Power Control in Wireless Networks via Dual Averaging

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

Abstract— **We propose a simple, novel and distributed power control algorithm that efficiently incorporates past information and regulates power to achieve better stability. We provide an analytical framework that examines in depth the properties of the proposed algorithm, and we establish its convergence for both deterministic, time-invariant channels and stochastic, time-varying environments. In the former (deterministic) case, if the channel is feasible, the proposed dual averaging algorithm converges to the optimal power vector. More importantly, in the latter (stochastic) case, we show that if the channel is feasible on average, then the proposed dual averaging power control algorithm converges almost surely to a deterministic optimal power vector, even though existing power control algorithms (such as Foschini–Miljanic) may fail to converge (to a distribution) altogether. Finally, we provide a set of simulations that further demonstrate various interesting and desirable properties of the proposed algorithm.**

I. Introduction

Distributed power control has been a core issue in wireless communications ever since the early stages of wireless network design. It is widely agreed that distributed power control has emerged as the predominant paradigm for good reasons: centralized coordination is extremely difficult in typical large-scale wireless networks; the exchange of information required might cause significant power expenditures on its own; a single point of failure in a centralized power allocator can have devastating effects; and the list goes on [1], [2]. Consequently, much effort has been devoted to designing distributed power control algorithms that achieve good performance (quality-of-service) guarantees.

The well-known Foschini-Miljanic (FM) algorithm [3], which has greatly influenced the subsequent literature [4], [5], [6], is a simple and elegant distributed power control scheme with strong convergence properties. In the past two decades or so, game theory and utility-based models [7], [8], [9], [10] have also been employed to analyze and design distributed power control algorithms. However, all those distributed algorithms tend to rely implicitly only on the previous power iterate when selecting the power for the current iteration. In some respects, this is easily understood: the resulting power control algorithm is simple to use and implement, and it does not need to keep track of all the past histories (a real memory constraint in wireless transmission devices). However, an important and well-known downside is that this approach tends to be unstable because a single power iterate is the sole power selection criterion; see Section II-C for a discussion. As such, the implicit use of all past iterates holds great promise for the overall stability of the algorithm, as the influence of the last iterate cannot have dominating impact of the previous iterates (provided they are utilized in an intelligent, memory-efficient way).

Our goal is to provide a distributed algorithm satisfying the above desiderata. This task faces at least two challenges from a practical perspective. First, such an algorithm cannot assume that each transmitter has access to the power used by all the other transmitters, as such an information exchange is infeasible in practice. This dictates that a transmitter can only use aggregate information, such as measurements of their SINR and/or total interference and noise (i.e. information that can be sensed by each individual link). Second (and 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 other words, this information must be exploited in an implicit, parsimonious manner.

A. Our Contributions

Our contributions are threefold. First, we propose a simple, novel distributed power control algorithm which we call dual averaging and which makes efficient use of all past information in order to regulate power in a stable manner. The dual averaging algorithm proposed here satisfies the above-mentioned requirements. In particular, the past power iterates information are represented in the most parsimonious form possible: it takes only a constant amount of memory independent of transmission horizon. In fact, the amount of memory it takes is no larger than that of the FM algorithm, even though the latter does not use any previous iterate information. Further, the dual averaging algorithm has the desired baseline convergence and performance guarantee: in the deterministic and time-invariant case, if the channel is feasible, then dual averaging converges to the optimal power vector (Theorem 3).

Second, we extend the model and analysis to the stochastic and time-varying case, where we consider a general channel environment that is driven by a stationary and ergodic stochastic process. We then establish the surprising and highly desirable property of the proposed dual averaging algorithm: if the channel is feasible on average, then the

Z. Zhou, S. Mehdian, N. Bambos and P. Glynn are with the Department of Electrical Engineering and Department of Management Science and Engineering, Stanford University, CA, 94305, USA.

