Robust Power Management
via Learning and Game Design

Zhengyuan Zhou

Department of Electrical Engineering, Stanford University, Stanford, CA 94305, zyzhou@stanford.edu

Panayotis Mertikopoulos

Univ. Grenoble Alpes, CNRS, Grenoble INP, Inria, LIG, F-38000 Grenoble, France, panayotis.mertikopoulos@imag.fr

Aris L. Moustakas

Department of Physics, University of Athens and Institute of Accelerating Systems and Applications (IASA), Athens, Greece, arislm@phys.uoa.gr

Nicholas Bambos, Peter Glynn

Department of Management Science and Engineering, Stanford University and Department of Electrical Engineering, Stanford, CA 94305, bambos@stanford.edu, glynn@stanford.edu

We consider the target-rate power management problem for wireless networks and we propose two simple, distributed power management schemes that regulate power in a provably robust manner by efficiently leveraging past information. Both schemes are obtained via a combined approach of learning and "game design" whereby power management is reformulated as a Nash equilibrium problem for a game with suitably designed payoff functions, and the networks' users employ a no-regret learning algorithm to maximize their individual utility over time. To establish convergence, we focus on the well-known online learning algorithm called online (lazy) gradient descent (OGD) in a general class of (weighted) monotone games. In this class of games (which properly includes all concave potential games), we show that multi-agent OGD converges to the unique Nash equilibrium, even when players only have imperfect/stochastic feedback at their disposal.

In the context of power management in static (deterministic) networks, we show that the designed games are monotone when the network is feasible in the first place (i.e. when all users can concurrently attain their target rates). Otherwise, in time-varying (stochastic) networks where channel quality fluctuates ergodically over time, the designed games are monotone even when the network is only feasible on average (i.e. even though users may be unable to meet their requirements with positive probability). As a result, the proposed power management algorithms converge almost surely to the joint optimal transmission power when the network is feasible on average. This comes in stark contrast to existing algorithms (like the seminal Foschini–Miljanic algorithm and its variants) that may fail to converge altogether.

Key words: power management; online learning; Nash equilibrium; monotone games.

1. Introduction

Viewed abstractly, power management (or *power control*) is a collection of techniques that allows the users of a wireless network to achieve their performance requirements (e.g. in terms of throughput) while minimizing the power consumed by their equipment. Thus, given the key part played by transmitted power in increasing battery life and network capacity, power control has been a core aspect of network design ever since the early development stages of wireless networks.

In this general context, *distributed* power management has proven to be the predominant power management paradigm, and with good reason: centralized coordination is extremely difficult to achieve in large-scale wireless networks, a single point of failure in a centralized allocator could have devastating network-wide effects, the communication overhead alone becomes unmanageable in cellular networks, and the list goes on (Rappaport 2001, Goldsmith 2005). Consequently, considerable effort has been devoted to designing distributed power management algorithms that are provably capable of attaining various performance guarantees required by the network's users while leaving a sufficiently small footprint at the device level.

The problem of distributed power management becomes even more important and challenging in the exciting era Internet of Things (IoT) (Bradley et al. 2013), which paints the compelling vision (in part already under way) of embedding uniquely identifiable wireless devices in the world around us and connecting these devices/sensors to the existing Internet infrastructure to form an intelligent and coherently functional entity – as in "smart cities" (Deakin 2013), patient monitoring (Byrne and Lim 2007) and digital health (Au-Yeung et al. 2010). Thus, as the exponentially growing number of wireless communicating "things" has been pushing the entire wireless ecosystem from the low-traffic, low-interference regime to the high-traffic, high-interference limit, the power management question grows ever more urgent: how can the power of battery-driven (and

hence energy-constrained) devices be regulated in a *distributed, real-time* manner so as to achieve the quality-of-service guarantees using minimum power in the presence of the inherent stochastic fluctuations of the underlying wireless network?

The current gold standard for power management is the seminal method of Foschini and Miljanic (1993), which, owing to its elegance, simplicity and strong convergence properties, is still widely deployed (in one variant or another). The original Foschini–Miljanic (FM) algorithm was expanded upon in a series of subsequent works (Mitra 1994, Yates 1996, Ulukus and Yates 1998) that considered different variants of the problem (e.g. working with maximum power constraints, allowing devices to asynchronously update their power, etc.). Thereafter, various other objectives related to power management have been considered in wireless networks (as well as in the closely related wireline networks), resulting in more sophisticated models and more complex algorithms addressing issues related to throughput (El Gamal et al. 2006a,b, Seferoglu et al. 2008, Reddy et al. 2008), fairness (Eryilmaz et al. 2006), delays (Eryilmaz et al. 2008, Altman et al. 2010), backlog (Bambos et al. 2000, Gitzenis and Bambos 2002, Reddy et al. 2012, Gopalan et al. 2015) and weighted-sum-of-rates (Candogan et al. 2010, Weeraddana et al. 2012).[1]

Importantly, most of the existing (distributed) power management/control schemes tend to rely implicitly only on the previous power iterate when selecting the power for the current iteration (especially when pertaining to the target-rate power management problem that we consider here). In some respects, there is good reason to do so: the resulting power control algorithm is simple to use and implement; it does not need to use up memory to keep track of the entire radiated power history (a scarce resource in small wireless devices); and, as a pleasant by-product, the algorithm becomes much easier to analyze theoretically.

Notwithstanding, a crucial and well-known downside to this approach is that such algorithms tend to be unstable because a single power iterate is the sole power selection criterion. In particular, despite the various convergence and optimality guarantees observed when the underlying wireless network is static/slowly-varying (see e.g. Chiang et al. 2008, Tan 2015, and references therein), this

instability is acutely manifested in time-varying, stochastic networks, especially when the number of mobile devices is large and/or power control is occasionally infeasible due to device mobility (a case that is much less studied and understood; cf. Section 2.3 for a detailed discussion). As we discuss in the rest of this paper, this instability is in some sense the root cause for the lack of both convergence and optimality results when the network is stochastic and time-varying. Under this light, the explicit use of *all* past iterates holds great promise for the overall stability of a power management policy, as the influence of the last iterate cannot have a dominating effect over the algorithm's previous iterates (provided they are utilized in an intelligent, memory-efficient way).

Our aim in this paper is to provide a distributed power control algorithm satisfying the above desiderata. This task faces two key challenges from a practical perspective: First, such an algorithm cannot take for granted that each transmitter has access to the power characteristics of all other transmitters, as such information is typically unavailable in practical scenarios. This dictates that any given transmitter can only make explicit use of link-specific information such as its signal-to-interference plus noise ratio (SINR) and/or the total interference and noise at the receiver (i.e. information that can be sensed by each individual link in the network). Second (and perhaps more stringently), a transmitter should not be required to store all past information (including past power, SINR and/or interference measurements) in an explicit fashion. If met, this requirement is highly desirable for memory-constrained wireless transmitter where such bookkeeping is in general infeasible – in words, past information must be exploited in an implicit, parsimonious manner.

1.1. Our Contributions and Related Work

Our main contributions are highlighted below in the context of related work on the topic:

1. **A Learning and Game Design Approach to Power Management.**   To tackle the power management question (in both deterministic and stochastic networks) we map the original problem to that of finding the unique Nash equilibrium of a suitably designed continuous game (i.e. a game with continuous action spaces) and then apply a well-known no-regret learning algorithm. To streamline the presentation of the results, we start by formally describing

in Section 3 the two simple and efficient power control algorithms based on *dual averaging* (Variants A and B), both of which only require very little operational overhead while leveraging past information in an efficient way. In particular, the information on past power iterates is represented in the most parsimonious form possible: it takes only a constant amount of memory independent of the number of time steps involved. In fact, the amount of memory required is the same as that of the algorithms using only the last power iterates (e.g. FM), even though the latter do not make explicit use of past power information.

Furthermore, by mapping the proposed power management algorithms to a game-theoretic learning problem (our second contribution; see below and Section 4), we are able to examine in depth the properties of the two variant dual averaging algorithms and establish theoretical guarantees in both deterministic (static) and stochastic (time-varying) environments. More precisely, in the case of static channels, we show that feasibility of a network directly implies convergence of dual averaging to the joint optimum transmission power vector (Theorem 6).[2] Furthermore (and perhaps most surprisingly), in the stochastic case (where the fluctuating network is sometimes feasible and sometimes not), we show that convergence to a *deterministic* power vector is still guaranteed with probability 1 as long as the network is feasible *in the mean* – i.e. even if the network is infeasible with positive probability (Theorem 7).

This last property is particularly appealing because it incorporates elements of both stability and optimality (and hence robustness): the former because the proposed algorithms converge almost surely to a *fixed* constant power vector (despite the persistent, random fluctuations in the network); the latter because the algorithm's end state is an optimal solution of the network's power management problem with respect to the network's mean value. This comes in sharp contrast to the FM algorithm which, when the channel is feasible on average, may fail to converge altogether – or, at best, only converges in distribution to a power profile that is not optimal in any way (Zhou et al. 2016).

Importantly, establishing the algorithms' convergence takes us through a distinct (albeit related) line of research, namely utility-based models of power management in wireless networks (Famolari et al. 1999, Alpcan et al. 2002, 2006, Fan et al. 2006, Menache and Ozdaglar

2010, Han et al. 2014). This flourishing literature has taken the view that each device is a self-interested entity with its individual objective depending on how much power it uses as well as how much power all the other devices use (via the interference that they create). In this literature, the cost function of each device (as a function of, say, signal-to-interference ratio) is modeled explicitly and pertains to the utility of each user of that device: in other words, the resulting game is *a priori* assumed to model the underlying reality. In contrast, the "game design" approach that we take in this paper leads to a virtual game that only enters the problem as a theoretical tool to aid the design of robust power management algorithms (which are implemented at the device level and are not subject to game-theoretic rationality postulates). Nevertheless, the resulting games do admit a natural interpretation as they effectively measure the distance betweeen the users' achieved throughput and their target rates.

The idea of "game design" has been explored before and our approach here is inspired by the remarkable works of Candogan et al. (2010) and Li and Marden (2013). In more detail, Candogan et al. (2010) designed a near-potential game for the maximum weighted-sum-of-rates problem and used best-response dynamics to derive a power control algorithm. Li and Marden (2013) designed a potential game and used a particular method (called gradient-play) to derive a distributed optimization algorithm for the network consensus problem. In addition to considering a different problem altogether (power minimization subject to rate constraints), our work here differs from the above in several key aspects: First, our game-theoretic analysis is not limited to potential (or near-potential) games, but it instead applies to all (weighted) monotone games (cf. Section 4). Second, by focusing on online gradient descent (see below), our results can also be framed in the context of no-regret learning in games. Finally, to tackle the stochastic regime, we consider a more general framework with imperfect feedback which is only accurate up to a bounded-variance error (again, see below). The above elements cannot be treated with the techniques of Candogan et al. (2010) and/or Li and Marden (2013), leading us to introduce a new set of tools from stochastic approximation, martingale limit theory and variational analysis.

2. **Online Gradient Descent and No-Regret Learning in Games.** To establish the convergence properties of the proposed power management algorithms, we consider a much more general framework for no-regret learning in games with continuous action sets and individually concave (but otherwise unknown) payoff functions. Working at this level of generality achieves two distinct goals: First, by abstracting away the problem's functional details, it shows that distributed power control can be interpreted in a natural way as a multi-agent regret minimization problem (more on this below). Second, it allows us to provide a general convergence result for no-regret learning in concave games with imperfect first-order, gradient feedback. As such, our theoretical investigation here falls within the broader inquiry of game-theoretic learning, an area that stands at the intersection of learning and game theory and that seeks to answer the following question: what is the evolution of play when every player adopts a no-regret learning algorithm? In particular, *if all players of a repeated game employ an updating rule that leads to no regret, do their actions converge to a Nash equilibrium of the one-shot game?*

The first (and, to this date, the most widely studied) algorithm for no-regret learning in online decision processes with convex action spaces is the online gradient descent (OGD) method pioneered by Zinkevich (2003).[3] Up to a multiplicative constant related to the problem's dimensionality, the $\mathcal{O}(\sqrt{T})$ regret minimization rate of OGD is well-known to be min-max optimal (Shalev-Shwartz et al. 2012), so OGD has seen extensive use in online decision problems. With this in mind, our first key observation is that the proposed power control algorithms can be seen as OGD schemes for a specific utility function capturing each user's power control objective (the "game design" component described above).

The importance of this observation is clear: a no-regret learning algorithm provides a player with sharp worst-case guarantees, so it is natural to expect that every player will adopt such an algorithm in selecting their actions in the decision process. On that account, a flourishing line of research tackling this problem has emerged in the past decade or so. However, most

of this literature has focused on the convergence of the time-averaged sequence of play $\bar{\mathbf{x}}^t = \sum_{k=1}^{t} \gamma_k \mathbf{x}^k / \sum_{k=1}^{t} \gamma^k$ (where γ^t is the step-size and $\mathbf{x}^t$ is the joint action at time t) as opposed to the actual sequence of actions employed by the players. For instance, in the context of finite games, it is well-known that if each player employs a no-regret learning algorithm, the time-averaged sequence of play converges to the game's Hannan set, also known as the set of *coarse correlated equilibria* (Cesa-Bianchi and Lugosi 2006). Refining this result, if each player plays employs a no-internal-regret learning algorithm (a stronger regret notion), the time-averaged sequence of play converges to the game's set of correlated equilibria, a strict refinement of Hannan-consistent strategies (Cesa-Bianchi and Lugosi 2006). On the other hand, convergence to Nash equilibrium is, in the words of Cesa-Bianchi and Lugosi (2006) "considerably more difficult", because Nash equilibrium is significantly finer than its correlated variants:[4] in general, these convergence results do not extend to Nash equilibria.

To account for this, a growing corpus of literature has been devoted to studying the problem of equilibrium convergence under no-regret learning, typically focusing on special classes of games (such as convex potential games, zero-sum games, routing games, etc.). Here again, most efforts have focused on the convergence of the time-averaged sequence of play because of its connection to the players' regret (Krichene et al. 2015, Balandat et al. 2016, Lam et al. 2016). Nevertheless, the convergence of the actual sequence of play is also crucial to investigate for several reasons: First, from a practical standpoint, since the players' payoffs are determined by the sequence of chosen actions and not some fictitious variant thereof, convergence to Nash equilibrium should also be stated in terms of the actual sequence of play. Second, from a theoretical standpoint, convergence of the latter implies convergence of the former, so it is a stronger result.

Our focal point in this paper is the class of λ-*monotone* games, i.e. games that satisfy Rosen's diagonal strict concavity condition (Rosen 1965a). Mathematically, this condition implies that the profile of the players' individual payoff gradients is a monotone operator in the sense

of variational analysis (possibly up to a player-specific weight coefficient); as an immediate corollary, if a game is monotone, it admits a unique Nash equilibrium. Importantly, going back to the problem of power control, the key tie-in observation is that our game-theoretic reformulation leads to a monotone game whenever power control is feasible on average (Lemma 3). Thus, by establishing the convergence of OGD in monotone games, we readily recover the convergence of our proposed power control policies when power control is feasible in the first place.

Motivated by the random, time-varying nature of channel feasibility in realistic wireless environments, we frame all of the above in a bona fide stochastic setting where exact gradient information is not available, either because the players' payoffs are themselves stochastic in nature, or because the players' feedback is contaminated by noise, observation errors, and/or other exogenous stochastic effects. To model all this, we consider a noisy feedback model where players only have access to a first-order oracle providing unbiased, bounded-variance estimates of their payoff gradients at each step. Apart from this, players are assumed to operate in a "black box" setting, without any knowledge of the game's structure or their payoff functions (or even that they are playing a game).

