Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Robust Power Management via Learning and Game Design

Zhengyuan Zhou, Panayotis Mertikopoulos, Aris L. Moustakas, Nicholas Bambos, Peter Glynn

informs
http://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Vol. 69, No. 1, January–February 2021, pp. 331–345
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Methods

Robust Power Management via Learning and Game Design

Zhengyuan Zhou,[a] Panayotis Mertikopoulos,[b] Aris L. Moustakas,[c] Nicholas Bambos,[d] Peter Glynn[d]

[a] Department of Technology, Operations, and Statistics, Stern School of Business, New York University, New York, New York 10012; [b] University Grenoble Alpes, CNRS, Grenoble INP, Inria, LIG, F-38000 Grenoble, France; [c] Department of Physics, University of Athens, Athens, Greece; [d] Department of Management Science and Engineering, Stanford University, Stanford, California 94305

Contact: zzhou@stern.nyu.edu, https://orcid.org/0000-0002-0005-9411 (ZZ); panayotis.mertikopoulos@imag.fr, https://orcid.org/0000-0003-2026-9616 (PM); arislm@phys.uoa.gr (ALM); bambos@stanford.edu (NB); glynn@stanford.edu (PG)

Received: August 22, 2017
Accepted: September 18, 2020
Published Online in Articles in Advance: December 24, 2020

Subject Classifications: computer science: artificial intelligence; games/group decisions
Area of Review: Machine Learning

https://doi.org/10.1287/opre.2020.1996

Abstract. 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" where we (1) design a game with suitable payoff functions such that the optimal joint power profile in the original power management problem is the unique Nash equilibrium of the designed game; (2) derive distributed power management algorithms by directing the networks' users to employ a no-regret learning algorithm to maximize their individual utility over time. To establish convergence, we focus on the well-known online eager gradient descent learning algorithm in the class of weighted strongly monotone games. In this class of games, we show that when players only have access to imperfect stochastic feedback, multiagent online eager gradient descent converges to the unique Nash equilibrium in mean square at a $O(\frac{1}{T})$ rate.

In the context of power management in static networks, we show that the designed games are weighted strongly monotone if the network is feasible (i.e., when all users can concurrently attain their target rates). This allows us to derive a geometric convergence rate to the joint optimal transmission power. More importantly, in stochastic networks where channel quality fluctuates over time, the designed games are also weighted strongly monotone and the proposed algorithms converge in mean square to the joint optimal transmission power at a $O(\frac{1}{T})$ rate, even when the network is only feasible on average (i.e., users may be unable to meet their requirements with positive probability). This comes in stark contrast to existing algorithms (like the seminal Foschini–Miljanic algorithm and its variants) that may fail to converge altogether.

Funding: P. Mertikopoulos is partially supported by the COST Action [CA16228] European Network for Game Theory (GAMENET) for this work, and is grateful for financial support by the French National Research Agency (ANR) in the framework of the starting grant ORACLESS [ANR–16–CE33–0004–01, the Investissements d'avenir program [ANR-15-IDEX-02], the LabEx PERSYVAL [ANR-11-LABX-0025-01], and MIAI@Grenoble Alpes [ANR-19-P3IA-0003]. A. Moustakas is supported by the U.S. Office of Naval Research through the grant [N62909-18-1-2141]. N. Bambos is partially supported by Koret Foundation under the Digital Living 2030 project.

Keywords: power management • wireless network • online learning • Nash equilibrium

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 emerging era of the Internet of Things (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; Reddy et al. 2008; Seferoglu et al. 2008), fairness (Eryilmaz et al. 2006), delays (Eryilmaz et al. 2008, Altman et al. 2010), backlog (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; compare (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 a memory-constrained wireless transmitter where such bookkeeping is in general infeasible—in other words, past information must be exploited in an implicit, parsimonious manner.

1.1. Related Work and Our Contributions

Our contributions are three-fold, and we discuss them in the context of existing work below.

First, we propose two novel power control algorithms (Variants A and B in Algorithm 2), both of which require meager, $O(1)$ 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. The proposed algorithms are quite simple and lend themselves easily to being implemented on wireless devices.

Second, we provide theoretic performance guarantees that establish the robustness and optimality of the proposed algorithms. More precisely, in the case of a static network environment that is feasible (i.e., when all users can concurrently attain their target rates), the proposed algorithms converge to the joint optimum transmission power vector at a geometric rate $O(\kappa^{-T})$ (for some $\kappa > 1$). More importantly, in a stochastic iid network environment (where the fluctuating network is sometimes feasible and sometimes not), we show that the proposed algorithms still converge to a *deterministic* power vector in mean square at a $O(1/T)$ rate, provided only that the network is feasible *in the mean*, that is, even if the network is infeasible with positive probability. We find this guarantee particularly appealing because it incorporates elements of both stability and optimality. The former (stability) is because the proposed algorithm converges to a *fixed* constant power vector despite the persistent, random fluctuations in the network (and, in particular, even if power control is *not* feasible for a given channel realization). The latter (optimality) is because the algorithms' end state is an *optimal* solution to the network's power management problem with respect to the channels' mean value statistics. 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. 2016c).

We obtain these theoretic guarantees via a combined approach of learning and "game design" where we design a game with suitable reward functions such that the optimal joint power profile in the original power management problem is the unique Nash equilibrium of the designed game. We then show that if the network environment is feasible, the designed game is *weighted strongly monotone* (WSM): this is an important class of games where convergence to a Nash equilibrium can be characterized under suitable no-regret online learning algorithms. We then establish the equivalence between the proposed power control algorithms and multiagent online eager gradient descent (EGD) as applied to the designed games: by showing that the latter converges to the unique Nash equilibrium of a WSM game, and because the unique Nash equilibrium by design is the optimal joint power profile, we readily obtain all the desired results for the proposed power control algorithms.

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 (Alpcan et al. 2002, 2006; Fan et al. 2006; Menache and Ozdaglar 2010; Han et al. 2014; Zhou and Bambos 2015; Zhou et al. 2016a, b, 2018a). 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 theoretic 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 between 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 work 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 games, but it instead applies to WSM games. 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. 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.

Our third contribution is to establish quantitative Nash equilibrium convergence results for multiagent no-regret learning in the class of WSM games. Specifically, we consider a general game-theoretic learning framework where each agent is endowed with a continuous action set and does not have the knowledge of its (or any other agent's) reward function. Instead, each agent only has access to an imperfect gradient feedback during the learning process.[2] In this setting, we focus on online eager gradient descent, a learning algorithm with tight regret bounds: when the sequence of reward functions are concave (as a function of an agent's own action), then up to a multiplicative constant related to the problem's dimensionality, it achieves the $O(\sqrt{T})$ regret bound that is min-max optimal; the stronger $O(\log T)$ regret bound can be achieved if the sequence of reward functions are strongly concave (see Shalev-Shwartz 2012, Hazan 2016).

In this context, we show that if each agent employs online eager gradient descent to maximize their cumulative reward, the last iterate of the joint action profile of all agents (now a random variable as a result of stochastic first-order feedback) converges in mean square to the unique Nash equilibrium at a $O(1/T)$ rate, provided the underlying game is WSM. In comparison, much of the existing literature on no-regret game-theoretic learning (including other no-regret algorithms) focuses on the convergence of the time-averaged sequence of the joint action $\bar{\mathbf{x}}^t = \sum_{k=1}^t \gamma_k \mathbf{x}^k / \sum_{k=1}^t \gamma^k$ (where γ^t is the step-size and $\mathbf{x}^t$ is the joint action at time t) in different special classes of games, such as convex potential games, zero-sum games, routing games, and so on (Cesa-Bianchi and Lugosi 2006, Krichene et al. 2015, Balandat et al. 2016, Lam et al. 2016). Such time-average convergence results, although useful in their own right, are weaker than the corresponding last-iterate convergence results (i.e., convergence of $\mathbf{x}^t$) and are not fine grained enough to characterize the actual joint evolution of the agents' sequence of play.