P. Mertikopoulos is with the French National Center for Scientific Research (CNRS) and with Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG, F-38000, Grenoble, France

A. Moustakas is with Department of Physics, University of Athens and Institute of Accelerating Systems and Applications (IASA), Athens, Greece.

dual averaging algorithm converges almost surely to a deterministic optimal power vector (Theorem 4). This means that how the channel environment fluctuates from time to time (the particular underlying distribution) is of no importance, so long as on average the channel is feasible. And when the channel is feasible on average, dual averaging will stabilize almost surely at the optimal power vector as if the channel were deterministic and time-invariant with the channel parameters given by means of the corresponding stochastic channel. This comes in sharp contrast with FM and demonstrates thus dual averaging's superior stability: on one hand, even if the average of the stochastic channel is feasible, FM may fail to converge in any reasonable sense; on the other hand, even in average-feasible channel cases where FM does converge, it converge at best to a (stationary) distribution, as opposed to a fixed power profile, which is highly desirable in wireless applications. Furthermore, precisely due to this instability, no performance guarantees can be established on this stationary distribution (to which FM converges), whereas dual averaging yields the optimal power profile in the mean sense.

Third, we provide a set of simulations that demonstrates various interesting properties of the proposed dual averaging power control algorithm. We consider both the deterministic, time-varying channel case, where further examine the impact of maximum power bounds on convergence, and the stochastic, time-varying channel case, where characterize the behavior of dual averaging under different channel distributions. Discussions and comparisons with FM are also included.

II. Model, Background, Motivation

We consider the classical power control setting in wireless communications [11], [12]. After introducing the problem and some minimum notation in II-A, we discuss in Section II-B the well-known Foschini-Miljanic (FM) power control algorithm [3]. This discussion will provide, both quantitatively and qualitatively, an account of some of the drawbacks of the FM algorithm, which then serves as the motivation of the paper, as outlined in Section II-C.

A. Power Control in Wireless Communications

Consider a wireless network of N communication links, each comprised of a transmitter and an intended receiver. The power vector for transmission is denoted by $\mathbf{p} = (p_1, \dots, p_N)$ where p_i is the power used by the transmitter of link i. Throughout the paper, it will be assumed that $\mathbf{p} \in \mathbb{R}_+^N$, where $\mathbb{R}_+$ denotes non-negative reals.

The most commonly used measure of link service quality is the signal to interference and noise ratio (SINR). Given a power vector $\mathbf{p}$, link i's SINR $R_i(\mathbf{p})$ is given by the following ratio:

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

where η_i is the thermal noise associated with the receiver of link i and $G_{ij} \geq 0$ is[1] the power gain between transmitter j and receiver i, representing the interference caused to receiver i by transmitter j per unit transmission power used. For notational convenience, we collect all the power gains G_{ij} into the gain matrix $\mathbf{G}$ and all the noises into η. The power gain matrix $\mathbf{G}$ depends on the underlying topology of the communicating wireless links. Furthermore, each link has a predetermined target SINR threshold $r_i^* > 0$: the corresponding acceptable link service quality threshold.

Two quantities that will be useful later are:

1) The re-weighted gain matrix W, where

$$W_{ij} := \begin{cases} 0, & i = j \\ r_i^* \frac{G_{ij}}{G_{ii}}, & i \neq j, \end{cases} \quad (2)$$

for $i, j \in \{1, 2, \dots, N\}$.

2) The re-weighted noise vector γ, where

$$\gamma_i = R_i^* \frac{\eta_i}{G_{ii}}, i \in \{1, 2, \dots, N\}. \quad (3)$$

Under the aforementioned notation, the standard power control problem [11], [12] can then be formulated to find a power assignment $\mathbf{p}$, such that the quality of service constraints hold:

$$r_i(\mathbf{p}) \geq r_i^*, \forall i. \quad (4)$$

In order to find a power vector $\mathbf{p}$ that satisfies the quality of service constraints, such a $\mathbf{p}$ must exist in the first place: the notion of *channel feasibility*, formalized in the next definition, characterizes such scenarios.