In this general framework, our main results are as follows: For a general class of continuous games called λ-monotone games (first studied by Rosen 1965b), OGD converges to the unique Nash equilibrium in the perfect gradient case (Theorem 4). Further, for λ-monotone games, even when only a noisy and unbiased gradient is available, OGD converges to the unique Nash equilibrium almost surely (Theorem 5).

2. Model, Background, Motivation

We describe below the target-rate power management problem on wireless networks (Weeraddana et al. 2012, Tan 2014). After introducing the problem in Section 2.1, we discuss in Section 2.2 the well-known Foschini-Miljanic (FM) power control algorithm (Foschini and Miljanic 1993). This discussion will provide an account of some of the drawbacks of the FM algorithm, both quantitative and qualitative, and will serve as the motivation of the paper (cf. Section 2.3).

10

Zhou, Mertikopoulos, Moustakas, Bambos and Glynn: *Robust Power Management via Learning and Game Design*
Article submitted to *Operations Research*; manuscript no. (Please, provide the manuscript number!)

2.1. Setup

Consider a wireless network of N communication links, each link consisting of a transmitter and an intended receiver. Assume further that the i-th transmitter transmits with power p_i and let $\mathbf{p} = (p_1, \ldots, p_N) \in \mathbf{R}_+^N$ denote the joint power profile of all users (transmitters) in the network. In this context, the most commonly used measure of link service quality is the signal-to-interference-and-noise ratio (SINR). Intuitively, link i's SINR depends not only on how much power its transmitter is employing, but also on how much power all the other transmitters are concurrently employing. Specifically, link i's SINR, which we denote by $r_i(\mathbf{p})$, is given by the following ratio:

$$r_i(\mathbf{p}) = \frac{G_{ii} p_i}{\sum_{j \neq i} G_{ij} p_j + \eta_i}, \tag{1}$$

where η_i is the thermal noise associated with the receiver of link i and $G_{ij} \geq 0$ is the power gain between transmitter j and receiver i, representing the interference caused to receiver i by transmitter j per unit transmission power used. We further assume throughout the paper that $G_{ii} > 0$ for otherwise transmission for link i is meaningless. We collect all the power gains G_{ij} into the gain matrix $\mathbf{G}$ and all the thermal noises into the noise vector η. Note that the power gain matrix $\mathbf{G}$ depends on the underlying network topology of the wireless links. Each link has a target SINR threshold $r_i^* > 0$: the minimum acceptable service quality threshold for that link.

The target-rate power management problem (Weeraddana et al. 2012, Tan 2014) then lies in finding a power assignment $\mathbf{p}$, such that the following quality-of-service constraints hold:

$$r_i(\mathbf{p}) \geq r_i^*, \forall i. \tag{2}$$

In order to find a joint transmission power $\mathbf{p}$ that satisfies the quality-of-service constraints in Equation (2), such a $\mathbf{p}$ must exist in the first place: the notion of wireless network channel feasibility, formalized in the next definition, characterizes such scenarios.

DEFINITION 1. The channel given by $(\mathbf{G}, \eta)$ is *feasible* with respect to a target SINR vector $r^* = (r_1^*, \ldots, r_N^*)$ if there exists a $\mathbf{p}$ satisfying Equation (2). The channel is otherwise said to be *infeasible*.

In their original paper, Foschini and Miljanic (1993) presented a simple necessary and sufficient condition for deciding when a channel is feasible. To state it, we first need a convenient and equivalent characterization of a wireless network channel:

DEFINITION 2. A wireless network channel (or channel for short) specified by $(\mathbf{G}, \eta)$ can be alternatively represented by the pair (W, γ) consisiting of the following components:

1. The re-weighted gain matrix W, where for $i, j \in \{1, 2, \ldots, N\}$:

$$W_{ij} := \begin{cases} 0, & i = j \\ r_i^* \frac{G_{ij}}{G_{ii}}, & i \neq j. \end{cases} \tag{3}$$

2. The re-weighted noise vector γ, where

$$\gamma_i = r_i^* \frac{\eta_i}{G_{ii}}, i \in \{1, 2, \ldots, N\}. \tag{4}$$

THEOREM 1. *A channel is feasible with respect to r^* if and only if the largest[5] eigenvalue $\lambda_{\max}(W)$ of the re-weighted gain matrix W satisfies $\lambda_{\max}(W) < 1$.*

When the channel is feasible, there is a special joint transmission power $\mathbf{p}^*$ that satisfies the quality-of-service constraints in Equation 2:

FACT 1. Given a feasible channel $(\mathbf{G}, \eta)$, we have $\lambda_{\max}(W) < 1$, which implies that $(\mathbf{I} - W)^{-1}$ exists and is component-wise strictly positive. This further implies that the joint transmission power $\mathbf{p}^* = (\mathbf{I} - W)^{-1} \gamma$ satisfies the quality-of-service constraints given in Equation (2) and it is component-wise strictly positive. Finally, $\mathbf{p}^*$ is the unique vector that satisfies the following property: if $\mathbf{p}$ is any vector satisfying Equation (2), then $\mathbf{p}^* \leq \mathbf{p}$ (component-wise).

In other words, $\mathbf{p}^*$ is the "smallest" joint transmission power that satisfies the quality-of-service constraints. To highlight the importance of this quantity and to recognize the fact that the results in this paper will mostly pertain to $\mathbf{p}^*$, we have the following definition:

DEFINITION 3. In a feasible channel $(\mathbf{G}, \eta)$ (or equivalently (W, γ)), $\mathbf{p}^*$ defined in Fact 1 is called the *optimal joint transmission power vector*.

2.2. Foschini-Miljanic Power Control Algorithm

We now present the well-known Foschini-Miljanic (FM) power control algorithm, which finds the optimal joint transmission power if one exists (i.e. in a feasible channel). Following the standard convention in wireless communications literature (Han et al. 2011), the transmission power p_i for transmitter i is assumed to lie in a compact interval $\mathcal{P}_i = [0, p_i^{\max}]$. Therefore, $\mathbf{p}$ is constrained to lie in the feasible support set $\mathcal{P} \triangleq \prod_{i=1}^N \mathcal{P}_i = \prod_{i=1}^N [0, p_i^{\max}]$. We shall adopt this convention for the rest of the paper. The FM algorithm is then formally given in Algorithm 1 below:

Algorithm 1 FM Algorithm: Bounded Power Support

1: Each link i chooses an initial power $p_i^0 \in [0, p_i^{\max}]$.

2: **for** $t = 0, 1, 2, \ldots$ **do**

3:    **for** $i = 1, \ldots, N$ **do**

4:        $p_i^{t+1} = \min(p_i^t \frac{r_i^*}{r_i(\mathbf{p}^t)}, p_i^{\max})$

5:    **end for**

6: **end for**

In the classical power control setting, the channel $(\mathbf{G}, \eta)$ is assumed to be deterministic and time-invariant, i.e. $(\mathbf{G}, \eta)$ remains the same from iteration to iteration. By a monotonicity argument, Foschini and Miljanic (1993) provide the following characterization:

THEOREM 2. *Let the channel $(\mathbf{G}, \eta)$ be deterministic and time-invariant.*

- *If the channel is feasible with respect to r^* and if the power support includes the optimal power vector (i.e. $\mathbf{p}^* \in \mathcal{P}$), then the joint power iterate $\mathbf{p}^t$ in Algorithm 1 converges to the optimal joint transmission power $\mathbf{p}^*$, irrespective of the initial point p^0.*

- *If the channel is infeasible with respect to r^* or if the power support does not include the optimal power vector (i.e. $\mathbf{p}^* \notin \mathcal{P}$), then the joint power iterate in Algorithm 1 converges to $p^{\max}$, irrespective of the initial point p^0.*

2.3. Motivation of the Paper

Although the FM algorithm enjoys good convergence properties in a deterministic and time-invariant channel (as given in Theorem 2 above), it quickly loses its appeal in cases where the channel is stochastic and time-varying (i.e. when $(\mathbf{G}^t, \eta^t)_{t=0}^{\infty}$ are stochastic processes): this can occur when the transmitters and receivers in the wireless network are moving while communicating with each other, thereby causing the channel gain matrix (and potentially the thermal noises) to fluctuate from iteration to iteration. This is because FM, in only using the last power iterate $\mathbf{p}^t$ to determine the next power iterate $\mathbf{p}^{t+1}$, has certain inherent instability, which is particularly manifested when the underlying channel is stochastic and time-varying.

When the channel environment is stochastic and time-varying, we use W^t and γ^t to denote the random re-weighted gain matrix and the random re-weighted noise at iteration t, respectively. We will use P^t to denote the random power vector at iteration t generated by some power control algorithm. More concisely, denoting $\Pi_{\mathcal{P}}(\mathbf{y}) = \arg\min_{\mathbf{x} \in \mathcal{P}} \|\mathbf{x} - \mathbf{y}\|$ and denoting the projection operator $\Pi_{\mathcal{P}}(\cdot)$ for $\mathcal{P}$, the FM update can be written as

$$P^{t+1} = \Pi_{\mathcal{P}}(W^t P^t + \gamma^t). \tag{5}$$

In a stochastic and time-varying channel, the power iterates generated by FM will be random variables and may fail to converge altogether. Even when FM does converge, it will at best, under uncertain conditions of the channel, converge to a stationary distribution (Zhou et al. (2016)), as opposed to a deterministic power vector, which is more desirable. This reveals two main drawbacks of FM. First, as mentioned previously, convergence to a limiting stationary distribution is not the most desired case. Consequently, when operating in a stochastic and time-varying channel, FM is not very stable. As alluded to before, the fact that FM uses only the last power iterate is the reason that such stochastic stability type of results is the best one can hope for on FM. Second, even when FM converges to a stationary distribution, it is not clear what performance guarantees are achieved by that limiting power distribution. Specifically, the stationary distribution that FM converges to is not optimal in any sense.

The above two drawbacks lead to the problem of designing a distributed power control algorithm that has both stability and performance guarantees. More specifically, in light of the root cause of the instability of FM, it is natural to ask whether it is possible to incorporate all the past power iterates to synthesize a distributed power control scheme so as to stabilize the power iterate more quickly, such that the resulting power iterate converges (almost surely) to a fixed vector even in the presence of a stochastic and time-varying channel? If so, would this vector be optimal in some (average) sense? As we shall see in the next section, both questions have affirmative answers.

3. Power Control via Dual Averaging

In this section, we present two new and closely-related distributed power control algorithms that utilize all the power iterates in the past to achieve better stability and optimality guarantees. The design of such a distributed algorithm (that uses past information) faces at least two challenges. First, such an algorithm cannot assume that each transmitter has access to the power used by all the other transmitters, as such communications is infeasible in practice. This dictates that a transmitter can only use the aggregate information, such SINR and/or total interference and noise (i.e. information that can be sensed by each individual link) as opposed to the exact individual powers. Second, more stringently, one should not expect a transmitter to store all the past information that is available to it, which could include its own past transmission powers, past SINRs etc. This additional constraint, if met, can be highly desirable in practice because for a memory-constrained wireless transmitter, such bookkeeping is in general infeasible. This second constraint further implies that an economic representation that incorporates all such information is needed.

Here we propose the dual averaging power control algorithm that satisfies those two constraints. The dual averaging algorithm has two variants, which we call Variant A and Variant B. For each variant, the past information is represented and stored in the most economic form possible: it takes only constant amount of memory independent of time steps. For ease of exposition, and to highlight the specific modeling assumptions, we break the discussion into two cases: one where the channel is deterministic and time-invariant, the other where the channel is stochastic and time-varying. However, each variant of the dual averaging stays the same (except for the slight difference on the notation) across different environments.

3.1. Deterministic and Time-Invariant Channel

Here we present the dual-averaging algorithms for the case where the channel $(\mathbf{G}, \eta)$ is deterministic and time-invariant. Algorithm 2 provides a formal description of variant A of dual averaging.

Algorithm 2 Dual Averaging Variant A: Deterministic and Time-Invariant Channel

1: Each link i chooses an initial $y_i^0 \in \mathbf{R}$.

2: **for** $t = 0, 1, 2, \ldots$ **do**

3: **for** each link i **do**

4:   $p_i^t = \Pi_{\mathcal{P}_i}(y_i^t)$

5:   $y_i^{t+1} = y_i^t - \frac{1}{t+1}(G_{ii}p_i^t - r_i^*(\sum_{j \neq i} G_{ij}p_j^t + \eta_i))$

6: **end for**

7: **end for**

REMARK 1. Several remarks are in order. First, note that y_i^t's serve to fulfill the role of keeping a compact representation that aggregates all the past information at any given time t, thereby eliminating the need to keep track of the past $\mathbf{p}^t$'s. Second, as we shall see and make precise in Section 5, each y_i^t is the weighted average of the gradients of a certain cost function. Consequently, $\mathbf{y}^t$ (the vector of all individual y_i^t's) resides in the dual space of the space $\mathcal{P}$ of possible transmission power; hence the name "dual averaging". In fact, Line 4 shows that the projection transforms a point in this dual space to a point in the action space (i.e. $\mathcal{P}$). In particular, while $\mathcal{P}$ is a bounded set, the dual space that $\mathbf{y}^t$ is the whole $\mathbf{R}^N$. Finally, note that to perform the update in Line 5 does not require transmitter i to know the transmission powers used by others: it need only know the interference and noise as a whole as well the SINR.

We now give the description of Variant B in Algorithm 2: the only difference lies in updating the discounted gradient y_i^t. All the points in Remark 1 apply to this variant except the last one: note that in Variant B, only the SINR $\frac{\sum_{j \neq i} G_{ij}p_j^t + \eta_i}{G_{ii}p_i^t}$ is needed, because from this SINR and p_i^t, the ratio $\frac{\sum_{j \neq i} G_{ij}p_j^t + \eta_i}{G_{ii}}$ can be recovered. Consequently, the requirement to implement Variant B is even less

stringent than that of Variant A. However, as we shall see in later (Theorem 7), Variant B needs

slightly stronger assumption on the channel gains for convergence and optimality.

Algorithm 3 Dual Averaging Variant B: Deterministic and Time-Invariant Channel

1: Each link i chooses an initial $y_i^0 \in \mathbf{R}$.

2: **for** $t = 0, 1, 2, \ldots$ **do**

3:     **for** each link i **do**

4:         $p_i^t = \Pi_{\mathcal{P}_i}(y_i^t)$

5:         $y_i^{t+1} = y_i^t - \frac{1}{t+1}\left(p_i^t - r_i^* \frac{\sum_{j \neq i} G_{ij} p_j^t + \eta_i}{G_{ii}}\right)$

6:     **end for**

7: **end for**

3.2. Stochastic and Time-Varying Channel

Next we present the two variants of the dual averaging algorithm for the stochastic and time-varying channel case. Before stating the algorithm, we shall first present the model of the stochastic and time-varying channel within which the algorithm will operate. We adopt the same stochastic model of a wireless network channel as in Zhou et al. (2016):

ASSUMPTION 1. $(\mathbf{G}^t, \eta^t)$ *is drawn **iid** from an arbitrary (discrete or continuous, bounded or unbounded) support on $\mathbf{R}_+^{N \times N} \times \mathbf{R}_+^N$, satisfying the following assumptions:*

1. *Finite mean:* $\forall i, j, \forall t, \mathbf{E}[G_{ij}^t] < \infty, \mathbf{E}[\eta_i^t] < \infty$.

2. *Finite variance:* $\forall i, j, \forall t, \mathbf{Var}[G_{ij}^t] < \infty, \mathbf{Var}[\eta_i^t] < \infty$.