Other related considerations date as far back as the work of Arrow and Hurwicz (1960) who considered a special class of multiagent games that is a generalization of two-player zero-sum games and showed that if each agent applies gradient descent (without projection), then the continuous-time approximation of the discrete-time dynamics (modelled by a differential equation) converges to the unique Nash equilibrium of the game. However, whether the discrete-time iterates converge to Nash in that class is unclear; our convergence analysis can be seen as a partial answer to this question. More recently, Antipin (2002) considered a class of two-player game with a generalized potential function and established that gradient prediction-type projection method, a centralized algorithm the authors developed, converges to a Nash equilibrium at a geometric rate. However, this gradient prediction-type projection method cannot be decentralized: in each iteration, one player applies a partial gradient to update its action and the second player applies its own partial gradient but using the already updated action value from the first player. Consequently, this cannot be used in an online learning setting where each player updates its own action simultaneously.

At around the same time as Antipin (2002), Flåm (2002) considered a class of multiagent games called evolutionary stable games (a very broad class of games that include WSM games as a special case) and established local convergence results. In particular, they showed that if each player applies eager gradient descent with *perfect* feedback, convergence to a Nash equilibrium is guaranteed if one starts in the vicinity of a Nash equilibrium. Facchinei and Pang (2007) strengthened both the assumption and the conclusion and gave explicit convergence rates: translating their results (originally in variational inequalities) into games, they established that in strongly monotone games, if each player applies eager gradient descent, then convergence to the unique Nash equilibrium is guaranteed at a geometric rate, a result we apply to give sharper performance guarantees for our designed algorithms in the static network environment case. Cui et al. (2008) considered a particular game in medium access control and investigated a distributed algorithm called gradient-play (in addition to other algorithms, such as best response and Jacobi play) and established geometric convergence to the unique Nash equilibrium of the game when each agent applies the gradient-play algorithm. However, this gradient-play algorithm is different from gradient descent in that each player applies a partial gradient only to part of one's reward function; because of this, it is unclear whether this algorithm is no-regret if applied in an online learning setting.

Importantly, all works discussed above are concerned with the case where agents receive perfect gradient feedback on their actions (some time with observability of a nonindividual gradient). From a dynamical system viewpoint, Kar et al. (2012) considered the more general case where the gradient is corrupted by a stochastic but unbiased noise and established that if the game is strongly monotone, then if each agent applies gradient descent with noise (without projection), almost sure convergence to the unique Nash equilibrium is guaranteed. Recently, (Zhou et al. 2017b, Mertikopoulos and Zhou 2019) established that in variationally stable games (a broad class of games that includes strongly monotone games), if each agent applies gradient descent with lazy projection, then even with stochastic (unbiased) noise, almost sure convergence to Nash is guaranteed. Zhou et al. (2017a, 2018b) further considered other types of imperfect feedback in variationally stable games that include delay and loss and characterized the robustness of the algorithms therein. However, all these convergence results are asymptotic and finite-time rates are impossible at such generality. Further, to the best of our knowledge, a qualitative $O(1/T)$ convergence rate is not known for multiagent online gradient descent (with our without projection) in WSM games. Our work fills in this gap in the literature with a relatively simple proof.

For convenience, we summarize this thread of findings in Table 1.

Table 1. Number of Iterations Required by the FM and EGD Algorithms to Reach an ϵ-Optimal State in Wireless Networks with Fixed (Deterministic) or Fluctuating (Stochastic) Channels

	Deterministic	Stochastic
Foschini–Miljanic algorithm	$\log(1/\epsilon)$	——
Eager gradient descent (our paper)	$\log(1/\epsilon)$	$O(1/\epsilon)$

Note. In the static case, both algorithms converge at a geometric rate; in the stochastic case, the FM algorithm may fail to stabilize (even in probability), whereas the proposed EGD policy converges within $O(1/\epsilon)$ iterations.

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 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).

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 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) consisting of the following components:

1. The reweighted 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 reweighted noise vector γ, where

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

3. Define $\lambda_{\max}(W)$ to be the spectral radius of the matrix W: $\lambda_{\max}(W)$ is the eigenvalue with the largest absolute value. As noted in Tan (2014), the reweighted gain matrix W is a nonnegative (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.

Theorem 1 (Tan 2014). *A channel is feasible with respect to r^* if and only if the largest eigenvalue $\lambda_{\max}(W)$ of the reweighted gain matrix W satisfies $\lambda_{\max}(W) < 1$.*

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, in a feasible channel, $\mathbf{p}^*$ 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 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:

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, that is, $(\mathbf{G}, \eta)$ remains the same from iteration to iteration. By a monotonicity argument, we can leverage the results in Foschini and Miljanic (1993) to give the following characterization:

Theorem 2 (Foschini and Miljanic 1993)**.** *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.*

Remark 1. The original result in Foschini and Miljanic (1993) assumes there is no contraint on the maximum power (and hence in their original setting, the joint power iterate in Algorithm 1 would diverge to infinity if the channel is not feasible). However, the corresponding results when there is a constrained feasible power support set (as given in Theorem 2) is almost an immediate corollary therefrom via a simple monotonicity argument. Because in practice, maximum power is always bounded, and because the subsequent power management literature (Tan 2014) in wireless networks do assume bounded power, we follow this convention in this paper too.

2.3. Motivation of the Paper

Although FM enjoys good convergence properties in deterministic channels (as given in Theorem 2), it quickly loses its appeal in stochastic channels (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 $\mathbf{p}^t$ to determine the next power iterate $\mathbf{p}^{t+1}$, has certain inherent instability, which is particularly manifested when the underlying channel is stochastic. In stochastic channels, we use W^t and γ^t to denote the random reweighted gain matrix and the random reweighted noise at iteration t, respectively. P^t denotes the random power vector at iteration t. More concisely, denoting $\Pi_{\mathcal{P}}(\mathbf{y}) = \arg\min_{\mathbf{x} \in \mathcal{P}} \|\mathbf{x} - \mathbf{y}\|$ and the projection operator $\Pi_{\mathcal{P}}(\cdot)$ for $\mathcal{P}$, the FM update can be written as:

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

In a stochastic 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. 2016c) rather than a deterministic quantity. In fact, convergence to a stationary distribution is already the best one can hope for when using FM. This reveals two main drawbacks of FM. First, as mentioned previously, when operating in a stochastic channel, FM is not very stable. Second, convergence to a limiting stationary distribution is not fully desirable. Even when FM does converge 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 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. Proposed Power Control Algorithms

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.[3] This dictates that a transmitter can only use the aggregate information, such as SINR and/or total interference and noise (i.e., information that can be sensed by each individual link), as opposed to the 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 is highly desirable and necessary in practice because, for a memory-constrained wireless transmitter, such bookkeeping is in general infeasible. In fact, the popularity of the FM algorithm stems in large part from the very limited memory it requires in performing each update. Consequently, for a new algorithm to be practically useful, an economic representation that incorporates all such information from the past in a way compact way is needed.

Here we propose a robust power control algorithm that satisfies those two constraints. The proposed algorithm has two variants, which we subsequently label as 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 space concerns, we will directly consider the stochastic channel case (specializing the result into the deterministic channel case is straightforward). Before stating the algorithm, we first present the model of the stochastic channel within which the algorithm operates, which is the same model as in Zhou et al. (2016c):

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 2 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.