Definition 1: The channel, given by $\mathbf{G}, \eta$ is said to be feasible with respect to a target SINR vector $r^* = (r_1^*, \dots, r_N^*)$ if there exists a $\mathbf{p}$ satisfying Equation 4. The channel is otherwise said to be infeasible.

A simple necessary and sufficient condition (given in [3]) exists on deciding when a channel is feasible (note that it does not depend on the noise vector η):

Theorem 1: A channel is feasible with respect to r^* if and only if the largest[2] 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 power vector $\mathbf{p}^*$ that satisfies the quality of service constraints in Equation 4, as given by the following fact:

Fact 1: Given a feasible channel $\mathbf{G}, \eta$, the power vector $\mathbf{p}^* = (\mathbf{I} - W)^{-1}\gamma$ satisfies the quality of service constraints given in Equation 4. Furthermore, it is component-wise strictly positive and the unique vector that satisfies the following property: if $\mathbf{p}$ is any vector satisfying Equation 4, then $\mathbf{p}^* \leq \mathbf{p}$, where inequality is component-wise.

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

Definition 2: $\mathbf{p}^*$ defined in Fact 1 is called the optimal power vector.

[1]We further assume that $G_{ii} > 0$ for otherwise transmission for link i is meaningless.

[2]As noted in [13], [12], 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.

B. Foschini-Miljanic Power Control Algorithm

We now present the well-known FM power control algorithm (Algorithm 1), which finds the optimal power vector (if one exists). Following the standard convention in wireless communications [14], the transmission power p_i for transmitter i is assumed to lie in a compact interval $[0, p_i^{\max}]$. Therefore, the power vector $\mathbf{p}$ is constrained to lie in the feasible support set $\mathcal{P} = \prod_{i=1}^{N}[0, p_i^{\max}]$. We shall adopt this assumption for the rest of the paper. To accommodate this maximum power constraint, the FM algorithm can be described as follows:

Algorithm 1 FM Algorithm: Bounded Power Support

```
Each link i chooses an initial power p_i^0.
for t = 0, 1, 2, ... do
    for i = 1, ..., N do
        p_i^{t+1} = min(p_i^t r_i^* / r_i(p^t), p_i^max)
    end for
end for
```

By a monotonicity argument and a result in [3], one has the following convergence characterization of the FM algorithm:

Theorem 2: Let the channel $\mathbf{G}, \eta$ be constant 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 power iterate $\mathbf{p}^t$ in Algorithm 1 converges to the optimal power vector $\mathbf{p}^*$, irrespective of the initial point $\mathbf{p}^0$.

C. Motivation of the Paper

Although the FM algorithm enjoys good convergence properties (cf. Theorem 2), it is in a certain sense not very stable because it uses only the last power iterate $\mathbf{p}^t$. This point is particularly manifested when the underlying channel is stochastic and time-varying. In this case (i.e. when $\{\mathbf{G}^t\}_{t=0}^{\infty}, \{\eta^t\}_{t=0}^{\infty}$ are stochastic processes), the power iterate P^t generated by FM will be a random variable satisfying:

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

where $\Pi_{\mathcal{P}}(y) = \arg\min_{x\in\mathcal{P}} \|x-y\|_2$ and W^t and γ^t are the random re-weighted gain matrix and the random re-weighted noise at iteration t, respectively.

Note here that the power iterate P^t will at best (under certain conditions of the channel) converge weakly to a stationary distribution (as opposed to a deterministic power vector). This is to be expected since the FM algorithm is sensitive to the last power iterate. Specifically, [15] considers the **iid** channel environment ($\{\mathbf{G}^t\}_{t=0}^{\infty}$ and $\{\eta^t\}_{t=0}^{\infty}$ are each assumed to be **iid** in which case $\{P^t\}_{t=0}^{\infty}$ from the FM algorithm forms a Markov chain) and studies sufficient conditions of the channel under which P^t converges weakly to a stationary distribution.