Note that under this model, (G_{ij}^t, η_k^t) can be arbitrarily correlated with $(G_{i'j'}^t, \eta_{k'}^t)$ and $\mathbf{G}^t$ can be correlated with η^t. Algorithm 4 gives the description of Variant A in the stochastic and time-varying channel case. Here we use the upper case for gradient and power to highlight the fact that all the quantities are now random variables.

Continuing with the same notation, we next give the description of Variant B as follows.

Algorithm 4 Dual Averaging Variant A: Stochastic and Time-Varying Channel

---

1: Each link i chooses an initial $Y_i^0 \in \mathbf{R}$.

2: **for** $t = 0, 1, 2, \ldots$ **do**

3:      **for** each link i **do**

4:         $P_i^t = \Pi_{\mathcal{P}_i}(Y_i^t)$

5:         $Y_i^{t+1} = Y_i^t - \frac{1}{t}(G_{ii}^t P_i^t - r_i^*(\sum_{j \neq i} G_{ij}^t P_j^t + \eta_i^t))$

6:      **end for**

7: **end for**

Algorithm 5 Dual Averaging Variant B: Stochastic and Time-Varying Channel

---

1: Each link i chooses an initial $Y_i^0 \in \mathbf{R}$.

2: **for** $t = 0, 1, 2, \ldots$ **do**

3:      **for** each link i **do**

4:         $P_i^t = \Pi_{\mathcal{P}_i}(Y_i^t)$

5:         $Y_i^{t+1} = Y_i^t - \frac{1}{t}(P_i^t - r_i^* \frac{\sum_{j \neq i} G_{ij}^t P_j^t + \eta_i^t}{G_{ii}^t})$

6:      **end for**

7: **end for**

---

4. Online Gradient Descent Learning in Continuous Games

In this section, we present the framework of online gradient descent (OGD) learning on continuous games that both serves as the foundation for the game-design approach of power control and enables us to subsequently establish (in Section 5 and Section 6) the theoretical guarantees of dual averaging power control algorithms in both the constant and time-invariant channel and the stochastic and time-varying channel. For both simplicity and direct applicability to the power control application, the presentation here will be on the simplest setting needed: the action set of each player is assumed to lie in $\mathbf{R}$. However, all the results in this section generalize in a straightforward manner to the settings where each player's underlying action set is some finite dimensional real vector space. Wherever possible, we shall use the notation that matches the power

control setting to make explicit the connection between OGD learning on concave games and power control as well as to ease the subsequent transition from the former to the latter.

4.1. λ-Monotone Games

We start with the definition of a λ-Monotone game, which will set the stage for both the theoretical study in this section and the practical design and application for the subsequent sections.

DEFINITION 4. A continuous game $\mathcal{G}$ is a tuple $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^{N} \mathcal{X}_i, \{u_i\}_{i=1}^{N})$, where $\mathcal{N}$ is the set of N players $\{1, 2, \ldots, N\}$, $\mathcal{X}$ is the joint action space with $\mathcal{X}_i$ being the action space for player i and $u_i : \mathcal{X} \to \mathbf{R}$ is the utility function for player i, such that the following assumptions hold:

1. Each $\mathcal{X}_i$ is a compact and convex subset of $\mathbf{R}$.

2. Each u_i is continuous in $\mathbf{x}$ and continuously differentiable in x_i and $\frac{\partial u_i(\mathbf{x})}{\partial x_i}$ is continuous in $\mathbf{x}$.

Throughout the paper, $\mathbf{x}_{-i}$ denotes the joint action of all players but player i. Consequently, the joint action $\mathbf{x}$ will frequently be written as $(x_i, \mathbf{x}_{-i})$. Further, we denote $\mathbf{v}(\mathbf{x})$ to be the column vector of partial derivatives of the utility functions: $\mathbf{v}(\mathbf{x}) = (v_1(\mathbf{x}), \ldots, v_N(\mathbf{x}))$, where $v_i(\mathbf{x}) \triangleq \frac{\partial u_i(\mathbf{x})}{\partial x_i}$. Note that per the definition of a continuous game (Definition 4), the gradient $\mathbf{v}(\mathbf{x})$ always exists and is a continuous function on the joint action space $\mathcal{X}$.

DEFINITION 5. A continuous game $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^{N} \mathcal{X}_i, \{u_i\}_{i=1}^{N})$ is called λ-monotone for some $\lambda \in \mathbf{R}_{++}^{N}$ if the following holds (with equality if and only if $\mathbf{x} = \mathbf{y}$):

$$\sum_{i=1}^{N} \lambda_i (v_i(\mathbf{x}) - v_i(\mathbf{y}))(x_i - y_i) \leq 0, \forall \mathbf{x}, \mathbf{y} \in \mathcal{X}. \tag{6}$$

Condition given in Equation (6) is essentially the diagonal stricty concavity condition first introduced in Rosen (1965b). Note that this condition necessarily implies that all utility functions u_i are concave. In the special case that $\lambda = \mathbf{1}$, we simply say the game is monotone. We use the terminology monotone because when $\lambda = \mathbf{1}$, this condition is equivalent to that $v(\mathbf{x})$ is a monotone operator, a key concept in variational analysis Rockafellar and Wets (2009). Next, we recall the definition of a Nash equilibrium, a key quantity that will concern us for the rest of the paper.

DEFINITION 6. Given a continuous game $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^{N} \mathcal{X}_i, \{u_i\}_{i=1}^{N})$, $\mathbf{x}^* \in \mathcal{X}$ is called a Nash equilibrium if for each $i \in \mathcal{N}$, $u_i(x_i^*, \mathbf{x}_{-i}^*) \geq u_i(x_i, \mathbf{x}_{-i}^*), \forall x_i \in \mathcal{X}_i$.

The celebrated result established in Rosen (1965b) is that there is always a unique Nash equilibrium for every λ-monotone game:

THEOREM 3. *(**Rosen**)* *A λ-monotone game $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^{N} \mathcal{X}_i, \{u_i\}_{i=1}^{N})$ admits a unique Nash equilibrium $\mathbf{x}^*$.*

Rosen (1965b) also gives a convenient sufficient condition ensuring that a continuous game is λ-monotone, a condition that is used later to establish that the designed games are λ-monotone.

LEMMA 1. *Given a continuous game $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^{N} \mathcal{X}_i, \{u_i\}_{i=1}^{N})$, where each u_i is twice continuously differentiable. For each $\mathbf{x} \in \mathcal{X}$, define the λ-weighted Hessian matrix $H^\lambda(\mathbf{x})$ as follows:*

$$H_{ij}^\lambda(\mathbf{x}) = \frac{1}{2}\lambda_i \frac{\partial v_i(\mathbf{x})}{\partial x_j} + \frac{1}{2}\lambda_j \frac{\partial v_j(\mathbf{x})}{\partial x_i}. \tag{7}$$

If $H^\lambda(\mathbf{x})$ is negative-definite for every $\mathbf{x} \in \mathcal{X}$, then $\mathcal{G}$ is λ-monotone.

4.2. Online Gradient Descent: Perfect and Imperfect Information

When players play a repeated game with each stage game being a fixed continuous game, it is an interesting question as to what learning dynamics the players would adopt. A well-known class of learning dynamics that enjoys the no-regret property (see Zinkevich 2003 and Shalev-Shwartz et al. 2012 for a precise formulation and proof) when the utility function is concave is online gradient descent[6] (OGD), as presented below:

Several comments are in order here. First, y_i^t can be seen as an auxiliary (dual) variable that accumulates gradient steps discounted by the step-size sequence $\{\alpha^t\}_{t=1}^{\infty}$; the chosen actions are then given by the (lazy) projection of y_i^t onto $\mathcal{X}_i$. As for the step-size sequence α^t, Algorithm 7 can be run with any positive, non-increasing sequence that satisfies the widely used non-summable-but-square-summable ("$\ell^2 - \ell^1$") condition below:

Algorithm 6 Online Gradient Descent under Perfect Information

---

1: Each player i chooses an initial y_i^0.

2: **for** $t = 0, 1, 2, \ldots$ **do**

3:      **for** $i = 1, \ldots, N$ **do**

4:         $x_i^t = \arg\min_{x_i \in \mathcal{X}_i} \|y_i^t - x_i\|$

5:         $y_i^{t+1} = y_i^t + \alpha^t v_i(\mathbf{x}^t)$

6:      **end for**

7: **end for**

DEFINITION 7. A positive and non-increasing sequence $\{\alpha^t\}_{t=0}^\infty$ is called *slowly vanishing* if $\sum_{t=0}^\infty \alpha^t = \infty$ and $\sum_{t=0}^\infty (\alpha^t)^2 < \infty$.

We can further generalize OGD to allow for imperfect information, where an exact gradient is not available. Algorithm 7 gives a formal description of this generalized version. The main difference between Algorithm 7 and Algorithm 6 lies in Step 5, where only a noisy gradient is available. In addition, the iterates are capitalized to make explicit the fact that due to the noisy gradient used in Step 5, they are now random variables. Specifically, we have used the capital letters X_i^t and Y_i^t in Algorithm 7 because these iterates are now random variables as a result of the noisy gradients $\tilde{v}_i$. Of course, in order for convergence to be guaranteed, $\tilde{v}_i(\mathbf{X}^t)$ cannot be just any noisy perturbation of the gradient. Here we make a rather standard assumption on the noisy gradient:

Algorithm 7 Online Gradient Descent under Imperfect Information

---

1: Each player i chooses an initial Y_i^0.

2: **for** $t = 0, 1, 2, \ldots$ **do**

3:      **for** $i = 1, \ldots, N$ **do**

4:         $X_i^t = \arg\min_{X_i \in \mathcal{X}_i} \|Y_i^t - X_i\|$

5:         $Y_i^{t+1} = Y_i^t + \alpha^t \tilde{v}_i(\mathbf{X}^t)$

6:      **end for**

7: **end for**

ASSUMPTION 2. *Let $\mathcal{F}^t$ be the canonical filtration induced by the (random) iterates up to time t:*

$\mathbf{X}^0, \mathbf{X}^1, \ldots, \mathbf{X}^t$. *We assume:*

1. *The noisy gradients are conditionally unbiased:*

$$\forall i \in \mathcal{N}, \forall t = 0, 1, \ldots, \mathbf{E}[\tilde{v}_i(\mathbf{X}^t) \mid \mathcal{F}^t] = v_i(\mathbf{X}^t), a.s.. \tag{8}$$

2. *The noisy gradients are bounded in mean square:*

$$\forall i \in \mathcal{N}, \forall t = 0, 1, \ldots, \mathbf{E}[\|\tilde{v}_i(\mathbf{X}^t)\|^2 \mid \mathcal{F}^t] \leq V, a.s., \tag{9}$$

for some constant $V > 0$, where $\| \cdot \|$ is some finite dimensional norm (note that all finite

dimensional norms are equivalently up to a multiplicative constant).

REMARK 2. An equivalent and useful characterization of Assumption 2 is that the noisy gradient

can be decomposed as $\tilde{v}_i(\mathbf{X}^t) = v_i(\mathbf{X}^t) + \xi_i^{t+1}$, where the noise $\xi^t = (\xi_i^t)_{i=1}^N$ satisfies:

1. $\forall t = 0, 1, \ldots, \mathbf{E}[\xi^{t+1} \mid \mathcal{F}^t] = 0$ (a.s.).

2. $\forall t = 0, 1, \ldots, \mathbf{E}[\|\xi^{t+1})\|^2 \mid \mathcal{F}^t] \leq \Xi$ (a.s.).

4.3. λ-**Fenchel Coupling**

We now introduce an important notion, λ-Fenchel coupling, that will serve as an energy function

and will play a indispensable role in establishing the convergence of the OGD dynamics. For ease

of exposition, we will directly define λ-Fenchel coupling in the current continuous game context

(where each $\mathcal{X}_i$ is a subset of $\mathbf{R}$), although it shall be instantly clear that one can easily extend it

to more general contexts where each $\mathcal{X}_i$ is a finite dimensional real vector space.

We begin with some basic definitions and notation:

DEFINITION 8. Let $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathcal{X}_i, \{u_i\}_{i=1}^N)$ be a continuous game and, for each player i, let

$$h_i(x_i) = \frac{1}{2}\|x_i\|^2 \tag{10}$$

denote the Euclidean squared norm on $\mathcal{X}_i$. Then:

1. The *convex conjugate* $h_i^* : \mathbf{R} \to \mathbf{R}$ of h_i is defined as:

$$h_i^*(y_i) = \max_{x_i \in \mathcal{X}_i} \{\langle y_i, x_i \rangle - h_i(x_i)\} \quad \text{for all } y_i \in \mathbf{R}. \tag{11}$$

2. The *choice function* $C_i : \mathbf{R} \to \mathcal{X}_i$ for player i is defined as the (lazy) projection map

$$C_i(y_i) = \arg\min_{x_i \in \mathcal{X}_i} \|x_i - y_i\| = \arg\max_{x_i \in \mathcal{X}_i} \{\langle y_i, x_i \rangle - h_i(x_i)\} \quad \text{for all } y_i \in \mathbf{R}. \tag{12}$$

3. The *λ-Fenchel coupling* $F^\lambda : \mathcal{X} \times \mathcal{R}^N \to \mathbf{R}$ is defined as:

$$F^\lambda(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^N \lambda_i (h_i(x_i) + h_i^*(y_i) - \langle y_i, x_i \rangle), \forall \mathbf{x} \in \mathcal{X}, \forall \mathbf{y} \in \mathbf{R}^N. \tag{13}$$

REMARK 3. Two things worth noting is that although the domain of h_i is $\mathcal{X}_i \subset \mathbf{R}$, the domain of its conjugate h_i^* is $\mathbf{R}$. Further, since h_i is proper, lower semi-continuous and convex, it follows that $(h_i^*)^* = h_i$ (Theorem 11.1 in Rockafellar and Wets (2009)), a result to be used later. Second, The choice function C_i projects y_i from the gradient space to x_i in the decision space, and corresponds to Line 4 in both Algorithm 6 and Algorithm 7.

The two key properties of λ-Fenchel coupling that will be useful later are:

LEMMA 2. *For all $i \in \{1, \ldots, N\}$, and all $\mathbf{x} \in \mathcal{X}$, $\tilde{\mathbf{y}}, \mathbf{y} \in \mathbf{R}^N$, we have:*

1. $F^\lambda(\mathbf{x}, \mathbf{y}) \geq \frac{1}{2} \sum_{i=1}^N \lambda_i \|C_i(y_i) - x_i\|_i^2 \geq \frac{1}{2}(\min_i \lambda_i) \sum_{i=1}^N \|C_i(y_i) - x_i\|_i^2$.
2. $F^\lambda(\mathbf{x}, \tilde{\mathbf{y}}) \leq F^\lambda(\mathbf{x}, \mathbf{y}) + \sum_{i=1}^N \lambda_i(\tilde{y}_i - y_i)(C_i(y_i) - x_i) + \frac{1}{2}(\max_i \lambda_i) \sum_{i=1}^N (\|\tilde{y}_i - y_i\|_i)^2$.

REMARK 4. Collecting each individual choice map into a vector, we obtain the aggregate choice map $C : \mathbf{R}^N \to \mathcal{X}$, with $C(\mathbf{y}) = (C_1(y_1), \ldots, C_N(y_N))$. Since each space $\mathcal{X}_i$ is normed, we can define the induced aggregate norm $\|\cdot\|$ on the joint space $\mathcal{X}$ as $\|\mathbf{x}\| = \sum_{i=1}^N \|x_i\|_i$. Henceforth, it shall be clear that the convergence in the joint space (e.g. $C(\mathbf{y}^t) \to \mathbf{x}$, $\mathbf{y}^t \to \mathbf{y}$) will be defined under the respective aggregate norm.