Algorithm 2 (Robust Power Control in Stochastic Channels)

1: Initialize $\gamma_0 > 0$.
2: Each link i chooses an initial $P_i^1 \in \mathcal{P}_i$.
3: **for** $t = 1, 2, \ldots$ **do**
4:     **for** each link i **do**
5:         **Variant A:** $\tilde{P}_i^{t+1} = P_i^t - \frac{\gamma_0}{t}(G_{ii}^t P_i^t - \frac{\gamma_0}{t}(G_{ii}^t P_i^t - r_i^*(\sum_{j\neq i} G_{ij}^t P_j^t + \eta_i^t))$
6:         **Variant B:** $\tilde{P}_i^{t+1} = P_i^t - \frac{\gamma_0}{t}(P_i^t - r_i^* \frac{\sum_{j\neq i} G_{ij}^t P_j^t + \eta_i^t}{G_{ii}^t})$
7:         $P_i^{t+1} = \Pi_{\mathcal{P}_i}(\tilde{P}_i^{t+1})$
8:     **end for**
9: **end for**

Remark 2. First, note that $\tilde{P}_i^t$'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 P_i^t's. Second, as we shall make precise later, each $\tilde{P}_i^t$ is the weighted average of the gradients of a certain cost function. Consequently, $\tilde{P}^t$ (the vector of all individual $\tilde{P}_i^t$'s) resides in the dual space of the space $\mathcal{P}$ of feasible joint transmission power. In fact, line 7 shows that the projection transforms a point in this dual space to a point in the action space (i.e., $\mathcal{P}$). Third, to perform the update in line 5 (Variant A) 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. For Variant B, only the SINR $\frac{\sum_{j\neq i} G_{ij}^t P_j^t + \eta_i}{G_{ii}^t P_i^t}$ is needed, because from this SINR and P_i^t, the ratio $\frac{\sum_{j\neq i} G_{ij}^t P_j^t + \eta_i}{G_{ii}^t}$ can be recovered. Consequently, the requirement to implement Variant B is even less stringent than that of Variant A. However, as we shall see, Variant B needs a slightly stronger assumption on the channel gains for convergence and optimality.

Remark 3. Even though we make the stochastic channel environment as stated in Assumption 1, Algorithm 2 also applies in an arbitrary time-varying environment: an environment where the channel gains G_{ij}^t and the noise η^t do not follow from a stationary stochastic process and can arbitrarily change. It turns out even in that case, the designed algorithm has the inherited the no-regret guarantee with respect to a particular cost function. Although the arbitrary time-varying environment is not the focus of this paper, this guarantee is still desirable and will become clear once we have made the connection to online learning on games.

4. Eager Gradient Descent Learning in Weighted Strongly Monotone Games

We present the framework of EGD learning on the class of WSM games. The results in this section serve as the foundation for the game-design approach of power control and enable us to subsequently establish the theoretical guarantees of the proposed power control algorithm in Algorithm 2. Wherever possible, we shall use the notation that matches the power control setting to make explicit the connection between EGD learning on WSM games and power control.

4.1. Weighted Strongly Monotone Games

We start with the definition that 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 (λ, β)-weighted strongly monotone game (hereafter referred to as (λ, β)-WSM for short) $\mathcal{G}$ 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

1. Each $\mathcal{X}_i$ is a compact and convex subset of $\mathbf{R}^{d_i}$ and each u_i is continuous in $\mathbf{x}$ and continuously differentiable in x_i, with $\frac{\partial u_i(\mathbf{x})}{\partial x_i}$ continuous in $\mathbf{x}$.
2. There exists some $\lambda \in \mathbf{R}^N_{++}$ such that $\sum_{i=1}^{N} \lambda_i \langle v_i(\mathbf{x}) - v_i(\mathbf{y}), x_i - y_i \rangle \le -\frac{\beta}{2} \sum_{i=1}^{N} \lambda_i \|\mathbf{x}_i - \mathbf{y}_i\|_2^2, \forall \mathbf{x}, \mathbf{y} \in \mathcal{X}$, where $v_i(\mathbf{x}) \triangleq \frac{\partial u_i(\mathbf{x})}{\partial x_i}$.

Remark 4. The class of WSM games is a proper subclass of diagonal strict concave games, first introduced in Rosen (1965). $\mathcal{G}$ is a diagonal strict concave game if, instead of the second requirement in Definition 4, the weaker assumption $\sum_{i=1}^{N} \lambda_i \langle v_i(\mathbf{x}) - v_i(\mathbf{y}), x_i - y_i \rangle \le 0$ (with equality if and only if $\mathbf{x} = \mathbf{y}$) is imposed. Note that even this weaker condition implies that each u_i is concave in x_i. Next, a few words on notation. $\mathbf{x}_{-i}$ 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 $\mathbf{v}(\mathbf{x}) = (v_1(\mathbf{x}), \ldots, v_N(\mathbf{x}))$. Note that, per the definition, the joint partial gradient $\mathbf{v}(\mathbf{x})$ always exists and is a continuous function on the joint action space $\mathcal{X}$. Finally, recall that $\mathbf{x}^* \in \mathcal{X}$ is a Nash equilibrium if for each $i \in \mathcal{N}$, $u_i(x_i^*, \mathbf{x}_{-i}^*) \ge u_i(x_i, \mathbf{x}_{-i}^*)$, $\forall x_i \in \mathcal{X}_i$. The celebrated result by Rosen (1965) is that there is always a unique Nash equilibrium for every diagonally strict concave (DSC) game. Consequently, every (λ, β)-WSM game admits a unique Nash equilibrium.

We conclude this subsection with a simple sufficient condition ensuring that a game is (λ, β)-WSM. In Rosen (1965), a simple sufficient condition is given to ascertain whether a game is DSC. By an adaptation of that sufficient condition, we have a sufficient condition to check whether a given game is a WSM game.

Lemma 1. *Given $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^{N} \mathcal{X}_i, \{u_i\}_{i=1}^{N})$, 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:*

$$H_{ij}^{\lambda}(\mathbf{x}) = \frac{1}{2} \lambda_i \frac{\partial v_i(\mathbf{x})}{\partial x_j} + \frac{1}{2} \lambda_j \frac{\partial v_j(\mathbf{x})}{\partial x_i}. \tag{6}$$

If there exists some $\beta > 0$ such that $\lambda_{\max}(H^\lambda(\mathbf{x})) \le -\beta$, $\forall \mathbf{x} \in \mathcal{X}$, then $\mathcal{G}$ is (λ, β)-WSM.

Remark 5. Note that by definition, the λ-weighted Hessian matrix is always symmetric and hence all of its eigenvalues are real. Using the same notation as in the preceding power control section, $\lambda_{\max}(H^\lambda(\mathbf{x}))$ refers to its maximum eigenvalue.

4.2. Multiagent EGD Learning with Noisy Gradient

When players repeatedly interact with one another where the reward is given by a fixed (but possibly unknown) stage game, it is an important question as to what learning dynamics the players would adopt. One answer provided by the online learning literature is EGD, which enjoys the no-regret property (Shalev-Shwartz 2012) when the utility function is concave: each agent, when interacting with the environment (consisting of other agents), will achieve comparable performance to the best fixed action in hindsight, irrespectively of what the other agents do. When each agent adopts EGD, we obtain the multiagent learning dynamics as given in Algorithm 3.

Algorithm 3 (Multiagent Eager Gradient Descent)

Each player i chooses an initial $x_i^1 \in \mathcal{X}_i$.
for $t = 1, 2, \ldots$ **do**
  for $i = 1, \ldots, N$ **do**
    $\tilde{x}_i^{t+1} = x_i^t + \alpha^t v_i(\mathbf{X}^t)$
    $x_i^{t+1} = \operatorname{argmin}_{x_i \in \S_i} \|\tilde{x}_i^{t+1} - x_i\|_2$
  end for
end for

Note that $\tilde{X}_i^t$ takes a gradient step (with respect to i-th player's partial gradient only) from the previous action and is then immediately projected back to the feasible action space. This is called eager projection as the intermediate $\tilde{X}_i^t$ is immediate thrown away after the projection. This type of projection stands in contrast with lazy projection (Nemirovski et al. 2009), where $\tilde{X}_i^t$ (that can potentially be out of the feasible action space) is always kept and projected only when needed.

Remark 6. The convergence of multiagent EGD to Nash equilibria (and the corresponding convergence rate) in weighted strongly monotone games follows straightforwardly from the literature. Following a classical result in variational inequality (Facchinei and Pang 2007), we have a geometric convergence rate: let $\mathbf{x}^*$ be the unique Nash equilibrium of a (λ,β)-WSM game and let $\mathbf{x}^t$ be the iterates generated from Algorithm 3. Then $\|\mathbf{x}-\mathbf{x}^*\|_2^2 \le \epsilon$ after $O(\log\frac{1}{\epsilon})$ iterations if $0 < \inf_t \alpha^t < \frac{2\beta}{L^2}$, where L is the Lipschitz constant of the gradient $v(\mathbf{x})$. This linear rate means that convergence to Nash is exponentially fast. When $v(\mathbf{x})$ is not Lipschitz, convergence is still guaranteed, provided a diminishing step-size is chosen (note that in the Lipschitz case, the step-size α^t cannot decrease to zero: this is necessary in order for contraction to kick in, which ensures the geometric convergence rate). However, in the absence of Lipschitz gradient, exponential convergence is not achievable. Nevertheless, $O(\frac{1}{T})$ is still achievable, which in fact follows as a special case of our subsequent result for the stochastic case in Theorem 3.