However, convergence to a limiting stationary distribution is not the most desirable case for several reasons. First, this is a form of instability: even in stationarity, the users' power vector still fluctuates. Second, such fluctuations (i.e. the stationary distribution, if one exists) depend on the underlying distribution of the channel environment: a different stochastic channel process of the channel environment yields a different stationary distribution of power. Even though this is to be expected (mathematically at least), it also indicates that that FM is sensitive to the environment in which it operates. Finally, precisely due to this instability, it is hard to obtain long-run performance guarantees on the users' average SINR when operating in stochastic and time-varying environments.

As mentioned before, the fact that FM uses only the last power iterate is the reason that one cannot hope to do better than this type of weak stochastic convergence results for FM. This naturally calls for the design of a new algorithm that addresses the three above-mentioned issues. Specifically, we are led to the following questions:

- Would it be possible to incorporate all the past power iterates to synthesize a power control scheme that converges almost surely to a fixed power profile even in the presence of unpredictable stochastic fluctuations??
- Would this power profile be robust to the distribution of the underlying channel process?
- Can we obtain performance guarantees of this algorithm?

As we shall see in the next section, the answers to the above questions are all in the affirmative.

III. Dual Averaging Power Control Algorithm

In this section, we propose a novel distributed power control algorithm that aims to utilize all the power iterates in the past so as to achieve better stability. The design of such a distributed algorithm (that uses information on past power iterates) faces at least two challenges from a practical perspective. First, such an algorithm cannot assume that each transmitter has access to the power used by all other users, as such communications is infeasible in practice. This dictates that a transmitter can only use aggregate information such as SINR and/or total interference (i.e. information that can be sensed by each individual link). Second, more stringently, one should not expect a transmitter to store all the past information that is available to it, which includes its own past transmission powers, past SINRs etc. This additional constraint, if met, is highly desirable in practice because such bookkeeping is in general infeasible for memory-constrained wireless transmitters. This second constraint further implies that it is necessary to incorporate all such information in a parsimonious, memory-sensitive way.

As we shall see next, the Dual Averaging power control algorithm (see Remark 1 for an explanation of the naming of the algorithm) proposed here satisfies these two constraints. Further, the past power iterates information are represented 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.

A. Deterministic and Time-Invariant Channel

We start with the description of the algorithm in the deterministic and time-invariant channel case. In this case, Algorithm 2 provides a formal description of the Dual Averaging algorithm.

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

```
1: Each link i chooses an initial y_i^0.
2: for t = 0, 1, 2, ... do
3:    for i = 1, ..., N do
4:       p_i^t = Π_P(y_i^t)
5:       y_i^{t+1} = y_i^t − (1/t)(G_ii p_i^t − r_i^*(Σ_{j≠i} G_ij p_j^t + η_i))
6:    end for
7: end for
```

Remark 1: First, note that y_i^t's maintain 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, 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", as it is reminiscent of the well-known dual averaging optimization algorithm proposed in [16]. 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 $\mathcal{P}$, the dual space that $\mathbf{y}^t$ is in, need not be. 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 (both of which can be detected by the receiver).

B. Stochastic and Time-Varying Channel

We extend here the model to the stochastic and time-varying channel where same dual averaging power control algorithm operates. We adopt the following stochastic model of a wireless communications network where $(\mathbf{G}^t, \eta^t)$ is a stationary and ergodic stochastic process drawn from a arbitrary (discrete or continuous, bounded or unbounded) support on $\mathbb{R}_+^{N\times N} \times \mathbb{R}_+^N$, where the following two (rather weak) assumptions are satisfied:

1) Finite mean: $\forall i, j, \forall t, 0 < \mathbf{E}[G_{ij}^t] = \bar{G}_{ij} < \infty, \mathbf{E}[\eta_i^t] = \bar{\eta}_i < \infty$.
2) Finite variance: $\forall i, j, \forall t, \mathbf{Var}[G_{ij}^t] < \infty, \mathbf{Var}[\eta_i^t] < \infty$.