4.4. Convergence of Online Gradient Descent with Perfect Information

In this subsection, we tackle the problem of convergence of OGD when the perfect gradient information is available (i.e. $v_i(\mathbf{x}^t)$ is known exactly in Step 5 of Algorithm 6). Our main result here is that when a given continuous game is λ-monotone, then OGD given in Algorithm 6 converges to the unique Nash equilibrium.

THEOREM 4. *Let* $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^{N} \mathcal{X}_i, \{u_i\}_{i=1}^{N})$ *be a* λ*-monotone game with the unique Nash equilibrium* $\mathbf{x}^*$*. If the step-size sequence* α^t *of Algorithm 6 is slowly vanishing, the OGD sequence of play* $\mathbf{x}^t$ *converges to* $\mathbf{x}^*$ *for any initial condition* $\mathbf{y}^0$*.*

Since this theorem is a special case[7] of Theorem 5, we will defer its proof to that of Theorem 5.

4.5. Almost Sure Convergence of Online Gradient Descent under Imperfect Information

In this subsection, we tackle the even harder problem of convergence of OGD under imperfect information, i.e. when an exact gradient is not known but, instead, only a noisy gradient $\tilde{v}_i(\mathbf{x}^t)$ is available in Step 5 of Algorithm 7). Our main result here is that, when a given continuous game is λ-monotone, OGD given in Algorithm 7 converges to the unique Nash equilibrium almost surely.

4.5.1. Mathematical Preliminaries We begin by collecting in one place the minimal set of definitions and results that will be used later to prove the almost-sure convergence result. The first two came from Benaïm (1999).

DEFINITION 9. A semiflow ϕ on a metric space (M, d) is a continuous map $\phi : \mathbf{R}_+ \times M \to M$:

$$(t, x) \to \phi_t(x),$$

such that the semi-group properties hold: $\phi_0 = \text{identity}$, $\phi_{t+s} = \phi_t \circ \phi_s$ for all $(t, s) \in \mathbf{R}_+ \times \mathbf{R}_+$.

REMARK 5. A standard way to induce a semiflow is via an ordinary differential equation (ODE). Specifically, as mentioned in Benaïm (1999), if $F : \mathbf{R}^m \to \mathbf{R}^m$ is a continuous function and if the following ODE has a unique solution trajectory for each initial point $\tilde{x} \in \mathbf{R}^m$:

$$\frac{dx}{dt} = F(x),$$
$$x(0) = \tilde{x},$$

then $\phi_t(\tilde{x})$ defined by the solution trajectory $x(t) \in \mathbf{R}^m$ as follows is a semiflow: $\phi_t(\tilde{x}) \triangleq x(t)$ with $x(0) = \tilde{x}$. We say ϕ defined in this way is the semiflow induced by the corresponding ODE.

DEFINITION 10. *Let ϕ be a semiflow on the metric space (M, d). A continuous function $s : \mathbf{R}_+ \to M$ is an asymptotic pseudotrajectory for ϕ if for every $T > 0$, the following holds:*

$$\lim_{t \to \infty} \sup_{0 \leq h \leq T} d(s(t+h), \phi_h(s(t))) = 0. \tag{14}$$

REMARK 6. Per its definition, when s is an asymptotic pseudotrajectory for ϕ, s and ϕ are very close for sufficiently large t. Specifically, for each fixed $T > 0$, one can find a large enough t_0, such that $\forall t > t_0$, the curve $s(t+h)$ approximates the trajectory $\phi_h(s(t))$ on the interval $h \in [0, T]$ with any pre-specified degree of accuracy. A more thorough discussion on asymptotic pseudotrajectory can be found in Benaïm and Hirsch (1996).

4.5.2. Convergence Analysis We now turn to the main convergence result.

THEOREM 5. *Let $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^{N} \mathcal{X}_i, \{u_i\}_{i=1}^{N})$ be λ-monotone with the unique Nash equilibrium $\mathbf{x}^*$. If the following conditions are satisfied:*

1. *Assumption 2 holds.*

2. *Each $v_i(\mathbf{x})$ is Lipschitz continuous in $\mathbf{x}$ on $\mathcal{X}$.*

3. *The step size sequence $\{\alpha^t\}_{t=0}^{\infty}$ in Algorithm 6 is slowly vanishing.*

then the OGD sequence of play $\mathbf{X}^t$ converges to $\mathbf{x}^*$ almost surely, for any initial condition $\mathbf{Y}^0$.

REMARK 7. For clarity, we break the proof into several steps.

1. Let $B(\mathbf{x}^*, \epsilon) \triangleq \{\mathbf{x} \in \mathcal{X} \mid \|\mathbf{x} - \mathbf{x}^*\| < \epsilon\}$ be the open ball centered around $\mathbf{x}^*$ with radius ϵ. We show that any open ball around $\mathbf{x}^*$ is recurrent. Mathematically, for any $\epsilon > 0$ and any initial point $\mathbf{y}^0$, the iterate $\mathbf{X}^t$ visit $B(\mathbf{x}^*, \epsilon)$ infinitely often almost surely. Further, fix any $\delta > 0$ and consider the set $\tilde{B}(\mathbf{x}^*, \delta) \triangleq \{C(\mathbf{y}) \mid F^\lambda(\mathbf{x}^*, \mathbf{y}) < \delta\}$. Although $F^\lambda(\mathbf{x}^*, \mathbf{y})$ is not a metric, $\tilde{B}(\mathbf{x}^*, \delta)$ contains an open ball within it. Mathematically, the claim is that for any $\delta > 0$, $\exists \epsilon(\delta) > 0$ such that: $B(\mathbf{x}^*, \epsilon) \subset \tilde{B}(\mathbf{x}^*, \delta)$. Putting the above two pieces together, it follows that for any $\delta > 0$, $\mathbf{X}^t$ must almost surely visit $\tilde{B}(\mathbf{x}^*, \delta)$ infinitely often, irrespective of the initial point.

2. Next we consider the ordinary differential equation (ODE) approximation of OGD as follows:

$$\dot{\mathbf{y}} = v(\mathbf{x}),$$

$$\mathbf{x} = C(\mathbf{y}).$$

Note that this can be written as $\dot{\mathbf{y}} = v(C(\mathbf{y}))$, which we can then verify that it admits a unique solution trajectory for any given initial condition. Consequently, per Remark 5, this solution induces a semiflow[8], which we denote $\phi_t(\mathbf{y})$: it is the state at time t given it starts at $\mathbf{y}$ initially. Note that we have used $\mathbf{y}$ as the initial point (as opposed to $\mathbf{y}^0$) to indicate that the semiflow representing the solution trajectory should be viewed as a function of the initial point $\mathbf{y}$.

3. We now relate the iterates generated by OGD to the above ODE's solution. First, we connect linearly the OGD iterates $Y^0, Y^1, Y^2, \ldots, Y^k, \ldots$ at times $0, \alpha^0, \alpha^0 + \alpha^1, \ldots, \sum_{i=0}^{k-1} \alpha^i, \ldots$ respectively to form a continuous, piecewise affine curve. Namely:

$$Y(t) = Y^k + (t - \sum_{i=0}^{k-1} \alpha^i) \frac{Y^{k+1} - Y^k}{\alpha^k}, \text{for } t \in [\sum_{i=0}^{k-1} \alpha^i, \sum_{i=0}^{k} \alpha^i), k = 0, 1, \ldots,$$

where we adopt the usual convention that $\sum_{i=0}^{-1} \alpha^i = 0$. We then show that $Y(t)$ (a random trajectory) is almost surely an asymptotic pseudotrajectory of the semi-flow ϕ induced by the above ODE, under the metric induced by the dual norm $\|\cdot\|^*$ defined in Remark 4. Mathematically, we establish that

$$\forall T > 0, \lim_{t \to \infty} \sup_{0 \leq h \leq T} \|Y(t+h), \phi_h(Y(t))\|^* = 0, \text{a.s.}.$$

4. Having characterized the relation between the OGD trajectory (affine interpolation of the discrete OGD iterates) and the ODE trajectory (the semi-flow), we now turn to studying the latter (the semiflow given by the ODE trajectory). A desirable property of $\phi_t(\mathbf{y})$ is that the distance $F^\lambda(\mathbf{x}^*, \phi_t(\mathbf{y}))$ between the primal variable $\mathbf{x}^*$ and the dual variable $\phi_t(\mathbf{y}))$ (as measured by the Lyapunov function λ-Fenchel coupling) can never increase as a function of t. We refer to this as the monotonicity property of λ-Fenchel coupling under the ODE trajectory, to be contrasted to the discrete-time dynamics, where such monotonicity is absent (even when perfect information on the gradient is available). More formally, we show that $\forall \mathbf{y}, \forall 0 \leq s \leq t$,

$$F^\lambda(\mathbf{x}^*, \phi_s(\mathbf{y})) \geq F^\lambda(\mathbf{x}^*, \phi_t(\mathbf{y})). \tag{15}$$

5. Continuing on the previous point, not only the distance $F^\lambda(\mathbf{x}^*, \phi_t(\mathbf{y}))$ can never increase as t increases, but also, provided that $\phi_t(\mathbf{y})$ is not too close to $\mathbf{x}^*$ (under the λ-Fenchel coupling divergence measure), $F^\lambda(\mathbf{x}^*, \phi_t(\mathbf{y}))$ will decrease no slower than linearly. This suggests that either $\phi_t(\mathbf{y})$ is already close to $\mathbf{x}^*$ (and hence $x(t) = C(\phi_t(\mathbf{y}))$ is close to $\mathbf{x}^*$), or their distance will be decreased by a meaningful amount in (at least) the ensuing short time-frame. We formalize this discussion into the following mathematical claim: $\forall \epsilon > 0, \forall \mathbf{y}, \exists s > 0$, such that:

$$F^\lambda(\mathbf{x}^*, \phi_s(\mathbf{y})) \leq \max\{\frac{\epsilon}{2}, F^\lambda(\mathbf{x}^*, \mathbf{y}) - \frac{\epsilon}{2}\}. \tag{16}$$

6. Now consider an arbitrary fixed horizon T. If at time t, $F^\lambda(\mathbf{x}^*, \phi_0(Y(t)))$ is small, then by the monotonicity property in Claim 4, $F^\lambda(\mathbf{x}^*, \phi_h(Y(t)))$ will remain small on the entire interval $h \in [0, T]$. Since $Y(t)$ is an asymptotic pseudotrajectory of ϕ (Claim 3), $Y(t+h)$ and $\phi_h(Y(t))$ should be very close for $h \in [0, T]$, at least for t large enough. This means that $F^\lambda(\mathbf{x}^*, Y(t+h))$ should also be small on the entire interval $h \in [0, T]$, if λ-Fenchel coupling has a regular enough structure. It turns out that this is indeed the case. This can be made precise as follows: $\forall \epsilon, T > 0, \exists \tau(\epsilon, T) > 0$ such that $\forall t \geq \tau, \forall h \in [0, T]$:

$$F^\lambda(\mathbf{x}^*, Y(t+h)) < F^\lambda(\mathbf{x}^*, \phi_h(Y(t))) + \frac{\epsilon}{2}, \text{ a.s.}. \tag{17}$$

7. Finally, we are ready to put the above pieces together. Claim 6 gives us a way to control the amount by which the two λ-Fenchel coupling functions differ on the interval $[0, T]$. Claim 4 and Claim 5 together allow us to extend such control over successive intervals $[T, 2T), [2T, 3T), \ldots$, thereby establishing that, at least for t large enough, if $F^\lambda(\mathbf{x}^*, Y(t))$ is small, then $F^\lambda(\mathbf{x}^*, Y(t+h))$ will remains small $\forall h > 0$. As it turns out, this means that after long enough time, if $\mathbf{X}^t$ ever visits $\tilde{B}(\mathbf{x}^*, \epsilon)$, it will (almost surely) be forever trapped inside the neighborhood twice that size (i.e. $\tilde{B}(\mathbf{x}^*, 2\epsilon)$). Since Claim 1 ensures that $\mathbf{X}^t$ visits $\tilde{B}(\mathbf{x}^*, \epsilon)$ infinitively often (almost surely), the hypothesis is guaranteed to be true. Consequently, this leads to the following formal claim: $\forall \epsilon > 0, \exists \tau_0$ (a positive integer), such that:

$$F^\lambda(\mathbf{x}^*, Y(\tau_0 + h)) < \epsilon, \forall h \in [0, \infty), \text{ a.s.}. \tag{18}$$

To conclude, Equation (18) in Claim 7 implies that $F^\lambda(\mathbf{x}^*, Y^t) \to 0$, a.s. as $t \to \infty$, where the OGD iterates Y^t are values at integer time points of the affine trajectory $Y(\tau)$. Per Statement 1 in Lemma 2, this leads to that $\|C(Y^t) - \mathbf{x}^*\| \to 0$, a.s. as $t \to \infty$, thereby establishing that $X^t = C(Y^t) \to \mathbf{x}^*$, a.s. as $t \to 0$.

5. Theoretical Guarantees of Dual Averaging Power Control: Deterministic and Time-Invariant Channels

In this section, we establish the theoretical guarantees of the two variants of the dual averaging power control algorithm in the deterministic and time-invariant channel case. Our approach lies in designing two λ-monotone games such that each variant of the dual averaging algorithm can be interpreted as a special instance of the OGD learning dynamics for the corresponding game. We emphasize that it is *not* the case that the transmitters are playing a repeated game where their utilities are prescribed by the designed function. Instead, this is merely used as an analytical framework to study the proposed algorithms. In fact, this is the analytical framework we use to design and derive the dual averaging power control algorithms in the first place.

5.1. Designed λ-Monotone Games

Under the notation introduced in Section 2, we consider the following two games $\mathcal{G}^1, \mathcal{G}^2$ as given below, where the set $\mathcal{N}$ of players is the set of links in the power control contexts:

1. $\mathcal{G}^1 = (\mathcal{N}, \mathcal{P}, \{u_i\}_{i=1}^N)$, where

$$u_i(\mathbf{p}) = -\frac{1}{2G_{ii}}(G_{ii}p_i - r_i^*(\sum_{j \neq i} G_{ij}p_j + \eta_i))^2. \tag{19}$$

2. $\mathcal{G}^2 = (\mathcal{N}, \mathcal{P}, \{\tilde{u}_i\}_{i=1}^N)$, where

$$\tilde{u}_i(\mathbf{p}) = -\frac{1}{2G_{ii}^2}(G_{ii}p_i - r_i^*(\sum_{j \neq i} G_{ij}p_j + \eta_i))^2. \tag{20}$$

We first point out an interesting feature shared by both of these games when the channel is feasible: per Fact 1, a (necessarily unique) optimal joint transmission power $\mathbf{p}^*$ exists. Since the optimal joint transmission power matches the quality-of-service constraints exactly, every player's

utility will be 0 if they transmit according to $\mathbf{p}^*$. This implies that $\mathbf{p}^*$ must be a Nash equilibrium, since 0 is the highest utility that can be possibly achieved for any given player (link). In fact, at $\mathbf{p}^*$, not only will any player fail to obtain better utility by unilaterally deviating from $\mathbf{p}^*$, the players cannot achieve better utility through collusion of any type. Furthermore, $\mathbf{p}^*$ is the unique Nash equilibrium, because if $\mathbf{p}$ were any other Nash equilibrium, then necessarily one player's utility is below 0. For this player, he will have the incentive to transmit at a higher power compared to the current prescribed transmission power so as to achieve better utility.