Next, we consider the case where gradient is imperfect and noisy.

Algorithm 4 (Multiagent Noisy Eager Gradient Descent)

1: Initialize $\gamma_0 > 0$.
2: Each player i chooses an initial $X_i^1 \in \mathcal{X}_i$.
3: **for** $t = 1, 2, \ldots$ **do**
4:     **for** $i = 1, \ldots, N$ **do**
5:         $\tilde{X}_i^{t+1} = X_i^t + \frac{\gamma_0}{t}\tilde{v}_i(\mathbf{X}^t)$
6:         $X_i^{t+1} = \operatorname{argmin}_{x_i\in\mathcal{X}_i}\|\tilde{X}_i^{t+1} - x_i\|_2$
7:     **end for**
8: **end for**

In Algorithm 4, we have used the capital letters X_i^t and $\tilde{X}_i^t$ 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:

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 the noisy gradients are*

1. *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.,*
2. *bounded in mean square: $\forall i \in \mathcal{N}, \forall t = 0, 1, \ldots, \mathbf{E}[\|\tilde{v}_i(\mathbf{X}^t)\|^2 \mid \mathcal{F}^t] \le \Xi_i$, a.s., 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 7. 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 $(\xi_i^t)_{i=1}^N$ satisfies (for each i):

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

Note that here we only need the noise to be martingale noise, rather than iid noise.

4.3. Convergence of Multiagent Noisy EGD to Nash Equilibrium

We tackle the convergence issue for multiagent EGD with noisy gradient and characterize the convergence rate explicitly. Our main result is that, in a WSM game, multiagent eager EGD given in Algorithm 4 converges in last-iterate to the unique Nash equilibrium in mean square at a $O(\frac{1}{T})$ rate. To that end, we start by recalling two preliminary results in the literature: the first is a fact of sequences first established in Chung (1954) (although it does not appear to be widely known); the second is a simple and well-known variational characterization of a Nash equilibrium.

Lemma 2 (Chung 1954). *Let a_t be a nonnegative sequence such that $a_{t+1} \le a_t(1 - \frac{P}{t^p}) + \frac{Q}{t^{p+q}}$, where $P > q > 0, 0 < p \le 1$, $Q > 0$. Then,*

$$a_t \le \begin{cases} \frac{Q}{P}\frac{1}{t^q} & \text{if } 0 < p < 1 \\ \frac{Q}{P-q}\frac{1}{t^q} & \text{if } p = 1. \end{cases} \tag{7}$$

Lemma 3 (Facchinei and Pang 2003). *Let $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathcal{X}_i, \{u_i\}_{i=1}^N)$ be a continuous game that has $\mathbf{x}^*$ as a Nash equilibrium. Then $\langle v_i(\mathbf{x}^*), x_i - x_i^*\rangle \le 0$ for each $i = 1,, 2\ldots, N$.*

Remark 8. Lemma 3 follows directly from Facchinei and Pang (2003) and is a consequence of the definition of a Nash equilibrium: when no player has any incentive to unilaterally deviate at $\mathbf{x}^*$, then if any player i deviates to x_i, the angle formed by the gradient and its deviation is nonnegative, thereby indicating that his or her individual deviation could not have done better.

Before stating the final convergence result, we make note of a few constants. First, because each $v_i(\cdot)$ is a continuous function over the convex and compact set $\mathcal{X}$, $v_i(\mathbf{x})$ is bounded and denotes $G \triangleq \sum_{i=1}^N \lambda_i \sup_{x\in\mathcal{X}} \|v_i(x)\|_2^2$. Next, per Assumption 2 and Remark 7, we denote $V \triangleq \sum_{i=1}^N \lambda_i V_i$, which is an upper bound for the (weighted) sum of all the variances of the noises. We are now ready to state and prove the convergence result.

Theorem 3. *Let $\mathcal{G}$ be a (λ,β)-WSM game with $\mathbf{x}^*$ being the unique Nash equilibrium. Let X^t's be the sequence of iterates generated from multiagent noisy EGD as given in Algorithm 4. If $\gamma_0 > \frac{1}{\beta}$, then X^t converges to $\mathbf{x}^*$ in mean square at $O(\frac{1}{T})$ rate: for any T, $\mathbf{E}[\|X^t - \mathbf{x}^*\|_2^2] \le \frac{(G+V)\gamma_0^2}{(\min \lambda_i)(\beta\gamma_0 - 1)}\frac{1}{T}$.*

Proof. The proof is not long, so we present it here. Consider the energy function $E_t = \frac{1}{2}\sum_{i=1}^N \lambda_i \|X_i^t - \mathbf{x}_i^*\|_2^2$.

Note that it is a random variable and we will look at how E_{t+1} relates to E_t. By definition, we have

$$
\begin{aligned}
E_{t+1} &= \frac{1}{2}\sum_{i=1}^{N}\lambda_i\|X_i^{t+1} - \mathbf{x}_i^*\|_2^2 \\
&= \frac{1}{2}\sum_{i=1}^{N}\lambda_i\|\mathrm{Proj}_{\mathcal{X}_i}\left(\tilde{X}_i^{t+1}\right) - \mathbf{x}_i^*\|_2^2 \\
&= \frac{1}{2}\sum_{i=1}^{N}\lambda_i\|\mathrm{Proj}_{\mathcal{X}_i}\left(\tilde{X}_i^{t+1}\right) - \mathrm{Proj}_{\mathcal{X}_i}(\mathbf{x}_i^*)\|_2^2, \quad (8)\\
&\le \frac{1}{2}\sum_{i=1}^{N}\lambda_i\|\tilde{X}_i^{t+1} - \mathbf{x}_i^*\|_2^2 = \frac{1}{2}\sum_{i=1}^{N}\lambda_i\|X_i^t \\
&\qquad + \frac{\gamma_0}{t}\tilde{v}_i(\mathbf{X}^t) - \mathbf{x}_i^*\|_2^2, \quad (9)\\
&= \frac{1}{2}\sum_{i=1}^{N}\lambda_i\left(\|X_i^t - \mathbf{x}_i^*\|_2^2 + \|\frac{\gamma_0}{t}\tilde{v}_i(\mathbf{X}^t)\|_2^2\right. \\
&\qquad \left. + 2\frac{\gamma_0}{t}\left\langle\tilde{v}_i(\mathbf{X}^t), X_i^t - \mathbf{x}_i^*\right\rangle\right), \quad (10)\\
&= E_t + \frac{1}{2}\frac{\gamma_0^2}{t^2}\sum_{i=1}^{N}\lambda_i\|\tilde{v}_i(\mathbf{X}^t)\|_2^2 + \frac{\gamma_0}{t}\sum_{i=1}^{N}\lambda_i \quad (11)\\
&\quad \times\left\langle\tilde{v}_i(\mathbf{X}^t), X_i^t - \mathbf{x}_i^*\right\rangle.
\end{aligned}
$$