By stationarity, the mean and variance across time steps are necessarily the same. Note that this model is quite general and includes as a special case the **iid** channel studied in [15] as well as the stationary Markov process commonly used in the literature. Note in particular, that (G_{ij}^t, η_k^t) can be arbitrarily correlated with $(G_{i'j'}^t, \eta_{k'}^t)$.

Algorithm 3 gives the description in the stochastic case. Here we use the upper case for gradient and power to highlight the fact that all the quantities are now random variables, while the dual averaging algorithm itself stays the same.

Algorithm 3 Dual Averaging Algorithm: Stochastic and Time-Varying Channel

```
Each link i chooses an initial Y_i^0.
for t = 0, 1, 2, ... do
   for i = 1, ..., N do
      P_i^t = Π_P(Y_i^t)
      Y_i^{t+1} = Y_i^t − (1/t)(G_ii^t P_i^t − r_i^*(Σ_{j≠i} G_ij^t P_j^t + η_i^t))
   end for
end for
```

IV. Convergence and Performance Guarantees

Here we characterize the convergence and performance results for the proposed dual averaging power control algorithm. Due to space limitation, we can only sketch the main elements behind the proofs.

A. Deterministic and Time-Invariant Channel Case

As indicated by the following theorem, in the deterministic and time-invariant channel case, dual averaging enjoys the same convergence and optimality guarantee as FM.

Theorem 3: Let the channel $(\mathbf{G}, \eta)$ be feasible with the optimal power vector $\mathbf{p}^* \in \mathcal{P}$. Then $\mathbf{p}^t \to \mathbf{p}^*$ as $t \to \infty$, where $\mathbf{p}^t$ is given in Algorithm 2.

Proof Elements: There are several ingredients to the proof. First, we define Fenchel coupling $F : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$:

$$F(\mathbf{p}, \mathbf{y}^t) = h(\mathbf{p}) + h^*(\mathbf{y}^t) - \langle \mathbf{y}^t, \mathbf{p} \rangle,$$

where $h(\mathbf{p}) = \sum_{i=1}^N h_i(p_i), h^*(\mathbf{y}) = \sum_{i=1}^N h_i^*(y_i), h_i(p_i) = \frac{1}{2}p_i^2,\ h_i^*(y_i) = \max_{p_i \in [0, p_i^{\max}]}\{p_i y_i - h_i(p_i)\}$ and $\langle \cdot, \cdot \rangle$ denotes inner product.

Using $F(\mathbf{p}, \mathbf{y}^t)$ as a Lyapunov function, we establish that the iterate $\mathbf{p}^t$ will enter the neighborhood $\tilde{B}(\mathbf{p}^*, \delta) \triangleq \{\Pi_{\mathcal{P}}(\mathbf{y}) \mid F(\mathbf{p}^*, \mathbf{y}) < \delta\}$ infinitely often. Further, by a detailed analysis, one can then show that, starting from iteration t, $\mathbf{p}^t$ will remain in $\tilde{B}(\mathbf{p}^*, \delta)$. Since this is true for any $\delta > 0$, we have $F(\mathbf{p}^*, \mathbf{y}^t) \to 0$ as $t \to \infty$. Per the properties of Fenchel coupling, this leads to that $\|\Pi_{\mathcal{P}}(\mathbf{y}^t) - \mathbf{p}^*\| \to 0$ as $t \to \infty$, thereby establishing that $\mathbf{p}^t \to \mathbf{p}^*$ as $t \to 0$.

■

B. Stochastic and Time-Varying Channel Case

The most remarkable properties of dual averaging are exhibited in the stochastic and time-varying channel case, as given by the following theorem.