The preceding discussion essentially establishes that when the channel is feasible, both of these games admit a unique Nash equilibrium $\mathbf{p}^*$, which is the optimal joint transmission power. However, it still remains a question as to whether they are λ-monotone. The following lemma presents a rather intriguing result: the feasibility of the channel not only guarantees the existence of a unique Nash equilibrium, but also, more importantly, implies that both games are λ-monotone.

LEMMA 3. *Fix* $\lambda = (\frac{1}{G_{11}}, \frac{1}{G_{22}}, \ldots, \frac{1}{G_{NN}})$. *Assume that the channel* $(\mathbf{G}, \eta)$ *is feasible and thereby let* $\mathbf{p}^*$ *be the optimal joint transmission power as defined in Definition 3, where* $\mathbf{p}^*$ *is assumed to lie in* $\mathcal{P}$. *The the following statements hold:*

1. *The designed game* $\mathcal{G}^1 = (\mathcal{N}, \mathcal{P}, \{u_i\}_{i=1}^N)$ *is a* λ*-monotone game with the unique Nash equilibrium* $\mathbf{p}^*$.

2. *The designed game* $\mathcal{G}^2 = (\mathcal{N}, \mathcal{P}, \{\tilde{u}_i\}_{i=1}^N)$ *is a monotone game with the unique Nash equilibrium* $\mathbf{p}^*$.

Proof: We first establish the first statement. For each i:

$$v_i(\mathbf{p}) = \frac{\partial u_i(\mathbf{p})}{\partial p_i} = -(G_{ii}p_i - r_i^*(\sum_{j \neq i} G_{ij}p_j + \eta_i)), \tag{21}$$

which can be easily seen as affine in p_i and hence concave, with all the smoothness assumptions satisfied. For each i, j:

$$\frac{\partial v_i(\mathbf{p})}{\partial p_i} = -G_{ii}, \frac{\partial v_i(\mathbf{p})}{\partial p_j} = r_i^* G_{ij}.$$

Computing the λ-weighted Hessian matrix of the designed game, we obtain:

$$H_{ij}^{\lambda}(\mathbf{p}) = \frac{1}{2G_{ii}}\frac{\partial v_i(\mathbf{p})}{\partial p_j} + \frac{1}{2G_{jj}}\frac{\partial v_j(\mathbf{p})}{\partial p_i}.$$

If $i = j$, then

$$H_{ii}^{\lambda}(\mathbf{p}) = 2 \times \frac{1}{2G_{ii}}\frac{\partial v_i(\mathbf{p})}{\partial p_i} = -1.$$

If $i \neq j$, then

$$H_{ij}^{\lambda}(\mathbf{p}) = \frac{1}{2G_{ii}}r_i^* G_{ij} + \frac{1}{2G_{jj}}r_j^* G_{ji} = \frac{1}{2}(r_i^*\frac{G_{ij}}{G_{ii}} + r_j^*\frac{G_{ji}}{G_{jj}}).$$

Let W be the re-weighted gain matrix as defined in Equation 3 and $\mathbf{I} \in \mathbf{R}^{N \times N}$ be the identity matrix. From the previous calculations, it follows that

$$H^{\lambda} = \frac{1}{2}(W - \mathbf{I}) + \frac{1}{2}(W^T - \mathbf{I}).$$

Since the channel $(\mathbf{G}, \eta)$ is feasible, per Theorem 1, $\lambda_{\max}(W) < 1$. Consequently, $(W - \mathbf{I})$ is negative definite (although not necessarily symmetric, i.e. may have complex eigenvalues): $\forall \mathbf{p} \in \mathbf{R}^N$,

$$\mathbf{p}(W - \mathbf{I})\mathbf{p} = \mathbf{p}W\mathbf{p} - \|\mathbf{p}\|_2 \leq (\lambda_{\max}(W) - 1)\|\mathbf{p}\|_2 < 0.$$

Similarly, $(W^T - \mathbf{I})$ is negative definite, thereby implying $H_{ij}(\mathbf{p})$ is negative definite. Since $H_{ij}(\mathbf{p})$ is negative definite for every $\mathbf{p}$ (since it is independent of $\mathbf{p}$), Lemma 1 establishes that the game is λ-monotone.

Since $\mathbf{p}^*$ results the maximum utility for every player: $u_i(\mathbf{p}^*) = 0$, $\mathbf{p}^*$ must be a Nash equilibrium and hence the unique Nash equilibrium.

Statement 2 follows by a similar argument by noting that in this case,

$$v_i(\mathbf{p}) = \frac{\partial \tilde{u}_i(\mathbf{p})}{\partial p_i} = -(p_i - r_i^*\frac{\sum_{j \neq i} G_{ij}p_j + \eta_i}{G_{ii}}). \tag{22}$$

Computing the β-weighted Hessian for $\beta = (1, 1, \dots, 1)$ yields:

$$H^{\beta} = \frac{1}{2}(W - \mathbf{I}) + \frac{1}{2}(W^T - \mathbf{I}),$$

thereby establishing the conclusion. ∎

5.2. Convergence of Dual Averaging in the Power Control Problem

With the above two designed games, we can now connect the dual averaging power control algorithms to the OGD learning dynamics, as made precise by the following lemma.

LEMMA 4. *Variant A of dual averaging (Algorithm 2) and Variant B of dual averaging (Algorithm 3) are OGD (Algorithm 6) for the designed games $(\mathcal{N}, \mathcal{P}, \{u_i\}_{i=1}^N)$ and $(\mathcal{N}, \mathcal{P}, \{\tilde{u}_i\}_{i=1}^N)$ respectively, where*

1. $h_i(p_i) = \frac{1}{2}p_i^2$.

2. $\alpha^t = \frac{1}{t+1}$.

Proof: First, note that it is a standard result in convex analysis that $h_i(p_i) = \frac{1}{2}p_i^2$ is 1-strongly convex with respect to the Euclidean 2-norm and that the corresponding choice map $C_i(y_i)$ is simply the closest-point projection: $C_i(y_i) = \arg\max_{p_i \in \mathcal{P}_i} \{p_i y_i - \frac{1}{2}p_i^2\} = \Pi_{\mathcal{P}_i}(y_i)$. The lemma therefore follows by noting that $\frac{\partial u_i(\mathbf{p})}{\partial p_i} = -(G_{ii}p_i - r_i^*(\sum_{j \neq i} G_{ij}p_j + \eta_i))$ and $\frac{\partial \tilde{u}_i(\mathbf{p})}{\partial p_i} = -(p_i - r_i^* \frac{\sum_{j \neq i} G_{ij}p_j + \eta_i}{G_{ii}})$. ∎

Combining all the pieces together, we therefore have the following convergence and optimality guarantees for dual averaging power control:

THEOREM 6. *Let the channel $(\mathbf{G}, \eta)$ be feasible with the optimal joint transmission power $\mathbf{p}^* \in \mathcal{P}$.*

1. $\mathbf{p}^t \to \mathbf{p}^*$ *as* $t \to \infty$, *where* $\mathbf{p}^t$ *is given in Algorithm 2.*

2. $\mathbf{p}^t \to \mathbf{p}^*$ *as* $t \to \infty$, *where* $\mathbf{p}^t$ *is given in Algorithm 3.*

Proof: Per Lemma 4, Algorithm 2 is OGD for the game $(\mathcal{N}, \mathcal{P}, \{u_i\}_{i=1}^N)$ and Algorithm 3 is OGD for the game $(\mathcal{N}, \mathcal{P}, \{\tilde{u}_i\}_{i=1}^N)$, where $\alpha^t = \frac{1}{t+1}$ and $h_i(p_i) = \frac{1}{2}p_i^2$. Note that the sequence $\{\frac{1}{t+1}\}_{t=0}^\infty$ is clearly slowly vanishing. The two results follow immediately from Theorem 4. ∎

6. Theoretical Guarantees of Dual Averaging Power Control: Stochastic and Time-Varying Channels

In this section, we establish the theoretical guarantees of the two variants of the dual averaging power control algorithm in the stochastic and time-varying channel case. Our approach again lies in designing two λ-monotone games, which build on the ones presented in the previous section.

6.1. Designed Concave Games

Note that in a stochastic and time-varying channel, $(\mathbf{G}, \eta)$ (or (W, γ)) is random. We consider the following two games $\mathcal{G}^1, \mathcal{G}^2$ as given below, where the set $\mathcal{N}$ of players is again the set of wireless links in the power control contexts:

1. $\mathcal{G}^1 = (\mathcal{N}, \mathcal{P}, \{u_i\}_{i=1}^N)$, $u_i(\mathbf{p}) = \mathbf{E}[-\frac{1}{2G_{ii}}(G_{ii}p_i - r_i^*(\sum_{j \neq i} G_{ij}p_j + \eta_i))^2]$.

2. $\mathcal{G}^2 = (\mathcal{N}, \mathcal{P}, \{\tilde{u}_i\}_{i=1}^N)$, $\tilde{u}_i(\mathbf{p}) = \mathbf{E}[-\frac{1}{2G_{ii}^2}(G_{ii}p_i - r_i^*(\sum_{j \neq i} G_{ij}p_j + \eta_i))^2]$.

Under the above two designed games, it is straightforward to verify that

$$\frac{\partial u_i(\mathbf{p})}{\partial p_i} = -\mathbf{E}[(G_{ii}p_i - r_i^*(\sum_{j \neq i} G_{ij}p_j + \eta_i))] = -(\mathbf{E}[G_{ii}]p_i - r_i^*(\sum_{j \neq i} \mathbf{E}[G_{ij}]p_j + \mathbf{E}[\eta_i])), \quad (23)$$

$$\frac{\partial \tilde{u}_i(\mathbf{p})}{\partial p_i} = -\mathbf{E}[(p_i - r_i^* \frac{\sum_{j \neq i} G_{ij}p_j + \eta_i}{G_{ii}})] = -(p_i - r_i^* \sum_{j \neq i}(\mathbf{E}[\frac{G_{ij}}{G_{ii}}]p_j + \mathbf{E}[\frac{\eta_i}{G_{ii}}])) = -(p_i - \sum_{j \neq i}(\mathbf{E}[W_{ij}]p_j + \mathbf{E}[\gamma_i])).$$
$$(24)$$

Before, we move on to establish the theoretical guarantees for dual averaging in the stochastic and time-varying channel case, we need a notion that characterizes the channel feasibility in this case. Since the channel is fluctuating, it would be too strong to require that each channel realization on any given time step is feasible. Instead, here we only impose the mild requirement that a channel is mean feasible: the channel can be feasible sometimes and infeasible some other times as long as it is feasible on average. The next definition formalizes it:

DEFINITION 11. A channel $(\mathbf{G}, \eta)$ (or equivalently (W, γ)) is:

1. Type-I mean-feasible if $(\mathbf{E}[\mathbf{G}], \mathbf{E}[\eta])$ is feasible, where expectation is taken component-wise.

2. Type-II mean-feasible if $(\mathbf{E}[W], \mathbf{E}[\gamma])$ is feasible, where expectation is taken component-wise.

The two types of mean-feasible channels are closely related: although in general neither implies the other, in the important special case that G_{ij}'s are independent of G_{ii}, Type-I mean-feasible is weaker than Type-II mean-feasible, as formalized by the following lemma:

LEMMA 5. *If for each $i \in \{1, 2, \ldots, N\}$, G_{ij} and G_{ii} are pairwise independent for each $j \neq i$, then a channel that is Type-II mean-feasible is Type-I mean-feasible.*

Proof: Let a channel $(\mathbf{G}, \eta)$ be Type-II mean-feasible. Defined the matrix $\tilde{W}$ to be:

$$\tilde{W}_{ij} := \begin{cases} 0, & i = j \\ r_i^* \dfrac{\mathbf{E}[G_{ij}]}{\mathbf{E}[G_{ii}]}, & i \neq j. \end{cases} \tag{25}$$

We have:

$$\mathbf{E}\left[\frac{G_{ij}}{G_{ii}}\right] = \mathbf{E}[G_{ij}]\mathbf{E}\left[\frac{1}{G_{ii}}\right] \geq \frac{\mathbf{E}[G_{ij}]}{\mathbf{E}[G_{ii}]}, \tag{26}$$

where the equality follows from independence and the inequality follows from Jensen's inequality and the fact that $f(x) = \frac{1}{x}$ is convex when x is positive. This immediately implies:

$$\mathbf{E}[W] \geq \tilde{W},$$

where inequality holds component-wise. Since each entry in both matrices is non-negative, we have:

$$\lambda_{\max}(\mathbf{E}[W]) = \max_{\mathbf{u} \in \mathbf{R}^N : \|\mathbf{u}\|_2 = 1} \mathbf{u}\mathbf{E}[W]\mathbf{u} = \max_{\mathbf{u} \in \mathbf{R}_+^N : \|\mathbf{u}\|_2 = 1} \mathbf{u}\mathbf{E}[W]\mathbf{u},$$

$$\lambda_{\max}(\tilde{W}) = \max_{\mathbf{u} \in \mathbf{R}_+^N : \|\mathbf{u}\|_2 = 1} \mathbf{u}\tilde{W}\mathbf{u} = \max_{\mathbf{u} \in \mathbf{R}_+^N : \|\mathbf{u}\|_2 = 1} \mathbf{u}\tilde{W}\mathbf{u}.$$

Further, for each $\mathbf{u} \in \mathbf{R}_+^N$, Equation (26) implies $\mathbf{u}\mathbf{E}[W]\mathbf{u} \geq \mathbf{u}\tilde{W}\mathbf{u}$, which then leads to

$$\lambda_{\max}(\mathbf{E}[W]) = \max_{\mathbf{u} \in \mathbf{R}_+^N : \|\mathbf{u}\|_2 = 1} \mathbf{u}\mathbf{E}[W]\mathbf{u} \geq \max_{\mathbf{u} \in \mathbf{R}_+^N : \|\mathbf{u}\|_2 = 1} \mathbf{u}\tilde{W}\mathbf{u} = \lambda_{\max}(\tilde{W}).$$

By Theorem 1, $\lambda_{\max}(\tilde{W}) < 1$ and hence the channel must be Type-I mean-feasible. ∎

6.2. Almost Sure Convergence of Dual Averaging Power Control: Stability and Optimality

We are now ready for the main result for the stochastic and time-varying channel case. We obtain this result by casting the dual averaging power control algorithms in the OGD learning framework.

THEOREM 7. *Given a stochastic and time-varying channel $(\mathbf{G}, \eta)$ (or equivalently (W, γ)) according to Assumption 1.*

1. *If the channel is Type-I mean feasible, with $\mathbf{p}^* \in \mathcal{P}$ being the optimal joint transmission power for $(\mathbf{E}[\mathbf{G}], \mathbf{E}[\eta])$, then $\mathbf{P}^t \to \mathbf{p}^*$ almost surely, as $t \to \infty$, where $\mathbf{P}^t$ is given in Algorithm 4.*

2. *If the channel is Type-II mean feasible, with* $\mathbf{p}^* \in \mathcal{P}$ *being the optimal joint transmission power for* $(\mathbf{E}[W], \mathbf{E}[\gamma])$ *and for each* i, *there exists some* $\underline{g}_i > 0$ *such that* $G_{ii} \geq \underline{g}_i$ *a.s., then* $\mathbf{P}^t \to \mathbf{p}^*$ *almost surely, as* $t \to \infty$, *where* $\mathbf{P}^t$ *is given in Algorithm 5.*

REMARK 8. Two things to note here. First, the almost sure convergence to a constant joint transmission power in the presence of persistent stochastic channel fluctuations is a manifestation of the stability of dual averaging. The intuition behind this stability is that as past powers are incorporated into the current power via a weighted sum, the random environments have less and and less impact on the current power iterate because the step-sizes are decreasing. Second, $G_{ii} \geq \underline{g}_i$ a.s. is a rather mild assumption as it means that power gain between each transmitter and its intended receiver is lower bounded by some positive constant.