Taking the expectation of both sides (by first conditioning on $\mathbf{X}^t$ and then averaging over $\mathbf{X}^t$) yields

$$
\begin{aligned}
\mathbf{E}[E_{t+1}] &\le \mathbf{E}[E_t] + \frac{1}{2}\frac{\gamma_0^2}{t^2}\sum_{i=1}^{N}\lambda_i\mathbf{E}\left[\|\tilde{v}_i(\mathbf{X}^t)\|_2^2\right] \\
&\quad + \frac{\gamma_0}{t}\sum_{i=1}^{N}\lambda_i\left\langle\mathbf{E}\left[\tilde{v}_i(\mathbf{X}^t)\right], X_i^t - \mathbf{x}_i^*\right\rangle, \quad (12)\\
&= \mathbf{E}[E_t] + \frac{1}{2}\frac{\gamma_0^2}{t^2}\sum_{i=1}^{N}\lambda_i\mathbf{E}\left[\|\tilde{v}_i(\mathbf{X}^t)\|_2^2\right] \\
&\quad + \frac{\gamma_0}{t}\sum_{i=1}^{N}\lambda_i\mathbf{E}\left[\left\langle\mathbf{E}\left[\tilde{v}_i(\mathbf{X}^t)\mid\mathbf{X}^t\right], X_i^t - \mathbf{x}_i^*\right\rangle\right], \quad (13)\\
&= \mathbf{E}[E_t] + \frac{1}{2}\frac{\gamma_0^2}{t^2}\sum_{i=1}^{N}\lambda_i\mathbf{E}\left[\|v_i(\mathbf{X}^t) + \xi_i^{t+1}\|_2^2\right] \\
&\quad + \frac{\gamma_0}{t}\mathbf{E}\left[\sum_{i=1}^{N}\lambda_i\langle v_i(\mathbf{X}^t), X_i^t - \mathbf{x}_i^*\rangle\right], \quad (14)\\
&\le \mathbf{E}[E_t] + \frac{\gamma_0^2}{t^2}\sum_{i=1}^{N}\lambda_i\left(\mathbf{E}\left[\|v_i(\mathbf{X}^t)\|_2^2\right] + \mathbf{E}\left[\|\xi_i^{t+1}\|_2^2\right]\right) \\
&\quad + \frac{\gamma_0}{t}\mathbf{E}\left[\sum_{i=1}^{N}\lambda_i\langle v_i(\mathbf{X}^t), X_i^t - \mathbf{x}_i^*\rangle\right], \quad (15)\\
&\le \mathbf{E}[E_t] + \frac{C\gamma_0^2}{t^2} + \frac{\gamma_0}{t}\mathbf{E}\left[\sum_{i=1}^{N}\lambda_i\langle v_i(\mathbf{X}^t), X_i^t - \mathbf{x}_i^*\rangle\right], \quad (16)
\end{aligned}
$$

where the first equality follows from Assumption 2 and the last inequality follows from the fact that both $\mathbf{E}[\|v_i(\mathbf{X}^t)\|_2^2]$ and $\mathbf{E}[\|\xi_i^{t+1}\|_2^2]$ are bounded: the former is bounded as a result of $\mathcal{X}$ being bounded and $v(\cdot)$ is continuous; the latter is bounded as a result of Assumption 2. Consequently, we use a constant C to denote the total upper bound. Note that it can be easily checked that $C \le G + V$.

Next, because the game $\mathcal{G}$ is (λ, β)-WSM, we have $\sum_{i=1}^{N}\lambda_i\langle v_i(\mathbf{x}) - v_i(\mathbf{y}), x_i - y_i\rangle \le -\frac{\beta}{2}\sum_{i=1}^{N}\lambda_i\|\mathbf{x}_i - \mathbf{y}_i\|_2^2$, $\forall \mathbf{x}, \mathbf{y} \in \mathcal{X}$. This immediately implies that

$$
\begin{aligned}
&\sum_{i=1}^{N}\lambda_i\langle v_i(\mathbf{x}) - v_i(\mathbf{x}^*), x_i - x_i^*\rangle \\
&\quad = \sum_{i=1}^{N}\lambda_i\langle v_i(\mathbf{x}), x_i - x_i^*\rangle - \sum_{i=1}^{N}\lambda_i\langle v_i(\mathbf{x}^*), x_i - x_i^*\rangle \\
&\quad \le -\frac{\beta}{2}\sum_{i=1}^{N}\lambda_i\|\mathbf{x}_i - \mathbf{x}_i^*\|_2^2, \forall \mathbf{x} \in \mathcal{X}. \quad (17)
\end{aligned}
$$

Consequently, plugging $\mathbf{X}^t$ into the Equation (17) yields

$$
\begin{aligned}
\sum_{i=1}^{N}\lambda_i\langle v_i(\mathbf{X}^t), X_i^t - x_i^*\rangle &\le \sum_{i=1}^{N}\lambda_i\langle v_i(\mathbf{x}^*), X_i^t - x_i^*\rangle \\
&\quad - \frac{\beta}{2}\sum_{i=1}^{N}\lambda_i\|\mathbf{X}_i^t - \mathbf{x}_i^*\|_2^2, \forall \mathbf{x} \in \mathcal{X}.
\end{aligned}
$$

Because $\mathbf{x}^*$ is a Nash equilibrium, it follows from Lemma 3 that $\sum_{i=1}^{N}\lambda_i\langle v_i(\mathbf{x}^*), X_i^t - x_i^*\rangle \le 0$ and hence $\sum_{i=1}^{N}\lambda_i\langle v_i(\mathbf{X}^t), X_i^t - x_i^*\rangle \le -\frac{\beta}{2}\sum_{i=1}^{N}\lambda_i\|\mathbf{X}_i^t - \mathbf{x}_i^*\|_2^2, \forall \mathbf{x} \in \mathcal{X}$. Note that even though $\mathbf{X}^t$ is random, the inequality holds surely. Consequently, it follows that

$$
\begin{aligned}
\mathbf{E}[E_{t+1}] &\le \mathbf{E}[E_t] + \frac{C\gamma_0^2}{t^2} + \frac{\gamma_0}{t}\mathbf{E} \\
&\qquad \times\left[\sum_{i=1}^{N}\lambda_i\langle v_i(\mathbf{X}^t), X_i^t - \mathbf{x}_i^*\rangle\right] \le \mathbf{E}[E_t] \\
&\qquad - \frac{\beta}{2}\frac{\gamma_0}{t}\mathbf{E}\left[\sum_{i=1}^{N}\lambda_i\|\mathbf{X}_i^t - \mathbf{x}_i^*\|_2^2\right] + \frac{C\gamma_0^2}{t^2}, \quad (18)\\
&= \mathbf{E}[E_t] - \frac{\beta\gamma_0}{t}\mathbf{E}[E_t] + \frac{C\gamma_0^2}{t^2} = \left(1 - \frac{\beta\gamma_0}{t}\right)\mathbf{E}[E_t] \\
&\quad + \frac{C\gamma_0^2}{t^2}. \quad (19)
\end{aligned}
$$

Consequently, applying Lemma 2 with $p = q = 1$ and $P = \beta\gamma_0, Q = C\gamma_0^2$ (noting that $\gamma_0 > \frac{1}{\beta}$) immediately yields

$$
\mathbf{E}[E_T] \le \frac{C\gamma_0^2}{\beta\gamma_0 - 1}\frac{1}{T}. \quad (20)
$$

As a result, it follows that $E_T = \|X^t - \mathbf{x}^*\|_2^2 \le \frac{1}{\min\lambda_i}\mathbf{E}[E_T] \le \frac{1}{\min\lambda_i}\frac{C\gamma_0^2}{\beta\gamma_0 - 1}\frac{1}{T}$. ■

5. Theoretical Guarantees of Robust Power Control

In this section, we establish the theoretical guarantees of the proposed power control algorithms. Our approach lies in designing two weighted strongly monotone games such that each variant of the algorithm can be interpreted as a special instance of the multiagent EGD 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 proposed power control algorithms in the first place.

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

When the channel is feasible, per Fact 1, a (necessarily unique) optimal joint transmission power $\mathbf{p}^*$ exists. Because the optimal joint transmission power matches the quality-of-service constraints exactly, every player's utility will be zero if they transmit according to $\mathbf{p}^*$. This implies that $\mathbf{p}^*$ must be a Nash equilibrium, because zero 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 zero. For this player, he or she will have the incentive to transmit at a higher power compared with 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 WSM. 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 WSM.

