3: $U^*_t(s_t)$ is strictly increasing in s_t for all t.

With the above preliminary results, we proceed in the next section showing our main optimality results.

III. Optimality Results: Diagonal $\hat{M}$

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

Theorem 1:

1) The optimal strategy of the user in problem (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 certain conditions on the resource dynamics $f_t, t = 1, 2, \ldots, T$, the optimal strategy of the user in problem (1) is to optimally respond to an adversary, 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 theorem 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.

A. Basic characterization

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

Lemma 4: Any SPE strategy $u^* \in \tilde{\mathcal{A}}^0$ for the adversary is such that either $u^*_0(s_t, t) = 1$ or $\mathrm{supp}(u^*(s_t, t)) \supseteq [n]$.

Proof: Let $(w^*, u^*) \in \tilde{\mathcal{A}} \times \tilde{\mathcal{A}}^0$ be a pair of SPE strategies. Assume that $u^*_0(s_t, t) < 1$, and let $\mathcal{N} := [n] - \mathrm{supp}(u^*(s_t, t))$. If $\mathcal{N} \neq \emptyset$, then $\mathrm{supp}(w^*(s_t, t)) \subseteq \mathcal{N}$, i.e., $\mathrm{supp}(w^*(s_t, t)) \cap \mathrm{supp}(u^*(s_t, t)) = \emptyset$. Otherwise, the payoff of any subgame (s_t, t) for the user using w^* provided u^*, which is given by

$$W^*_t(s_t) := \mathbb{E}\left\{-\sum_{\tau=t}^{T} w^*(s_\tau, \tau)^\top M u^*(s_\tau, \tau) - \epsilon(s_T) \;\middle|\; s_t\right\}$$
$$= \sum_{i=1}^{n} w^*_i(s_t, t)(-u^*_i(s_t, t)c_i + \sum_{j=0}^{n} u^*_j(s_t, t) W^*_{t+1}(f_t(s_t, j)))$$
$$= -\sum_{i=1}^{n} w^*_i(s_t, t) u^*_i(s_t, t) c_i + \sum_{j=0}^{n} u^*_j(s_t, t) W^*_{t+1}(f_t(s_t, j)), \tag{4}$$

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

Then, we have

$$U^*_t(s_t) = u^*_0(s_t, t) U^*_{t+1}(f_t(s_t, 0)) +$$
$$+ \sum_{i=1}^{n} u^*_i(s_t, t)(w^*_i(s_t, t)c_i + U^*_{t+1}(f_t(s_t, i))) \tag{5}$$

$$= \sum_{i \in \mathrm{supp}(u^*(s_t,t))} u_i^*(s_t,t) U_{t+1}^*(f_t(s_t,i)) < U_{t+1}^*(f_t(s_t,0)),$$

where the last inequality is due to Assumption 2 and Lemma 3. Hence, the payoff of the adversary can be strictly improved by choosing the null action with probability one, which contradicts the fact that u^* is a SPE strategy. Therefore, $\mathrm{supp}(u^*(s_t)) \supseteq [n]$. ∎

Lemma 5: For a pair of SPE strategies (w^*, u^*), if $\mathrm{supp}(u^*(s_t,t)) \supseteq [n]$, then $\mathrm{supp}(w^*(s_t,t)) = [n]$.

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 $i_1 \notin \mathrm{supp}(w^*(s_t,t))$. 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 adversary can strictly improve its payoff of any subgame (s_t,t), thus resulting in a contradiction. ∎

Lemma 6: 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 $\mathrm{supp}(u^*(s_t,t)) \supseteq [n]$ by Lemma 4, and thus $\mathrm{supp}(w^*(s_t,t)) = [n]$ by Lemma 5. Hence, referring to (4) 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 (5), 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 7: 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 $\mathcal{A}^\dagger$.

Proof: Assume that $0 < u_0^*(s_t,t) < 1$ for some $s_t \in \mathcal{S}_t$ and t. Then, by Lemma 6 we have $\mathrm{supp}(u^*) \supseteq [n]$ 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 adversary 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 (4), 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 adversary'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 Theorem 1, which we reproduce in the following theorem.

Theorem 2: The optimal strategy of the user for the problem (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 when $s_t \in \mathcal{S}_t$, and (b) takes the null action with probability one when $s_t \notin \mathcal{S}_t$.

In some application instances, we can reason that the adversary 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.