Proof: For the first claim, per Assumption 1, we can write $G_{ij}^t = \mathbf{E}[G_{ij}] + \tilde{G}_{ij}^t, \eta_i^t = \mathbf{E}[G_{ij}] + \tilde{\eta}_i^t$, where $\tilde{G}_{ij}^t$ and $\tilde{\eta}_i^t$ are both sequences of **iid**, zero-mean and finite-variance random variables. Consequently, the gradient update (Line 5) in Algorithm 4 can be equivalently written as:

$$Y_i^{t+1} = Y_i^t - \frac{1}{t}\left\{ (\mathbf{E}[G_{ii}] + \tilde{G}_{ii}^t)P_i^t - r_i^*(\sum_{j \neq i}(\mathbf{E}[G_{ij}] + \tilde{G}_{ij}^t)P_j^t + \mathbf{E}[\eta_i] + \tilde{\eta}_i^t) \right\} \tag{27}$$

$$= Y_i^t - \frac{1}{t}\left\{ \mathbf{E}[G_{ii}]P_i^t - r_i^*(\sum_{j \neq i}\mathbf{E}[G_{ij}]P_j^t) + \mathbf{E}[\eta_i] + \{\tilde{G}_{ii}^t P_i^t - r_i^*\sum_{j \neq i}\tilde{G}_{ij}^t P_j^t + \tilde{\eta}_i^t\} \right\}. \tag{28}$$

Denoting $\xi_i^{t+1} = \tilde{G}_{ii}^t P_i^t - r_i^*\sum_{j \neq i}\tilde{G}_{ij}^t P_j^t + \tilde{\eta}_i^t$, it follows that $\mathbf{E}[\xi_i^{t+1} \mid P^0, \ldots, P^t] = 0$, a.s. and $\mathbf{Var}[\xi_i^{t+1} \mid P^0, \ldots, P^t] < \infty$, a.s.. Since $\mathcal{P}$ is bounded, there exists a constant $B > 0$ such that $\mathbf{Var}[\xi_i^{t+1} \mid P^0, \ldots, P^t] \leq B$, a.s., $\forall t$. Consequently, the martingale noise ξ^t satisfies Assumption 2 per Remark 2. This implies that Algorithm 4 is a special case of Algorithm 7. Further, since the channel is Type-I mean feasible, Lemma 3 implies that $\mathcal{G}^1$ is $(\frac{1}{\mathbf{E}[G_{11}]}, \ldots, \frac{1}{\mathbf{E}[G_{NN}]})$-monotone with $\mathbf{p}^*$ being the unique Nash equilibrium. It therefore suffices to verify the conditions given in Theorem 5 are satisfied. Condition 1 is just verified above as ξ^t satisfies Assumption 2. Condition 2 holds $v_i(\mathbf{p})$ is still linear under expectation. Condition 3 was already verified in the proof to Theorem 6. The result therefore follows.

For the second claim, using the same line of reasoning, it follows that Algorithm 5 is a special case of Algorithm 7. The only thing to note here is that the bounded second moments assumption holds because $G_{ii} \geq \underline{g}_i$ a.s..

■

7. Conclusion

We close with two remarks. First, from a technical standpoint, our result for the almost sure convergence of online gradient descent (OGD) to Nash equilibrium in λ-monotone games was established via a multi-step argument that can be best described in two parts as follows. The first part is to establish the algorithm's recurrence: the game's (necessarily unique) Nash equilibrium is an accumulation point of the sequence of play generated by OGD. The second part is to use a mean-field approximation to push this recurrence result one step further and show that the algorithm does not admit any other accumulation points. Specifically, we do this by *a*) converting the stochastic discrete process into a deterministic continuous process; *b*) establishing the convergence of the resulting continuous trajectories; and, finally, *c*) using a stochastic approximation argument to map this "mean field" convergence back to the discrete process. Even though this general program differs on a case-by-case basis (our case is particularly involved due to the non-invertibility Euclidean projections), the underlying philosophy above has been successfully used in other important applications in operations research, the most notable of which is in queueing networks (see Gamarnik (2010) for a landmark survey).

Second, even though our focus in this paper was on power minimization subject to rate constraints, the general "mean field" framework above can be applied to a broad spectrum of other power management (and/or more general game-theoretic) problems. We intend to explore these directions in future work.

Endnotes

1. The literature on power control is too broad to review here; for a comprehensive survey, we refer the reader to Chiang et al. (2008).

2. Feasibility means here that there exists a power allocation vector under which every user's throughput requirements are satisfied.

3. In our case, the algorithm should be *ascending* the payoff to be maximized, but we use the more standard "descent" terminology to be consistent with the literature.

4. Specifically, a Nash equilibrium profile is a correlated equilibrium, which in turn is in the Hannan set; in general, all these inclusions are proper.

5. As noted in Tan (2014), the re-weighted gain matrix W is a non-negative (and without loss of generality, irreducible) matrix; thus by Perron-Frobenius theorem, there is a unique positive real eigenvalue ρ^* that has the largest magnitude.

6. To be precise, the version of OGD considered here is called lazy OGD. There is also a variant called eager OGD. Shalev-Shwartz et al. 2012 provides a nice tutorial on the (minor) differences. Note that the no-regret property holds for both variants; further, they have the same regret bound.

7. Strictly speaking, Theorem 5 has an additional assumption that v be Lipschitz continuous. It turns out that even when this assumption is violated, the result here (i.e. perfect gradient case) still holds, although it requires a different argument (i.e. a non-ODE approach). Due to space limitation and due to that in the application considered here v is indeed Lipschitz, we omit in the current draft this argument.

8. A crucial point to note is that since C is not invertible, there may not exist a unique solution for $\mathbf{x}(t)$.

References

Alpcan, Tansu, Tamer Basar, Subhrakanti Dey. 2006. A power control game based on outage probabilities for multicell wireless data networks. *IEEE transactions on wireless communications* **5**(4) 890–899.

Alpcan, Tansu, Tamer Başar, Rayadurgam Srikant, Eitan Altman. 2002. Cdma uplink power control as a noncooperative game. *Wireless Networks* **8**(6) 659–670.

Altman, Eitan, Tamer Başar, Francesco De Pellegrini. 2010. Optimal monotone forwarding policies in delay tolerant mobile ad-hoc networks. *Performance Evaluation* **67**(4) 299–317.

Au-Yeung, Kit Yee, Timothy Robertson, Hooman Hafezi, Gregory Moon, Lorenzo DiCarlo, Mark Zdeblick, George Savage. 2010. A networked system for self-management of drug therapy and wellness. *ACM Wireless Health*.

Balandat, Maximilian, Walid Krichene, Claire Tomlin, Alexandre Bayen. 2016. Minimizing regret on reflexive banach spaces and learning nash equilibria in continuous zero-sum games. *arXiv preprint arXiv:1606.01261* .

Bambos, Nicholas, Shou C Chen, Gregory J Pottie. 2000. Channel access algorithms with active link protection for wireless communication networks with power control. *Networking, IEEE/ACM Transactions on* **8**(5) 583–597.

Benaïm, Michel. 1999. Dynamics of stochastic approximation algorithms. *Seminaire de probabilites XXXIII*. Springer, 1–68.

BenaÏm, Michel, Morris W. Hirsch. 1996. Asymptotic pseudotrajectories and chain recurrent flows, with applications. *Journal of Dynamics and Differential Equations* **8**(1) 141–176. doi:10.1007/BF02218617. URL `http://dx.doi.org/10.1007/BF02218617`.

Bradley, Joseph, Joel Barbier, Doug Handler. 2013. Embracing the internet of everything to capture your share of \$14.4 trillion. *CISCO White Paper* .

Byrne, Christopher, Chin Leong Lim. 2007. The ingestible telemetric body core temperature sensor: a review of validity and exercise applications. *British journal of sports medicine* **41**(3) 126–133.

Candogan, Utku Ozan, Ishai Menache, Asuman Ozdaglar, Pablo A Parrilo. 2010. Near-optimal power control in wireless networks: A potential game approach. *INFOCOM, 2010 Proceedings IEEE*. IEEE, 1–9.

Cesa-Bianchi, Nicolo, Gábor Lugosi. 2006. *Prediction, learning, and games*. Cambridge university press.

Chiang, Mung, Prashanth Hande, Tian Lan, Chee Wei Tan. 2008. Power control in wireless cellular networks. *Foundations and Trends® in Networking* **2**(4) 381–533.

Coddington, A., N. Levinson. 1955. *Theory of Ordinary Differential Equations*. International series in pure and applied mathematics, R.E. Krieger. URL `https://books.google.com/books?id=AUAbvgAACAAJ`.

Deakin, M. 2013. *Smart Cities: Governing, Modelling and Analysing the Transition*. Taylor & Francis. URL `https://books.google.com/books?id=8QWCAAAAQBAJ`.

El Gamal, Abbas, James Mammen, Balaji Prabhakar, Devavrat Shah. 2006a. Optimal throughput-delay scaling in wireless networks-part i: The fluid model. *Information Theory, IEEE Transactions on* **52**(6) 2568–2592.

El Gamal, Abbas, James Mammen, Balaji Prabhakar, Devavrat Shah. 2006b. Optimal throughput-delay scaling in wireless networks-part i: The fluid model. *Information Theory, IEEE Transactions on* **52**(6) 2568–2592.

Eryilmaz, Atilla, Eytan Modiano, Asuman Ozdaglar. 2006. Randomized algorithms for throughput-optimality and fairness in wireless networks. *Decision and Control, 2006 45th IEEE Conference on*. IEEE, 1936–1941.

Eryilmaz, Atilla, Asuman Ozdaglar, Muriel Médard, Ebad Ahmed. 2008. On the delay and throughput gains of coding in unreliable networks. *Information Theory, IEEE Transactions on* **54**(12) 5511–5524.

Famolari, David, Narayan Mandayam, David Goodman, Viral Shah. 1999. A new framework for power control in wireless data networks: Games, utility, and pricing.

Fan, Xingzhe, Tansu Alpcan, Murat Arcak, TJ Wen, T Başar. 2006. A passivity approach to game-theoretic cdma power control. *Automatica* **42**(11) 1837–1847.

Foschini, G.J., Z. Miljanic. 1993. A simple distributed autonomous power control algorithm and its convergence. *Vehicular Technology, IEEE Transactions on* **42**(4) 641–646. doi:10.1109/25.260747.

Gamarnik, David. 2010. Fluid models of queueing networks. *Wiley Encyclopedia of Operations Research and Management Science* .

Gitzenis, Savvas, Nicholas Bambos. 2002. Power-controlled data prefetching/caching in wireless packet networks. *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3. IEEE, 1405–1414.

Goldsmith, Andrea. 2005. *Wireless communications*. Cambridge university press.

Gopalan, Aditya, Constantine Caramanis, Sanjay Shakkottai. 2015. Wireless scheduling with partial channel state information: large deviations and optimality. *Queueing Systems* **80**(4) 293–340.

Hall, P., C.C. Heyde. 1980. *Martingale limit theory and its application*. Probability and mathematical statistics, Academic Press. URL `https://books.google.com/books?id=xxbvAAAAMAAJ`.

Han, Z., D. Niyato, W. Saad, T.B. Ar, A.H. Rungnes. 2014. *Game Theory in Wireless and Communication Networks*. Cambridge University Press.

Han, Z., D. Niyato, W. Saad, T. Başar, A. Hjørungnes. 2011. *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*. Cambridge University Press. URL `https://books.google.com/books?id=mvaUAwAAQBAJ`.

Krichene, Syrine, Walid Krichene, Roy Dong, Alexandre Bayen. 2015. Convergence of heterogeneous distributed learning in stochastic routing games. *Communication, Control, and Computing (Allerton), 2015 53rd Annual Allerton Conference on*. IEEE, 480–487.

Lam, Kiet, Walid Krichene, Alexandre Bayen. 2016. On learning how players learn: estimation of learning dynamics in the routing game. *Cyber-Physical Systems (ICCPS), 2016 ACM/IEEE 7th International Conference on*. IEEE, 1–10.

Li, Na, Jason R Marden. 2013. Designing games for distributed optimization. *IEEE Journal of Selected Topics in Signal Processing* **7**(2) 230–242.

Menache, I., A. Ozdaglar. 2010. *Network Games: Theory, Models, and Dynamics*. Network Games: Theory, Models, and Dynamics, Morgan & Claypool.

Mitra, Debasis. 1994. An asynchronous distributed algorithm for power control in cellular radio systems. *Wireless and Mobile Communications*. Springer, 177–186.

Rappaport, Theodore. 2001. *Wireless Communications: Principles and Practice*. 2nd ed. Prentice Hall PTR, Upper Saddle River, NJ, USA.

Reddy, Akula Aneesh, Siddhartha Banerjee, Aditya Gopalan, Sanjay Shakkottai, Lei Ying. 2012. On distributed scheduling with heterogeneously delayed network-state information. *Queueing Systems* **72**(3-4) 193–218.

Reddy, Aneesh, Sanjay Shakkottai, Lei Ying. 2008. Distributed power control in wireless ad hoc networks using message passing: Throughput optimality and network utility maximization. *Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on*. IEEE, 770–775.

Rockafellar, R Tyrrell, Roger J-B Wets. 2009. *Variational analysis*, vol. 317. Springer Science & Business Media.

Rockafellar, Ralph Tyrell. 2015. *Convex analysis*. Princeton university press.

Rosen, J Ben. 1965a. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica: Journal of the Econometric Society* 520–534.

Rosen, J Ben. 1965b. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica: Journal of the Econometric Society* 520–534.

Seferoglu, H., A. Lakshmikantha, A. Ganesh, P. Key. 2008. Dynamic decentralized multi-channel mac protocols. *Information Theory and Applications Workshop, 2008*. 100–110. doi:10.1109/ITA.2008.4601031.

Shalev-Shwartz, Shai, et al. 2012. Online learning and online convex optimization. *Foundations and Trends®* *in Machine Learning* **4**(2) 107–194.

Tan, Chee Wei. 2014. Wireless network optimization by perron-frobenius theory. *Information Sciences and Systems (CISS), 2014 48th Annual Conference on*. IEEE, 1–6.

Tan, Chee Wei. 2015. Wireless network optimization by perron-frobenius theory. *Foundations and Trends in Networking* **9**(2-3) 107–218. doi:10.1561/1300000048. URL `http://dx.doi.org/10.1561/1300000048`.

Ulukus, Sennur, Roy D Yates. 1998. Stochastic power control for cellular radio systems. *Communications, IEEE Transactions on* **46**(6) 784–798.

Weeraddana, Pradeep Chathuranga, Marian Codreanu, Matti Latva-aho, Anthony Ephremides, Carlo Fischione, et al. 2012. Weighted sum-rate maximization in wireless networks: A review. *Foundations and Trends® in Networking* **6**(1–2) 1–163.

Yates, Roy D. 1996. A framework for uplink power control in cellular radio systems. *IEEE Journal on Selected Areas in Communications* **13** 1341–1347.

Zhou, Zhengyuan, Daniel Miller, Peter Glynn, Nicholas Bambos. 2016. A stochastic stability characterization of the foschini-miljanic algorithm in random wireless networks. *2016 IEEE Global Communications Conference (GLOBECOM)*. IEEE.