Lemma 4. *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 $\mathbf{p}^* \in \mathcal{P}$ be the optimal joint transmission power as defined in Definition 3. The we have*

1. *$\beta \triangleq -\lambda_{\max}(\frac{1}{2}(W - \mathbf{I}) + \frac{1}{2}(W^T - \mathbf{I})) > 0$, where W is the reweighted matrix in Equation (3).*
2. *The designed game $\mathcal{G}^1 = (\mathcal{N}, \mathcal{P}, \{u_i\}_{i=1}^N)$ is (λ, β)-WSM with the unique Nash equilibrium $\mathbf{p}^*$.*
3. *The designed game $\mathcal{G}^2 = (\mathcal{N}, \mathcal{P}, \{\tilde{u}_i\}_{i=1}^N)$ is $(\mathbf{1}, \beta)$-WSM with the unique Nash equilibrium $\mathbf{p}^*$.*

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

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

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 reweighted 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

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

Because 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. Because $H_{ij}(\mathbf{p})$ is negative definite for every $\mathbf{p}$ (because it is independent of $\mathbf{p}$), Lemma 1 establishes that the game is λ-DSC. Because $\mathbf{p}^*$ 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. Note that $H^\lambda(\mathbf{p})$ does not depend on $\mathbf{p}$ and is hence uniformly negative definite and has a uniform upper bound β on the largest eigenvalue.

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

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

Computing the $\tilde{\lambda}$-weighted Hessian for $\tilde{\lambda} = (1, 1, \ldots, 1)$ yields

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

thereby establishing the conclusion. ■

With the above two designed games, we can connect the proposed power control algorithms to the multiagent EGD learning dynamics and immediately derive fast convergence results in the deterministic channel environments. We start with some notation. Let M^1 and M^2 be two matrices such that $M^1_{ii} = G_{ii}, M^1_{ij} = -r^*_i G_{ij}$ when $i \neq j$ and $M^2_{ii} = 1, M^2_{ij} = -r^*_i \frac{G_{ij}}{G_{ii}}$ when $i \neq j$. Denote $\rho_1 \triangleq \frac{-2\lambda_{\max}(\frac{1}{2}(W-\mathbf{I})+\frac{1}{2}(W^T-\mathbf{I}))}{\lambda^2_{\max}(M^1)}, \rho_2 \triangleq \frac{-2\lambda_{\max}(\frac{1}{2}(W-\mathbf{I})+\frac{1}{2}(W^T-\mathbf{I}))}{\lambda^2_{\max}(M^2)}$ (note that both ρ_1 and ρ_2 are positive). One thing to note here is that in order to achieve a geometric convergence rate, instead of using diminishing step-size as in Algorithm 2, we need a small constant step-size.

Corollary 1. *Let $\mathbf{p}^t$ be the iterates generated from Algorithm 2 under a deterministic channel environment (i.e., $\mathbf{G}^t = \mathbf{G}$, $\eta^t = \eta$) and a constant step-size $\alpha^t = \alpha$. Then,*

1. *If $0 < \alpha < \rho_1$, then $\|\mathbf{p} - \mathbf{p}^*\|^2_2 = O(\kappa^{-T})$ under Variant A, for some $\kappa > 1$.*
2. *If $0 < \alpha < \rho_2$, then $\|\mathbf{p} - \mathbf{p}^*\|^2_2 = O(\kappa^{-T})$ under Variant B for some $\kappa > 1$.*

Remark 9. With some algebra, Corollary 1 follows as a consequence of the framework we have presented so far. In particular, Variant A and Variant B are multiagent EGD for the designed games $(\mathcal{N}, \mathcal{P}, \{u_i\}^N_{i=1})$ and $(\mathcal{N}, \mathcal{P}, \{\tilde{u}_i\}^N_{i=1})$, respectively, where $\alpha^t = \alpha$. To see this, note 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 $\frac{\partial \tilde{u}_i(\mathbf{p})}{\partial p_i} = -(p_i - r^*_i \frac{\sum_{j\neq i} G_{ij}p_j+\eta_i}{G_{ii}})$. Further, from Equation (21), it is clear that $v(\cdot)$ is ρ_1-Lipschitz because $\|v(\mathbf{p}) - v(\tilde{\mathbf{p}})\|_2 = \|M^1\mathbf{p} - M^1\tilde{\mathbf{p}}\|_2 \leq \lambda_{\max}(M^1)\|\mathbf{p} - \tilde{\mathbf{p}}\|_2$. Similarly, per Equation (22), $v(\cdot)$ is ρ_2-Lipschitz. Consequently, per Lemma 4 and Remark 6, we have the convergence results.

5.2. Channel Feasibility in Stochastic Environments

Before we move on to establish the theoretical guarantees in the stochastic channel environments, we need a notion that characterizes channel feasibility in such cases. Because 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 feasible on average (and hence can be feasible sometimes and infeasible some other times). The next definition formalizes it:

Definition 5. 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 and commonly occurring 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

$$\tilde{W}_{ij} := \begin{cases} 0, & i = j \\ r^*_i \frac{\mathbf{E}[G_{ij}]}{\mathbf{E}[G_{ii}]}, & i \neq j. \end{cases} \tag{23}$$

We have

$$\mathbf{E}\left[\frac{G_{ij}}{G_{ii}}\right] = \mathbf{E}\left[G_{ij}\right]\mathbf{E}\left[\frac{1}{G_{ii}}\right] \geq \frac{\mathbf{E}\left[G_{ij}\right]}{\mathbf{E}\left[G_{ii}\right]}, \tag{24}$$

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. Because each entry in both matrices is nonnegative, 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 (24) implies $\mathbf{u}\mathbf{E}[W]\mathbf{u} \geq \mathbf{u}\tilde{W}\mathbf{u}$, which then leads to

$$\begin{aligned}\lambda_{\max}(\mathbf{E}[W]) &= \max_{\mathbf{u}\in\mathbf{R}^N_+:\|\mathbf{u}\|_2=1} \mathbf{u}\mathbf{E}[W]\mathbf{u} \\ &\geq \max_{\mathbf{u}\in\mathbf{R}^N_+:\|\mathbf{u}\|_2=1} \mathbf{u}\tilde{W}\mathbf{u} = \lambda_{\max}(\tilde{W}).\end{aligned}$$

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

Remark 10. These two types of mean-feasible environments correspond to the respective environment under which Variant A and Variant B of the proposed algorithm admit theoretical performance guarantees. As it turns out, Type-I mean-feasible is required for Variant A and Type-II mean-feasible is required for Variant B. Per the previous lemma, Type-II mean-feasible is a slightly stronger condition than Type-I mean-feasible: this is to be expected because Variant B of the algorithm requires less information in choosing the power iterates than that of Variant A.

5.3. Convergence of Robust Power Control: Stability and Optimality

We are now ready to state the main result. We obtain this result by casting the proposed power control algorithms in the multiagent noisy EGD learning framework.

Theorem 4. *Given a stochastic channel* $(\mathbf{G}, \eta)$ *(or equivalently* (W, γ)*) according to Assumption 1: let* γ_0 *in Algorithm 2 be chosen such that* $\gamma_0 > \frac{1}{-\lambda_{\max}(\frac{1}{2}(W-\mathbf{I})+\frac{1}{2}(W^T-\mathbf{I}))}$.