Example. Assume that $f_t(s,i) < s$ for all $i \in [n]$, and $f_t(s,0) = s$. Let $\delta_{\min} := \inf_{s \geq 0, i \in [n]}(s - f_t(s,i))$, and assume that $\delta_{\min} > 0$. Let $T > \frac{s_1}{\delta_{\min}}$. 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 adversary 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 adversary'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 1 we have $g(u^*) \geq g(\hat{u}^*) - \epsilon_{\max}$ for any $\epsilon_{\max} > 0$, where $g(\hat{u}^*)$ is the optimal value of (2). Hence, u^* is an optimal solution to (2), and an optimal strategy of the user in (1) is to optimally respond to this belief on the adversarial behavior. In particular, it has the structure shown in Lemma 6 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 adversary at each round $\tau \geq t$ is 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

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

Then,

$$\begin{aligned}
U_t^*(s_t) &= \mathbb{E}\left\{ \sum_{\tau=t}^T w_{i_\tau}^*(s_\tau,\tau)c_{i_\tau} \;\middle|\; s_t \right\} \\
&= \frac{1}{\sum_{j=1} 1/c_j} \mathbb{E}\left\{ \sum_{\tau=t}^T \mathbf{1}(s_\tau \in \mathcal{S}_\tau) \;\middle|\; s_t \right\}
\end{aligned}$$

$$= \frac{1}{\sum_{j=1} 1/c_j} \mathbb{E}T(s_t),$$

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 [20].

B. Characterization with structure on the replenishment

The difficulty of applying Theorem 2 is that we have to determine whether the adversary chooses the null action with probability one even when all non-null actions are feasible. Intuitively, the only incentive for the adversary 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 8.

Assumption 3:

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 8: If (w^*, u^*) is a pair of SPE strategies and $u^* \in \mathcal{A}^\dagger$, 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 Theorem 1, which is repeated in the following theorem.

Theorem 3: The user's optimal strategy in (1) is to optimally respond to an adversary, 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 $s_t \notin \mathcal{S}_t$.

Proof: Lemma 8 directly proves the above claim for the perturbed problem (3). Using the same argument as shown in the example after Theorem 2, we conclude that the described adversarial strategy is also an equilibrium strategy in the original problem (1), and the result follows. ∎

The optimal strategy of the user is then given as in Lemma 6, where U^* can be similarly estimated using Monte-Carlo method as in [20].

C. Asymptotics

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

$$\kappa := \limsup_{T \to \infty} \min_{w \in \mathcal{A}} \max_{u \in \mathcal{A}^0} \mathbb{E} \left\{ \frac{1}{T} \sum_{t=1}^T w(h_t)^\top M u(h_t) \right\}. \quad (6)$$

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 $0 = d_0 \le d_1 \le \ldots \le d_n$. We also assume that $\gamma \ge d_1$. Hence, f satisfies Assumption 2 and 3. Let $s_{\text{th}} = \min\{d_n - \gamma, 0\}$, and then $\mathcal{S}_t = [s_{\text{th}}, \infty)$. Consequently, using Theorem 3, we can regard the attacker as behaving randomly and taking action from $[n]$ with the probability distribution q whenever $s \ge s_{\text{th}}$, 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 $X_t := \mathbf{1}(S_t \ge s_{\text{th}})$. 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

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

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.

IV. Extension and Open Problems

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 this work.

A. 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 9 ([24]): 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 $\hat{D} = \mathrm{diag}^{-1}(\mathbf{1}^\top \hat{M}^{-1}) = zI$, 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.

B. Conversion to a gain formulation

In this work, 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.

V. Concluding Remark

We presented the optimal strategy of the user to suffer the least worst-case cost from jamming attacks. The diagonal (i.e. binary collision) cost structure played a pivotal role in our techniques, and meanwhile we noticed they may not be directly applicable for an arbitrary interference model. A possibly different rationale in reasoning would be necessary to accommodate this further generalization as well as the promising notion of adversarial channel capacity, as shown in our previous discussion. Instead of considering the worst-case optimality, using the typical regret measure and investigating no-regret strategies is also an interesting direction of future research.

References

[1] W. Xu, W. Trappe, Y. Zhang, and T. Wood, "The Feasibility of Launching and Detecting Jamming Attacks in Wireless Networks," in *MobiHoc '05*, pp. 46–57, 2005.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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*, pp. 97–108, 2011.

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

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

[10] 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.

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

[12] 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.

[13] 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.

[14] 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.
[15] 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.
[16] 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.
[17] 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.
[18] N. Littlestone and M. K. Warmuth, "The Weighted Majority Algorithm," *Information and Computation*, vol. 108, no. 2, pp. 212–261, 1994.
[19] 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.
[20] J. D. Abernethy and M. K. Warmuth, "Repeated Games against Budgeted Adversaries," in *NIPS '10*, 2010.
[21] R. Mallik, R. Scholtz, and G. Papavassilopoulos, "Analysis of An On-off Jamming Situation as a Dynamic Game," *IEEE Transactions on Communications*, vol. 48, pp. 1360–1373, Aug 2000.
[22] A. Garnaev, Y. Hayel, E. Altman, and K. Avrachenkov, "Jamming Game in a Dynamic Slotted ALOHA Network," in *Game Theory for Networks* (R. Jain and R. Kannan, eds.), Springer Berlin Heidelberg, 2012.
[23] R. Selten, "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," 1975.
[24] A. Wilansky, "The Row-Sums of the Inverse Matrix," *The American Mathematical Monthly*, vol. 58, no. 9, pp. 614–615, 1951.