Zinkevich, Martin. 2003. Online convex programming and generalized infinitesimal gradient ascent. *ICML '03: Proceedings of the 20th International Conference on Machine Learning*. 928–936.

Appendix. Auxiliary Results and Missing Proofs

A. Martingale Convergence Theorems

We present two martingale convergence theorems that shall be useful later. These are adapted statements that come from Hall and Heyde (1980), which contains detailed proofs. The first one is a law of large number theorem for Martingales.

THEOREM 8. *Let $S^t = \sum_{k=0}^{t} X^k$ be a Martingale adapted to the filtration $\mathcal{S}^t$. Let $\{u^t\}_{t=0}^{\infty}$ be a nondecreasing sequence of positive numbers with $\lim_{t\to\infty} u^t = \infty$. If $\exists p \in [1,2]$ such that $\sum_{t=0}^{\infty} \frac{\mathbf{E}[|X^{t+1}|^p|\mathcal{S}^t]}{(u^t)^p} < \infty$, a.s. , then*

$$\lim_{t\to\infty} \frac{S^t}{u^t} = 0, \ a.s.,$$

The second one is Doob's Martingale convergence theorem.

THEOREM 9. *Let S^t be a submartingale adapted to the filtration $\mathcal{S}^t$, where $t = 0, 1, 2, \ldots$. If S^t is l^1-bounded: $\sup_{t\geq 0} \mathbf{E}[|S^t|] < \infty$, then S^t converges almost surely to a random variable S with $\mathbf{E}[|S|] < \infty$.*

B. Proof of Lemma 2

For the first statement, note that by a well-known result in convex analysis (see Rockafellar (2015)) when $x_i^* = C_i(y_i)$, it holds that y_i is a subgradient of h_i at the point x_i^*: $h_i(x_i) \geq h_i(x_i^*) + y_i(x_i - x_i^*)$. This then implies that, for any $\beta \in (0, 1]$:

$$\lambda_i(h_i(x_i^*) + y_i(\beta(x_i - x_i^*))) \leq \lambda_i h_i(x_i^* + \beta(x_i - x_i^*)) = \lambda_i h_i((1-\beta)x_i^* + \beta x_i) \leq \tag{29}$$

$$\lambda_i\{(1-\beta)h_i(x_i^*) + \beta h_i(x_i) - \frac{1}{2}\beta(1-\beta)\|x_i - x_i^*\|_i^2\}, \tag{30}$$

where the last inequality follows from the fact that h_i is 1-strongly convex.

Connecting the first and the last term of the above inequality chain, we have:

$$\lambda_i y_i \beta(x_i - x_i^*) \leq \lambda_i(-\beta)h_i(x_i^*) + \lambda_i \beta h_i(x_i) - \frac{1}{2}\lambda_i \beta(1-\beta)\|x_i - x_i^*\|_i^2.$$

Dividing both sides by β (since $\beta > 0$) and rearranging, we obtain:

$$\frac{1}{2}\lambda_i(1-\beta)\|x_i - x_i^*\|_i^2 \leq \lambda_i h_i(x_i) - \lambda_i h_i(x_i^*) - \lambda_i y_i(x_i - x_i^*).$$

Taking the limit that β approaches 0 (from above) results: $\lambda_i h_i(x_i) - \lambda_i h_i(x_i^*) - \lambda_i y_i(x_i - x_i^*) \geq \frac{1}{2}\lambda_i\|x_i - x_i^*\|_i^2$. Summing over all i's, we then obtain:

$$\sum_{i=1}^{N}\{\lambda_i h_i(x_i) - \lambda_i h_i(x_i^*) - \lambda_i y_i(x_i - x_i^*)\} \geq \sum_{i=1}^{N} \frac{1}{2}\lambda_i\|x_i - x_i^*\|_i^2.$$

The conclusion then follows by noting that:

$$F^\lambda(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{N} \lambda_i(h_i(x_i) - x_i y_i + h_i^*(y_i)) = \sum_{i=1}^{N} \lambda_i(h_i(x_i) - x_i y_i + x_i^* y_i - h_i(x_i^*)) \tag{31}$$

$$\geq \sum_{i=1}^{N} \frac{1}{2}\lambda_i\|x_i - x_i^*\|_i^2 \geq \frac{1}{2}(\min_i \lambda_i)\sum_{i=1}^{N}\|C_i(y_i) - x_i\|_i^2, \tag{32}$$

where the last inequality follows by noting that $x_i^* = C_i(y_i)$.

For the second statement, we start by citing here a useful result in Rockafellar and Wets (2009) (Theorem 12.60): For a proper, lower semi-continuous and convex function $f : \mathbf{R}^n \to \bar{\mathbf{R}}$ and a value $\sigma > 0$, f^* is σ-strongly convex (with respect to norm $\|\cdot\|^*$) if and only if f is differentiable and satisfies:

$$f(\tilde{\mathbf{x}}) \leq f(\mathbf{x}) + <\nabla f(\mathbf{x}), \tilde{\mathbf{x}} - \mathbf{x}> + \frac{1}{2\sigma}\|\tilde{\mathbf{x}} - \mathbf{x}\|^2, \forall \mathbf{x}, \tilde{\mathbf{x}},$$

where $\bar{\mathbf{R}} = [-\infty, \infty]$. Note that in the original statement, only the Euclidean norm $\|\cdot\|_2$ is used (Definition 12.58 in Rockafellar and Wets (2009) defined strong convexity implicitly in terms of the Euclidean norm), in which case $\|\cdot\|_2^* = \|\cdot\|_2$. However, as stated here, the same result holds true for any pair of norms $(\|\cdot\|, \|\cdot\|^*)$ by a straightforward adaptation of their proof.

Next, we note that in our case, each h_i is 1-strongly convex with respect to norm $\|\cdot\|_i$ and per Remark 3, $(h_i^*)^* = h_i$. Further, it can be easily checked that h_i^* is proper, lower semi-continuous and convex (since it is a point-wise maximum of affine functions per its definition), it therefore follows that the 1-strong convexity of $(h_i^*)^*$ (with respect to $\|\cdot\|_i^{**} = \|\cdot\|_i$) implies that h_i^* is differentiable and satisfies:

$$h_i^*(\tilde{y}_i) \leq h_i^*(y_i) + (h_i^*)'(y_i)(\tilde{y}_i - y_i) + \frac{1}{2}(\|\tilde{y}_i - y_i\|_i^*)^2, \forall y_i, \tilde{y}_i \tag{33}$$

$$= h_i^*(y_i) + C_i(y_i)(\tilde{y}_i - y_i) + \frac{1}{2}(\|\tilde{y}_i - y_i\|_i^*)^2, \forall y_i, \tilde{y}_i \tag{34}$$

where the equality follows because $(h_i^*)'(y_i) = C_i(y_i)$.

Therefore, it then follows that upon substituting the preceding inequality for each $h_i^*(\tilde{y}_i)$ into $F^\lambda(\mathbf{x}, \tilde{\mathbf{y}}) = \sum_{i=1}^{N} \lambda_i(h_i(x_i) - x_i\tilde{y}_i + h_i^*(\tilde{y}_i))$, we have:

$$F^\lambda(\mathbf{x}, \tilde{\mathbf{y}}) \leq \sum_{i=1}^{N} \lambda_i(h_i(x_i) - x_i\tilde{y}_i) + \sum_{i=1}^{N} \lambda_i\{h_i^*(y_i) + C_i(y_i)(\tilde{y}_i - y_i) + \frac{\lambda_i}{2}(\|\tilde{y}_i - y_i\|_i^*)^2\} \tag{35}$$

$$= \sum_{i=1}^{N} \lambda_i\{h_i(x_i) + h_i^*(y_i) - x_i y_i + x_i(y_i - \tilde{y}_i) + C_i(y_i)(\tilde{y}_i - y_i))\} + \sum_{i=1}^{N} \frac{\lambda_i}{2}(\|\tilde{y}_i - y_i\|_i^*)^2 \tag{36}$$

$$= F^\lambda(\mathbf{x}, \mathbf{y}) + \sum_{i=1}^{N} \lambda_i(\tilde{y}_i - y_i)(C_i(y_i) - x_i) + \sum_{i=1}^{N} \frac{\lambda_i}{2}(\|\tilde{y}_i - y_i\|_i^*)^2 \tag{37}$$

$$\leq F^\lambda(\mathbf{x}, \mathbf{y}) + \sum_{i=1}^{N} \lambda_i(\tilde{y}_i - y_i)(C_i(y_i) - x_i) + \frac{1}{2}(\max_i \lambda_i)\sum_{i=1}^{N}(\|\tilde{y}_i - y_i\|_i^*)^2. \tag{38}$$

∎

C. Proof of Theorem 5

We prove in turn each of the 7 claims laid out in Remark 7.

1. Let $Y^t = (Y_1^t, \ldots, Y_N^t), X^t = (X_1^t, \ldots, X_N^t)$ be the iterates generated in Algorithm 7. Fix an arbitrary $\epsilon > 0$. Assume for contradiction purposes that X^t only visits $B(\mathbf{x}^*, \epsilon)$ a finite number of times and hence let $t^0 - 1$ be the last time X^t is in $B(\mathbf{x}^*, \epsilon)$. Set $\beta^t = \frac{1}{2}(\alpha^t)^2 \max_i \lambda_i$.

   Using the representation $\tilde{v}(X^t) = v(X^t) + \xi^{t+1}$ given in Remark 2, we have that $\forall t \geq t^0$:

$$F^\lambda(\mathbf{x}^*, Y^{t+1}) = F^\lambda(\mathbf{x}^*, Y^t + \alpha^t \tilde{v}(X^t)) = F^\lambda(\mathbf{x}^*, Y^t + \alpha^t(v(X^t) + \xi^{t+1})) \tag{39}$$

$$\leq F^\lambda(\mathbf{x}^*, Y^t) + \sum_{i=1}^N \lambda_i(\alpha^t(v_i(X^t) + \xi_i^{t+1})(C_i(Y_i^t) - x_i^*) + \beta^t(\|\tilde{v}(X^t)\|^*)^2 \tag{40}$$

$$= F^\lambda(\mathbf{x}^*, Y^t) + \alpha^t \sum_{i=1}^N \lambda_i v_i(X^t)(X_i^t - x_i^*) + \alpha^t \sum_{i=1}^N \lambda_i \xi_i^{t+1}(X_i^t - x_i^*) + \beta^t(\|\tilde{v}(X^t)\|^*)^2 \tag{41}$$

$$\leq F^\lambda(\mathbf{x}^*, Y^t) + \alpha^t b_{\max}(\epsilon) + \alpha^t \sum_{i=1}^N \lambda_i \xi_i^{t+1}(X_i^t - x_i^*) + \beta^t(\|\tilde{v}(X^t)\|^*)^2 \tag{42}$$

$$\leq F^\lambda(\mathbf{x}^*, Y^{t^0}) + (\sum_{k=t^0}^t \alpha^k) b_{\max}(\epsilon) + \sum_{i=1}^N (\lambda_i \sum_{k=t^0}^t \alpha^k \xi_i^{k+1}(X_i^k - x_i^*)) + \sum_{k=t^0}^t \beta^k(\|\tilde{v}(X^k)\|^*)^2 \tag{43}$$

$$= F^\lambda(\mathbf{x}^*, Y^{t^0}) + (\sum_{k=t^0}^t \alpha^k) \left\{ b_{\max}(\epsilon) + \sum_{i=1}^N (\lambda_i \sum_{k=t^0}^t \frac{\alpha^k}{\sum_{k=t^0}^t \alpha^k} \xi_i^{k+1}(X_i^k - x_i^*)) \right\} + \sum_{k=t^0}^t \beta^k(\|\tilde{v}(X^k)\|^* \tag{44}$$

   where the last inequality follows from telescoping from $k = t^0$ to $k = t$.

   Next, note that $\forall i = 1, \ldots, N$, $\sum_{k=t^0}^t \alpha^k \xi_i^{k+1}(X_i^k - x_i^*)$ is a Martingale adapted to $\mathcal{F}^{t+1}$ because

$$\mathbf{E}[\alpha^k \xi_i^{t+1}(X_i^t - x_i^*)|\mathcal{F}^t] = \alpha^k(X_i^t - x_i^*)\mathbf{E}[\xi_i^{t+1}|\mathcal{F}^t] = 0.$$

Furthermore, setting $p = 2$ and $u^t = \sum_{k=t^0}^t \alpha^k$ (the first t^0 terms of u^t can be set arbitrarily and are not essential), it is clear that u^t is increasing and $\lim_{t \to \infty} u^t = \infty$ by assumption. It then follows that

$$\sum_{t=0}^\infty \frac{\mathbf{E}[|\alpha^t \xi_i^{t+1}|^p \mid \mathcal{F}^t]}{(u^t)^p} \leq \sum_{t=0}^\infty \frac{(\alpha^t)^2 \Xi}{(u^t)^p} \leq \Xi \sum_{t=0}^\infty \frac{(\alpha^t)^2}{(\sum_{k=t^0}^t \alpha^k)^2} < \infty, \text{ a.s.,}$$

where the last inequality follows from the fact that α^t is square-summable. Consequently, by Theorem 8,

$$\sum_{k=t^0}^t \frac{\alpha^k}{\sum_{k=t^0}^t \alpha^k} \xi_i^{k+1}(X_i^k - x_i^*) \to 0, \text{ a.s., } t \to \infty$$

thereby leading to

$$\sum_{i=1}^N (\lambda_i \sum_{k=t^0}^t \frac{\alpha^k}{\sum_{k=t^0}^t \alpha^k} \xi_i^{k+1}(X_i^k - x_i^*)) \to 0, \text{ a.s., } t \to \infty.$$

Since $\sum_{k=t^0}^{\infty} \alpha^k = -\infty$, we therefore have

$$
(\sum_{k=t^0}^{t} \alpha^k) \left\{ b_{\max}(\epsilon) + \sum_{i=1}^{N} (\lambda_i \sum_{k=t^0}^{t} \frac{\alpha^k}{\sum_{k=t^0}^{t} \alpha^k} \xi_i^{k+1}(X_i^k - x_i^*)) \right\} \to -\infty, \ \text{a.s.}, t \to \infty.
$$

Finally, note that since $\| \cdot \|^*$ is a norm, $S^t = \sum_{k=t^0}^{t} \beta^k (\|\tilde{v}(X^k)\|^*)^2$ is a submartingale adapted to $\mathcal{F}^{t+1}$. We check that S^t is l_1-bounded:

$$
\mathbf{E}[S^t] = \sum_{k=t^0}^{t} \beta^k \mathbf{E}[(\|\tilde{v}(X^k)\|^*)^2] \leq \sum_{k=t^0}^{t} \beta^k V < V \sum_{k=t^0}^{\infty} \beta^k < \infty, \forall t.
$$

Consequently, by Theorem 9, it follows that $S^t \to S, \ \text{a.s.}, t \to \infty$, with S finite almost surely. This then leads to

$$
(\sum_{k=t^0}^{t} \alpha^k) \left\{ b_{\max}(\epsilon) + \sum_{i=1}^{N} (\lambda_i \sum_{k=t^0}^{t} \frac{\alpha^k}{\sum_{k=t^0}^{t} \alpha^k} \xi_i^{k+1}(X_i^k - x_i^*)) \right\} + \sum_{k=t^0}^{t} \beta^k (\|\tilde{v}(X^k)\|^*)^2 \to -\infty, \ \text{a.s.}, t \to \infty.
$$

This further implies that $F^\lambda(\mathbf{x}^*, Y^t) \to -\infty, \ \text{a.s.}, t \to \infty$, which contradicts the first statement in Lemma 2.