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* $\mathbf{E}[\|\mathbf{P}^t \to \mathbf{p}^*\|_2^2] = O(\frac{1}{T})$*, where* $\mathbf{P}^t$ *is given by Variant A in Algorithm 2.*
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* $\mathbf{E}[\|\mathbf{P}^t \to \mathbf{p}^*\|_2^2] = O(\frac{1}{T})$*, where* $\mathbf{P}^t$ *is given by Variant B in Algorithm 2.*

Remark 11. Two things to note here. First, the mean-square convergence to a constant joint transmission power in the presence of persistent stochastic channel fluctuations is a manifestation of the stability of the proposed algorithms. 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 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. Because $(\mathbf{G}, \eta)$ (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

$$\begin{aligned}\frac{\partial u_i(\mathbf{p})}{\partial p_i} &= -\mathbf{E}\left[\left(G_{ii}p_i - r_i^*\left(\sum_{j\neq i} G_{ij}p_j + \eta_i\right)\right)\right] \\ &= -\left(\mathbf{E}[G_{ii}]p_i - r_i^*\left(\sum_{j\neq i}\mathbf{E}[G_{ij}]p_j + \mathbf{E}[\eta_i]\right)\right), \qquad (25)\end{aligned}$$

$$\begin{aligned}\frac{\partial \tilde{u}_i(\mathbf{p})}{\partial p_i} &= -\mathbf{E}\left[\left(p_i - r_i^*\frac{\sum_{j\neq i} G_{ij}p_j + \eta_i}{G_{ii}}\right)\right] \\ &= -\left(p_i - r_i^*\sum_{j\neq i}\left(\mathbf{E}\left[\frac{G_{ij}}{G_{ii}}\right]p_j + \mathbf{E}\left[\frac{\eta_i}{G_{ii}}\right]\right)\right) \\ &= -\left(p_i - \sum_{j\neq i}(\mathbf{E}[W_{ij}]p_j + \mathbf{E}[\gamma_i])\right). \qquad (26)\end{aligned}$$

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 2 can be equivalently written as

$$\begin{aligned}\tilde{X}_i^{t+1} &= X_i^t - \frac{\gamma_0}{t}\left\{\left(\mathbf{E}[G_{ii}] + \tilde{G}_{ii}^t\right)P_i^t - r_i^* \right.\\ &\quad \left.\times\left(\sum_{j\neq i}\left(\mathbf{E}\left[G_{ij}\right] + \tilde{G}_{ij}^t\right)P_j^t + \mathbf{E}\left[\eta_i\right] + \tilde{\eta}_i^t\right)\right\}, \qquad (27)\\ &= X_i^t - \frac{\gamma_0}{t}\left\{\mathbf{E}[G_{ii}]P_i^t - r_i^*\left(\sum_{j\neq i}\mathbf{E}\left[G_{ij}\right]P_j^t\right) + \mathbf{E}\left[\eta_i\right]\right.\\ &\quad \left.+\left\{\tilde{G}_{ii}^t P_i^t - r_i^*\sum_{j\neq i}\tilde{G}_{ij}^t P_j^t + \tilde{\eta}_i^t\right\}\right\}. \qquad (28)\end{aligned}$$

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. Because $\mathcal{P}$ 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 7. This implies that Algorithm 2 is a special case of Algorithm 4. Further, because the channel is Type-I mean-feasible, Lemma 4 implies that $\mathcal{G}^1$ is a WSM game with $\mathbf{p}^*$ being the unique Nash equilibrium. The result therefore follows by directly applying Theorem 3.

The second claim follows from a similar line of reasoning: the only thing to note here is that the bounded second moments assumption holds because $G_{ii} \geq \underline{g}_i$ a.s.. ■

6. Conclusion

We close with a few remarks. First, although we have focused on stochastic iid environments in this paper, the designed algorithm given in Algorithm 2 can still operate in an arbitrary time-varying environment. We believe our convergence results would generalize to the stationary and ergodic environment case, although making that fully rigorous is beyond the scope of the paper and requires new analysis techniques; we hence leave that for future work. Further, the recent empirical work (Ward et al. 2018) indicates that variants of the dual averaging power control algorithm can be made robust to delayed feedback, broadening its applicability even further. Second, our theoretical investigation on multiagent EGD learning with imperfect first-order feedback 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?* We aim to obtain quantitative convergence results for other no-regret learning algorithms in different classes of games in future work, with a particular goal in understanding when would the rate $O(\frac{1}{T})$ be achievable. Another interesting theoretical direction to take is to obtain convergence results when only zeroth-order feedback is obtained. Flaxman et al. (2005) characterized regret guarantees for online gradient descent in such cases. Studying the convergence issue in the multiagent setup would be interesting.

Finally, as mentioned in the introduction, the landscape of power management is vast and includes many other objectives. We believe that this game-design approach will be fruitful in those other problems: in order to design distributed algorithms that converge to some optimal/desired action, one can design a WSM game whose unique Nash equilibrium corresponds to the optimal action and thereby deriving an algorithm immediately via the online learning algorithm. Consequently, in this approach, EGD, and more broadly any no-regret online learning algorithm, can be viewed as a meta-algorithm that can be instantiated via the design of specific games.

Acknowledgments

The authors are grateful to the associate editor and the three reviewers for their extremely valuable comments and constructive feedback that led to significant improvement of the current paper.

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] Motivated by the random nature of network 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).

[3] In practice, only access to some aggregated statistics of the power used by all the other transmitters is feasible. The statistics typically take the form of some function of all the powers used by the transmitters, such as the one given in Algorithm 2.

References

Alpcan T, Basar T, Dey S (2006) A power control game based on outage probabilities for multicell wireless data networks. *IEEE Trans. Wireless Comm.* 5(4):890–899.

Alpcan T, Başar T, Srikant R, Altman E (2002) CDMA uplink power control as a noncooperative game. *Wireless Networks* 8(6): 659–670.

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

Antipin A (2002) Gradient approach of computing fixed points of equilibrium problems. *J. Global Optim.* 24(3):285–309.

Arrow KJ, Hurwicz L (1960) Stability of the gradient process in n-person games. *J. Soc. Indust. Appl. Math.* 8(2):280–294.

Au-Yeung KY, Robertson T, Hafezi H, Moon G, DiCarlo L, Zdeblick M, Savage G (2010) A networked system for self-management of drug therapy and wellness. *Wireless Health* (ACM, New York), 1–9.

Balandat M, Krichene W, Tomlin C, Bayen A (2016) Minimizing regret on reflexive Banach spaces and learning Nash equilibria in continuous zero-sum games. Preprint, submitted June 3, https://arxiv.org/abs/1606.01261.

Bradley J, Barbier J, Handler D (2013) Embracing the internet of everything to capture your share of $14.4 trillion. White Paper, Cisco Systems, Inc., San Jose, CA.

Byrne C, Lim CL (2007) The ingestible telemetric body core temperature sensor: a review of validity and exercise applications. *British. J. Sports Medicine* 41(3):126–133.

Candogan UO, Menache I, Ozdaglar A, Parrilo PA (2010) Near-optimal power control in wireless networks: A potential game approach. *Proc. 29th IEEE Conf. Comput. Comm.* (IEEE, Piscataway, NJ), 1–9.

Cesa-Bianchi N, Lugosi G (2006) *Prediction, Learning, and Games* (Cambridge University Press, Cambridge, UK).

Chiang M, Hande P, Lan T, Tan CW (2008) Power control in wireless cellular networks. *Foundations Trends® Networking* 2(4):381–533.

Chung KL (1954) On a stochastic approximation method. *Ann. Math. Statist.* 25(3):463–483.

Cui T, Chen L, Low SH (2008) A game-theoretic framework for medium access control. *IEEE J. Selected Areas Comm.* 26(7): 1116–1127.

Deakin M (2013) Smart Cities: Governing, Modelling and Analysing the Transition (Taylor & Francis, London).

El Gamal A, Mammen J, Prabhakar B, Shah D (2006a) Optimal throughput-delay scaling in wireless networks-part i: The fluid model. *Information Theory IEEE Trans. Inform. Theory* 52(6): 2568–2592.

El Gamal A, Mammen J, Prabhakar B, Shah D (2006b) Optimal throughput-delay scaling in wireless networks-part i: The fluid model. *Information Theory IEEE Trans. Inform. Theory* 52(6): 2568–2592.

Eryilmaz A, Modiano E, Ozdaglar A (2006) Randomized algorithms for throughput-optimality and fairness in wireless networks. *Proc. 45th IEEE Conf. Decision Control* (IEEE, Piscataway, NJ), 1936–1941.

Eryilmaz A, Ozdaglar A, Médard M, Ahmed E (2008) On the delay and throughput gains of coding in unreliable networks. *IEEE Trans. Inform. Theory* 54(12):5511–5524.

Facchinei F, Pang J-S (2007) *Finite-Dimensional Variational Inequalities and Complementarity Problems* (Springer Science & Business Media).

Fan X, Alpcan T, Arcak M, Wen TJ, Başar T (2006) A passivity approach to game-theoretic CDMA power control. *Automatica* 42(11):1837–1847.

Flåm SD (2002) Equilibrium, evolutionary stability and gradient dynamics. *Internat. Game Theory Rev.* 4(4):357–370.