APPENDIX

Proof of Lemma 8: 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 $\sigma \in \mathcal{S}_\tau$, $t \le \tau < T$, $u_0^*(\sigma,\tau) = 1$ 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 adversary. If $\tau = T-1$, then

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

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

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

Let $k \in [n]$ be such that $f_{T-1}(\sigma,k) \in \mathcal{S}_T$. Then $\tilde{U}_{T-1}(\sigma) \ge v + q_k v \ge v + q_{\min} v$. Hence $\tilde{U}_{T-1}(\sigma) > U_{T-1}^*(\sigma)$, which contradicts the fact that u^* is a SPE strategy. Thus, $\tau < T-1$. Now consider a particular subgame with the full label h_τ^σ such that $s_\tau = \sigma$. We will alternate u^* and construct inductively a sequence of strategies that only differ from u^* within this subgame. These alternative strategies will be in $\mathcal{A}^0 - \tilde{\mathcal{A}}^0$, i.e., it can depend on the past actions instead of only the resource level, and we will show that the last strategy of this sequence strictly improves the payoff of the adversary. 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 adversary at the node h_t of the game tree. To simplify our notation, since $w^* \in \mathcal{A}^\dagger \subseteq \tilde{\mathcal{A}}$, we will keep write $w^*(s,t)$ as the strategy of the user at some node h_t such that $s_t = s$. Note that

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

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

$$U_\tau^*(h_\tau^\sigma) = v + \sum_{i=1}^n w_i^*(\sigma,\tau) \sum_{j=1}^n q_j V_{\tau+2}(f_{\tau+1}(f_\tau(\sigma,0),j)).$$

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 $u_i^{(1)}(h_\tau^\sigma) = q_i$ for all $i \in [n]$ and $u_0^{(1)}(\langle h_\tau^\sigma, i, j, f_\tau(\sigma,j)\rangle) = 1$ for all $i,j \in [n]$. Then,

$$\begin{aligned} U_\tau^{(1)}(h_\tau^\sigma) &:= U_\tau(w^*,u^{(1)},h_\tau^\sigma) = \sum_{i=1}^n q_i w_i^*(\sigma,\tau) c_i + \\ &\quad + \sum_{i=1}^n \sum_{j=1}^n w_i^*(\sigma,\tau) q_j U_{\tau+1}^{(1)}\Big(\langle h_\tau^\sigma, i, j, f_\tau(\sigma,j)\rangle\Big) \\ &= v + \sum_{i=1}^n w_i^*(\sigma,\tau) \sum_{j=1}^n q_j U_{\tau+1}^{(1)}\Big(\langle h_\tau^\sigma, i, j, f_\tau(\sigma,j)\rangle\Big) \qquad (9) \\ &= v + \sum_{i=1}^n w_i^*(\sigma,\tau) \sum_{j=1}^n q_j \sum_{l=1}^n w_l^*(f_\tau(\sigma,j),\tau+1) \cdot \\ &\quad \cdot U_{\tau+2}^{(1)}\Big(\langle h_\tau^\sigma, i, j, f_\tau(\sigma,j), l, 0, f_{\tau+1}(f_\tau(\sigma,j),0)\rangle\Big) \\ &= v + \sum_{i=1}^n w_i^*(\sigma,\tau) \sum_{j=1}^n q_j \sum_{l=1}^n w_l^*(f_\tau(\sigma,j),\tau+1) \cdot \\ &\quad \cdot U_{\tau+2}^{(1)}\Big(\langle h_\tau^\sigma, i, j, f_\tau(\sigma,j), l, 0, f_{\tau+1}(f_\tau(\sigma,0),j)\rangle\Big) \end{aligned}$$

$U_{\tau+2}^{(1)}\big(\langle h_\tau^\sigma, i, j, f_\tau(\sigma,j), l, 0, f_{\tau+1}(f_\tau(\sigma,0),j)\rangle\big)$ 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 $u^{(1)} \in \tilde{\mathcal{A}}_{\tau+2}$ and $u^{(1)} = u^*$ at any node h_t with $t \ge \tau+2$ by construction, so we have

$$U_\tau^{(1)}(h_\tau^\sigma) = v + \sum_{i=1}^n w_i^*(\sigma,\tau) \sum_{j=1}^n q_j V_{\tau+2}(f_{\tau+1}(f_\tau(\sigma,0),j))$$