For the second part, assume for contradiction purposes no $B(\mathbf{x}^*, \epsilon)$ is contained in $\tilde{B}(\mathbf{x}^*, \delta)$, which means that for any $\delta > 0, \exists \mathbf{y}^l$, such that $\|Q(\mathbf{y}^l) - \mathbf{x}^*\| = \delta$ but $F^\lambda(\mathbf{x}^*, \mathbf{y}^l) \geq \epsilon$. This produces a sequence $\{\mathbf{y}^l\}_{l=0}^{\infty}$ such that $C(\mathbf{y}^l) \to \mathbf{x}^*$ but $F^\lambda(\mathbf{x}^*, \mathbf{y}^l) \geq \epsilon, \forall l$. However, this is a contradiction, because $C_i(h_i) = \Pi_{\mathcal{P}_i}(y_i)$ and hence $h_i^*(y_i) = \Pi_{\mathcal{P}_i}(y_i) y_i - \frac{1}{2}(\Pi_{\mathcal{P}_i}(y_i))^2$. Consequently, if $C(\mathbf{y}^t) \to \mathbf{p}$ for some $\mathbf{p} \in \mathcal{P}$, then $\Pi_{\mathcal{P}}(\mathbf{y}^t) = C(\mathbf{y}^t) \to \mathbf{p}$, leading to:

$$
F^\lambda(\mathbf{p}, \mathbf{y}^t) = \sum_{i=1}^{N} \lambda_i (h_i(p_i) - p_i y_i + h_i^*(y_i)) = \sum_{i=1}^{N} \lambda_i (\frac{1}{2} p_i^2 - p_i y_i + \Pi_{\mathcal{P}_i}(y_i) y_i - \frac{1}{2}(\Pi_{\mathcal{P}_i}(y_i))^2) \quad (45)
$$

$$
= \frac{1}{2} \sum_{i=1}^{N} \lambda_i (p_i^2 - (\Pi_{\mathcal{P}_i}(y_i))^2) + \frac{1}{2} \sum_{i=1}^{N} \lambda_i y_i (\Pi_{\mathcal{P}_i}(y_i) - p_i) \to 0, \quad (46)
$$

Consequently, the claim follows.

2. Since each $h_i(\cdot)$ is 1-strongly convex, $h(\cdot) = (h_1(\cdot), \ldots, h_N(\cdot))$ is 1-strongly convex. By a standard result in convex analysis Rockafellar (2015), $C(\cdot)$ is 1-Lipschitz continuous. Since v is lipschitz continuous by assumption, $v(Q(\cdot))$ is Lipschitz continuous. Consequently, standard results in differential equations (Coddington and Levinson (1955)) imply that a unique solution exists for the ODE.

3. Benaïm (1999) gives sufficient conditions that ensure a random trajectory to be an asymptotic pseudotrajectory of a semiflow almost surely. We shall state one set of sufficient conditions directly in the current context as follows.

   If for some $q \geq 2$, the following list of conditions are satisfied:

   (a) $\sup_t \mathbf{E}[(\|\xi^{t+1}\|^*)^q] < \infty$.

(b) $\sum_{n=0}^{\infty} (\alpha^t)^{1+\frac{q}{2}} < \infty$.

(c) $\sup_t \|\mathbf{x}^t\| < \infty$.

Then the affinely interpolated process $Y(t)$ is an asymptotic pseudotrajectory of the semi-flow ϕ induced by the ODE almost surely:

$$\forall T > 0, \lim_{t \to \infty} \sup_{0 \leq h \leq T} \|Y(t+h), \phi_h(Y(t))\|^* = 0, \text{a.s.}.$$

Choose $q = 2$, the above conditions can be easily verified: (a) holds by Assumption 2; (b) holds since α^t is square summable; (c) holds since the decision space $\mathcal{X}$ is compact. Therefore the claim follows.

4. By a well-known result in variational analysis (Rockafellar and Wets (2009)), each $h_i(\cdot)$ is differentiable and

$$\frac{dh_i^*(y_i)}{dy_i} = C_i(y_i). \tag{47}$$

Note further that since $\phi_t(\mathbf{y})$ is the solution to the ODE (under the initial condition $\mathbf{y}$), we have $\frac{d\phi_t(\mathbf{y})}{dt} = v(\mathbf{x}(t))$. Written out component-wise, we have

$$\frac{d(\phi_t(\mathbf{y}))_i}{dt} = v_i(\mathbf{x}(t)). \tag{48}$$

We can thus compute the derivative of λ-Fenchel coupling as follows:

$$\frac{F^\lambda(\mathbf{x}^*, \phi_t(\mathbf{y}))}{dt} = \frac{\sum_{i=1}^N \lambda_i \{h_i(x_i^*) - (\phi_t(\mathbf{y}))_i x_i^* + h_i^*((\phi_t(\mathbf{y}))_i)\}}{dt} \tag{49}$$

$$= \sum_{i=1}^N \lambda_i \left\{ -\frac{d(\phi_t(\mathbf{y}))_i}{dt} x_i^* + C_i(y_i) \frac{d(\phi_t(\mathbf{y}))_i}{dt} \right\} \tag{50}$$

$$= \sum_{i=1}^N \lambda_i \{-v_i(\mathbf{x}(t)) x_i^* + v_i(\mathbf{x}(t)) C_i(y_i)\} \tag{51}$$

$$= \sum_{i=1}^N \lambda_i v_i(\mathbf{x}(t))(x_i(t) - x_i^*) \leq 0, \tag{52}$$

where the second equality follows from Equation (47), the third equality follows from Equation (48), and the last inequality follows from $\mathbf{x}^*$ is λ-variationally stable. The monotonicity property therefore follows.

5. For any given $\epsilon > 0$, pick an $\hat{\epsilon} > 0$ such that $B(\mathbf{x}^*, \hat{\epsilon}) \subset \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2})$. By Equation (52), we have

$$\frac{F^\lambda(\mathbf{x}^*, \phi_t(\mathbf{y}))}{dt} = \sum_{i=1}^N \lambda_i v_i(\mathbf{x}(t))(x_i(t) - x_i^*) < 0, \forall \mathbf{x}(t) \neq \mathbf{x}^*.$$

Since $\mathcal{X} - B(\mathbf{x}^*, \hat{\epsilon})$ is a compact set and each $v_i(\cdot)$ is a continuous function, we have

$$\sum_{i=1}^N \lambda_i v_i(\mathbf{x}(t))(x_i(t) - x_i^*) \leq -a_{\tilde{\epsilon}}, \forall \mathbf{x}(t) \in \mathcal{X} - B(\mathbf{x}^*, \hat{\epsilon}), \tag{53}$$

for some positive constant $a_{\tilde{\epsilon}}$.

Starting at $\mathbf{y}$, by time s, there are two possibilities. The first possibility is that $\mathbf{x}(s) \in \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2})$. In this case, by definition,

$$F^\lambda(\mathbf{x}^*, \phi_s(\mathbf{y})) < \frac{\epsilon}{2}. \tag{54}$$

The second possibility is that $\mathbf{x}(s) \notin \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2})$. This implies that $\mathbf{x}(t) \notin B(\mathbf{x}^*, \hat{\epsilon}), \forall t \in [0, s]$, because otherwise, since $B(\mathbf{x}^*, \hat{\epsilon}) \subset \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2})$, it must be that $\mathbf{x}(s_0) \in \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2})$ for some $s_0 \in [0, s]$. This then implies that, by the monotonicity property established in Claim 4, $F^\lambda(\mathbf{x}^*, \phi_s(\mathbf{y})) \leq F^\lambda(\mathbf{x}^*, \phi_{s_0}(\mathbf{y}))$, thereby leading to $\mathbf{x}(s) \in \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2})$, a contradiction.

Since $\mathbf{x}(t) \notin \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2}), \forall t \in [0, s]$, we have $\mathbf{x}(t) \notin B(\mathbf{x}^*, \tilde{\epsilon}), \forall t \in [0, s]$, leading to that Equation 53 holds for $t \in [0, s]$. Therefore, taking $s = \frac{\epsilon}{2a_{\tilde{\epsilon}}}$, we obtain:

$$F^\lambda(\mathbf{x}^*, \phi_s(\mathbf{y})) \leq F^\lambda(\mathbf{x}^*, \mathbf{y}) - a_{\tilde{\epsilon}} s = F^\lambda(\mathbf{x}^*, \mathbf{y}) - \frac{\epsilon}{2}. \tag{55}$$

Equation (54) and Equation (55) together establish that:

$$F^\lambda(\mathbf{x}^*, \phi_s(\mathbf{y})) \leq \max\{\frac{\epsilon}{2}, F^\lambda(\mathbf{x}^*, \mathbf{y}) - \frac{\epsilon}{2}\}.$$

6. Let $R = \sup_{\mathbf{x} \in \mathcal{X}} \|\mathbf{x}\|$, which is finite since $\mathcal{X}$ is compact. By the definition of dual norm and denote $\lambda_{\max} = \max_i \lambda_i$, we have

$$\sum_{i=1}^N \lambda_i \left\{ (Y_i(t+h) - \phi_h^i(Y(t+h))) \right\} (C_i(\phi_h^i(Y(t+h))) - x_i^*) \leq \tag{56}$$

$$\sum_{i=1}^N \lambda_i \|(Y_i(t+h) - \phi_h^i(Y(t+h))\|_i^* \|(C_i(\phi_h^i(Y(t+h))) - x_i^*)\| \leq \lambda_{\max} R \|Y(t+h) - \phi_h(t+h)\|^*, \tag{57}$$

where $\phi_h^i(Y(t+h))$ is the i-th component of $\phi_h(Y(t+h))$.

Fix some $T > 0$ and define $K_\lambda = \max_i \lambda_i$ and $\delta = \frac{\sqrt{(\lambda_{\max}R)^2 + 2\epsilon K_\lambda} - K_\lambda R}{4 K_\lambda}$. Per Claim 3, we have

$$\forall T > 0, \lim_{t \to \infty} \sup_{0 \leq h \leq T} \|Y(t+h), \phi_h(Y(t))\|^* = 0, \text{a.s.}.$$

Consequently, choose $\tau(\delta, T)$ such that $\|Y(t+h), \phi_h(Y(t))\|^* < \delta, \forall t \geq \tau$. Expanding λ-Fenchel coupling, we obtain that $\forall t \geq \tau$ and $\forall h \in [0, T]$:

$$F^\lambda(\mathbf{x}^*, Y(t+h)) = F^\lambda(\mathbf{x}^*, \phi_h(Y(t)) + Y(t+h) - \phi_h(Y(t))) \leq F^\lambda(\mathbf{x}^*, \phi_h(Y(t))) \tag{58}$$

$$+ \sum_{i=1}^N \lambda_i \left\{ (Y_i(t+h) - \phi_h^i(Y(t))) \right\} (C_i(\phi_h^i(Y(t))) - x_i^*) + K_\lambda (\|Y(t+h) - \phi_h(Y(t))\|^*)^2 \tag{59}$$

$$\leq F^\lambda(\mathbf{x}^*, \phi_h(Y(t))) + \lambda_{\max} R \|Y(t+h) - \phi_h(t)\|^* + K_\lambda (\|Y(t+h) - \phi_h(Y(t))\|^*)^2 \tag{60}$$

$$\leq F^\lambda(\mathbf{x}^*, \phi_h(Y(t))) + \lambda_{\max} R \delta + K_\lambda \delta^2 \tag{61}$$

$$\leq F^\lambda(\mathbf{x}^*, \phi_h(Y(t))) + \lambda_{\max} R \frac{\sqrt{(\lambda_{\max} R)^2 + 2\epsilon K_\lambda} - \lambda_{\max} R}{4K_\lambda} + K_\lambda \left(\frac{\sqrt{(\lambda_{\max} R)^2 + 2\epsilon K_\lambda} - \lambda_{\max} R}{4K_\lambda} \right)^2 \tag{62}$$

$$< F^\lambda(\mathbf{x}^*, \phi_h(Y(t))) + \lambda_{\max} R \frac{\sqrt{(\lambda_{\max} R)^2 + 2\epsilon K_\lambda} - \lambda_{\max} R}{2K_\lambda} + K_\lambda \left(\frac{(\sqrt{(\lambda_{\max} R)^2 + 2\epsilon K_\lambda} - \lambda_{\max} R)^2}{4K_\lambda^2} \right) \tag{63}$$

$$= F^\lambda(\mathbf{x}^*, \phi_h(Y(t))) + \frac{\epsilon}{2}, \tag{64}$$

where the first inequality follows from Equation (57) and the last equality follows from straight-forward algebraic verification. The claim is therefore established.

7. We start by fixing an arbitrary $\epsilon > 0$. Per Claim 5, there exists an $s > 0$ (depending on ϵ) such that Equation (16) holds. Set the horizon $T = s$. Per Claim 6, there exists a τ (depending on both ϵ and T) such that Equation (17) holds $\forall t \geq \tau$. Now, per Claim 1, $\mathbf{X}^t$ visits $\tilde{B}(\mathbf{x}^*, \delta)$ infinitely often[9]. Therefore, pick an integer $\tau_0 \geq \tau$ such that $\mathbf{X}^{\tau_0} \in \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2})$. With this choice of τ_0, we know that by definition of $\tilde{B}$,

$$F^\lambda(\mathbf{x}^*, Y(\tau_0)) < \frac{\epsilon}{2}. \tag{65}$$

Our goal is to establish that $F^\lambda(\mathbf{x}^*, Y(\tau_0 + h)) < \epsilon, \forall h \in [0, \infty)$. To that end, partition the time $[0, \infty)$ into disjoint time intervals $[0, T), [T, 2T), \ldots, [nT, (n+1)T), \ldots$.

Per Claim 4, the monotonicity property given in Equation (15) implies that:

$$F^\lambda(\mathbf{x}^*, \phi_h(Y(\tau_0))) \leq F^\lambda(\mathbf{x}^*, \phi_0(Y(\tau_0))) = F^\lambda(\mathbf{x}^*, Y(\tau_0)) < \frac{\epsilon}{2}, \forall h \geq 0, \tag{66}$$

where the equality follows from the semi-group property of a semiflow.

Per Equation (17), for $h \in [0, T)$, we then have:

$$F^\lambda(\mathbf{x}^*, Y(\tau_0 + h)) < F^\lambda(\mathbf{x}^*, \phi_h(Y(\tau_0))) + \frac{\epsilon}{2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \tag{67}$$

where the last inequality follows from Equation (66).

Now assume inductively that Equation (67) holds for every $h \in [nT, (n+1)T)$, where n is a non-negative integer. We then have $\forall h \in [nT, (n+1)T)$:

$$F^\lambda(\mathbf{x}^*, Y(\tau_0 + T + h)) < F^\lambda(\mathbf{x}^*, \phi_T(Y(\tau_0 + h))) + \frac{\epsilon}{2} \leq \max\{\frac{\epsilon}{2}, F^\lambda(\mathbf{x}^*, Y(\tau_0 + h)) - \frac{\epsilon}{2}\} + \frac{\epsilon}{2} \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \tag{68}$$

where the first inequality follows from Equation (17), the second inequality follows from Equation (16), and the third inequality follows from the induction hypothesis $F^\lambda(\mathbf{x}^*, Y(\tau_0 + h)) < \epsilon$. Consequently, Equation (67) holds for every $h \in [(n+1)T, (n+2)T)$, thereby completing the induction and establishing that:

$$F^\lambda(\mathbf{x}^*, Y(\tau_0 + h)) < \epsilon, \forall h \in [0, \infty).$$

∎