Flaxman AD, Kalai AT, Kalai AT, McMahan HB (2005) *Online convex optimization in the bandit setting: Gradient descent without a gradient.Proc. 16th Annual ACM-SIAM Sympos. Discrete Algorithms* (Society for Industrial and Applied Mathematics, Philadelphia), 385–394.

Foschini GJ, Miljanic Z (1993) A simple distributed autonomous power control algorithm and its convergence. *IEEE Trans. Vehicular Tech.* 42(4):641–646.

Gitzenis S, Bambos N (2002) Power-controlled data prefetching/caching in wireless packet networks. *Proc. 21st Annual Joint Conf. IEEE Comp. Comm. Soc.*, vol. 3 (IEEE, Piscataway, NJ), 1405–1414.

Goldsmith A (2005) *Wireless Communications* (Cambridge University Press, Cambridge, UK).

Gopalan A, Caramanis C, Shakkottai S (2015) Wireless scheduling with partial channel state information: Large deviations and optimality. *Queueing Systems* 80(4):293–340.

Han Z, Niyato D, Saad W, Ar TB, Rungnes AH (2014) *Game Theory in Wireless and Communication Networks* (Cambridge University Press, Cambridge, UK).

Han Z, Niyato D, Saad W, Başar T, Hjørungnes A (2011) *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications* (Cambridge University Press, Cambridge, UK).

Hazan E (2016) *Introduction to online convex optimization*. *Foundations Trends Optim.* 2(3–4):157–325.

Kar S, Moura JMF, Ramanan K (2012) Distributed parameter estimation in sensor networks: Nonlinear observation models and imperfect communication. *IEEE Trans. Inform. Theory* 58(6):3575–3605.

Krichene S, Krichene W, Dong R, Bayen A (2015) Convergence of heterogeneous distributed learning in stochastic routing games. *53rd Annual Allerton Conf. Comm., Control, Comput. (Allerton)* (IEEE, Piscataway, NJ), 480–487.

Lam K, Krichene W, Bayen A (2016) On learning how players learn: Estimation of learning dynamics in the routing game. *2016 ACM/IEEE 7th Internat. Conf. Cyber-Physical Systems* (IEEE, Piscataway, NJ), 1–10.

Li N, Marden JR (2013) Designing games for distributed optimization. *IEEE J. Selected Topics Signal Processing* 7(2):230–242.

Menache I, Ozdaglar A (2010) Network Games: Theory, Models, and Dynamics(Morgan & Claypool, San Rafael, CA).

Mertikopoulos P, Zhou Z (2019) Learning in games with continuous action sets and unknown payoff functions. *Math. Programming* 173(1-2):465–507.

Mitra D (1994) An asynchronous distributed algorithm for power control in cellular radio systems. Holtzman JM, Goodman DJ, eds. *Wireless and Mobile Communications*, Springer Series in Engineering and Computer Science, vol. 277 (Springer, Boston), 177–186.

Nemirovski A, Juditsky A, Lan G, Shapiro A (2009) Robust stochastic approximation approach to stochastic programming. *SIAM J. Optim.* 19(4):1574–1609.

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

Reddy AA, Banerjee S, Gopalan A, Shakkottai S, Ying L (2012) On distributed scheduling with heterogeneously delayed network-state information. *Queueing Systems* 72(3-4):193–218.

Reddy A, Shakkottai S, Ying L (2008) Distributed power control in wireless ad hoc networks using message passing: Throughput optimality and network utility maximization. *42nd Annual Conf. Inform. Sci. Systems* (IEEE, Piscataway, NJ), 770–775.

Rosen BJ (1965) Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica* 33(3):520–534.

Seferoglu H, Lakshmikantha A, Ganesh A, Key P (2008) Dynamic decentralized multi-channel MAC protocols. *Inform. Theory Appl. Workshop* (IEEE, Piscataway, NJ), 100–110.

Shalev-Shwartz S (2012) Online learning and online convex optimization. *Foundations Trends Machine Learn.* 4(2):107–194.

Tan CW (2014) Wireless network optimization by Perron-Frobenius theory. *48th Annual Conf. Inform. Sci. Systems (CISS)* (IEEE, Piscataway, NJ), 1–6.

Tan CW (2015) *Wireless network optimization by Perron-Frobenius theory.Foundations Trends Networking* 9(2-3):107–218.

Ulukus S, Yates RD (1998) Stochastic power control for cellular radio systems. *IEEE Trans. Comm.* 46(6):784–798.

Ward A, Zhou Z, Mertikopoulos P, Bambos N (2018) Power control with random delays: Robust feedback averaging. *2018 IEEE Conf. Decision Control (CDC)* (IEEE, Piscataway, NJ), 7040–7045.

Weeraddana PC, Codreanu M, Latva-aho M, Ephremides A, Fischione C (2012) Weighted sum-rate maximization in wireless networks: A review. *Foundations Trends Networking* 6(1–2):1–163.

Yates RD (1996) A framework for uplink power control in cellular radio systems. *IEEE J. Selected Areas Commun.* 13(7):1341–1347.

Zhou Z, Bambos N (2015) Wireless communications games in fixed and random environments. *54th IEEE Conf. Decision Control (CDC)* (IEEE, Piscataway, NJ), 1637–1642.

Zhou Z, Bambos N, Glynn P (2016a) Dynamics on linear influence network games under stochastic environments. Zhu Q, Alpcan T, eds. *7th Internat. Conf. Decision Game Theory Security*, vol. 9996 (Springer-Verlag, Berlin, Heidelberg), 114–126.

Zhou Z, Bambos N, Glynn P (2018a) Deterministic and stochastic wireless network games: Equilibrium, dynamics, and price of anarchy. *Oper. Res.* 66(6):1498–1516.

Zhou Z, Glynn P, Bambos N (2016b) Repeated games for power control in wireless communications: Equilibrium and regret. *55th IEEE Conf. Decision Control (CDC)* (IEEE, Piscataway, NJ), 3603–3610.

Zhou Z, Miller D, Glynn P, Bambos N (2016c) A stochastic stability characterization of the Foschini-Miljanic algorithm in random wireless networks. *2016 IEEE Global Commun. Conf. (GLOBECOM)* (IEEE, Piscataway, NJ).

Zhou Z, Mertikopoulos P, Bambos N, Glynn PW, Tomlin C (2017a) Countering feedback delays in multi-agent learning. von Luxburg U, Guyon I, eds. *Proc. 31st Internat. Conf. Neural Inform. Processing Systems* (Curran Associates, Red Hook, NY), 6171–6181.

Zhou Z, Mertikopoulos P, Moustakas AL, Bambos N, Glynn P (2017b) Mirror descent learning in continuous games. *56th Annual IEEE Conf. Decision Control (CDC)* (IEEE, Piscataway, NJ), 5776–5783.

Zhou Z, Mertikopoulos P, Athey S, Bambos N, Glynn PW, Yinyu Y (2018b) Learning in games with lossy feedback. Bengio S, Wallach HM, eds. *Proc. 32nd Internat. Conf. Neural Inform. Processing Systems* (Curran Associates, Red Hook, NY), 5140–5150.

Zhengyuan Zhou is an assistant professor in the Stern School of Business at New York University. His research interests lie in learning, optimization, game theory, and stochastic systems.

Panayotis Mertikopoulos has been a CNRS researcher at the Laboratoire d'Informatique de Grenoble since 2011. His research interests lie in the areas of game theory, online and stochastic optimization, and algorithmic learning, as well as in their applications to operations research and wireless networks.

Aris Moustakas is currently a faculty member with the Physics Department at the National Kapodistrian University of Athens, Greece. His research interests include statistical physics, random matrix theory and wireless communications.

Nicholas Bambos is the current chair and the R. Weiland Professor of the Department of Management Science and Engineering at Stanford University. His research interests lie in communication networks, queueing theory, and applied probability.

Peter Glynn is the Thomas Ford Professor of Engineering in the Department of Management Science and Engineering at Stanford University. His research interests lie in simulation, computational probability, queueing theory, statistical inference for stochastic processes, and stochastic modeling.