$$U^{(k)}_{\tau+k}(h^\sigma_{\tau+k}) = \sum_{i=1}^n w^*_i(\sigma^{(k)}, \tau+k) U^{(k)}_{\tau+k+1}\Big(\langle h^\sigma_{\tau+k}, i, 0, f_{\tau+k}(\sigma^{(k)}, 0)\rangle\Big) \tag{7}$$
$$= v + \sum_{i=1}^n w^*_i(\sigma^{(k)}, \tau+k) \cdot \sum_{l=1}^n \sum_{j=1}^n w^*_l(f_{\tau+k}(\sigma^{(k)}, 0), \tau+k+1) q_j \cdot$$
$$\cdot U^{(k)}_{\tau+k+2}\Big(\langle h^\sigma_{\tau+k}, i, 0, f_{\tau+k}(\sigma, 0), l, j, f_{\tau+k+1}(f_{\tau+k}(\sigma^{(k)}, 0), j)\rangle\Big)$$
$$= v + \sum_{i=1}^n w^*_i(\sigma^{(k)}, \tau+k) \sum_{j=1}^n q_j V_{\tau+k+2}(f_{\tau+k+1}(f_{\tau+k}(\sigma^{(k)}, 0), j)),$$

$$U^{(k+1)}_{\tau+k}(h^\sigma_{\tau+k}) = \sum_{i=1}^n q_i w^*_i(\sigma^{(k)}, \tau+k) c_i + \sum_{i=1}^n \sum_{j=1}^n w^*_i(\sigma^{(k)}, \tau+k) q_j U^{(k+1)}_{\tau+k+1}\Big(\langle h^\sigma_{\tau+k}, i, j, f_{\tau+k}(\sigma^{(k)}, j)\rangle\Big) \tag{8}$$
$$= v + \sum_{i=1}^n w^*_i(\sigma^{(k)}, \tau+k) \sum_{j=1}^n q_j \sum_{l=1}^n w^*_l(f_{\tau+k}(\sigma^{(k)}, j), \tau+k+1) \cdot$$
$$\cdot U^{(k+1)}_{\tau+k+2}\Big(\langle h^\sigma_{\tau+k}, i, j, f_{\tau+k}(\sigma^{(k)}, j), l, 0, f_{\tau+k+1}(f_{\tau+k}(\sigma^{(k)}, j), 0)\rangle\Big)$$
$$= v + \sum_{i=1}^n w^*_i(\sigma^{(k)}, \tau+k) \sum_{j=1}^n q_j \sum_{l=1}^n w^*(f_{\tau+k+1}(\sigma^{(k)}, 0), \tau+k+1) \cdot$$
$$\cdot U^{(k+1)}_{\tau+k+2}\Big(\langle h^\sigma_{\tau+k}, i, j, f_{\tau+k}(\sigma^{(k)}, j), l, 0, f_{\tau+k+1}(f_{\tau+k}(\sigma^{(k)}, 0), j)\rangle\Big)$$
$$= v + \sum_{i=1}^n w^*_i(\sigma^{(k)}, \tau+k) \sum_{j=1}^n q_j V_{\tau+k+2}(f_{\tau+k+1}(f_{\tau+k}(\sigma^{(k)}, 0), j)) = U^{(k)}_{\tau+k}(h^\sigma_{\tau+k}).$$

$$= U^*_\tau(h^\sigma_\tau),$$

i.e., $u^{(1)}$ does not change the value of the subgame labeled by h^σ_τ, and also by (9), 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 adversary 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 $u^{(r)}_i(h^\sigma_{\tau+r-1}) = q_i$ for all $i \in [n]$ and $u^{(r)}_0(\langle h^\sigma_{\tau+r-1}, i, j, f_{\tau+r-1}(\sigma^{(r-1)}, j)\rangle) = 1$ for all $i, j \in [n]$, which implies that $u^{(r)} \in \tilde{\mathcal{A}}_{\tau+r+1}$ 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 $u^{(k+1)}_i(h^\sigma_{\tau+k}) = q_i$ for all $i \in [n]$ and $u^{(k+1)}_0(\langle h^\sigma_{\tau+k}, i, j, f_{\tau+k}(\sigma^{(\tau+k)})\rangle) = 1$ for all $i, j \in [n]$. Then, $u^{(k+1)} \in \tilde{\mathcal{A}}_{\tau+k+2}$ and $u^{(k+1)} = u^{(k)}$ at all nodes h_t with $t \geq \tau+k+2$. Consequently, we have (7), 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 (8). Hence, $U^{(k)}_\tau(h^\sigma_\tau) = U^*_\tau(h^\sigma_\tau)$. Also, by (8), $U^{(k+1)}_{\tau+k+1}(h^\sigma_{\tau+k+1})$ 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 $u^{(T-1-\tau)}_0(h^\sigma_{T-1}) = 1$ 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. ■