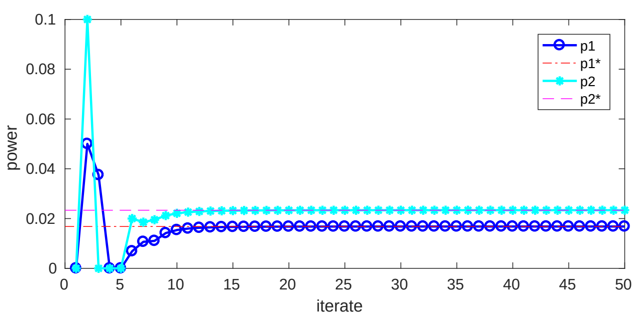
Theorem 4: Let the model of the stochastic and time-varying channel be given as in Section III-B. Let the mean channel $(\bar{\mathbf{G}}, \bar{\eta})$ be feasible with the optimal power vector $\mathbf{p}^* \in \mathcal{P}$. Then $\mathbf{P}^t \to \mathbf{p}^*$ almost surely, as $t \to \infty$, where $\mathbf{P}^t$ is given in Algorithm 3.

Proof Elements: The proof in this case is substantially more involved than the deterministic and time-invariant case, which can only be touched on very briefly here. Consider the cost function: $u_i(\mathbf{p}) = \mathbf{E}[-\frac{1}{2G_{ii}}(G_{ii}p_i - r_i^*(\sum_{j\neq i} G_{ij}p_j + \eta_i))^2]$. One can rewrite Algorithm 3 as a standard mirror descent update on this cost function. After that, one converts this discrete-time mirror descent update to a continuous-time ODE and establish convergence to a fluid limit (almost surely), which can then be converted back (with more analysis) to almost sure convergence in the discrete-time mirror-descent dynamics, thereby establishing the result. ■

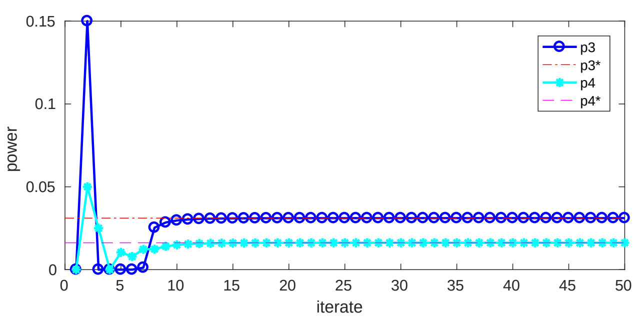
We highlight three important aspects. First, almost sure convergence to a deterministic power vector is ensured, even when the underlying channel environment follows an arbitrary stochastic process. Second, this deterministic power vector limit is not influenced by the specific distribution of the stochastic process; instead, it only depends on the mean (first moments) of the underlying distribution. Finally, this deterministic power vector limit is the optimal power vector of the mean channel. Consequently, law of large numbers ensure the performance guarantees of the dual averaging power control algorithm: using dual averaging over a long horizon is effectively the same as using that optimal power vector.

V. Simulation

To validate the convergence and performance guarantees of dual averaging, we provide here a series of numerical experiments illustrating the algorithm's various properties. In all the simulations, we select the target SINR $R^* = 0.5 \times \mathbf{1}$, where $\mathbf{1}$ is the all-ones vector with dimension equal to the number of links.

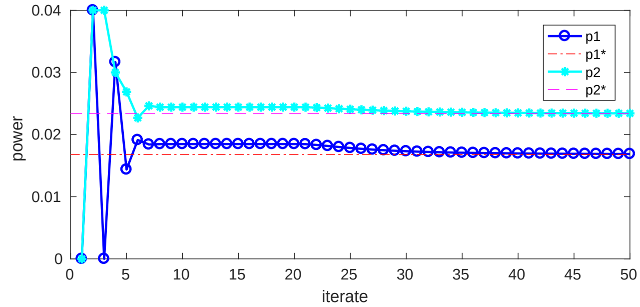


(a) Power marginals for link 1 and 2

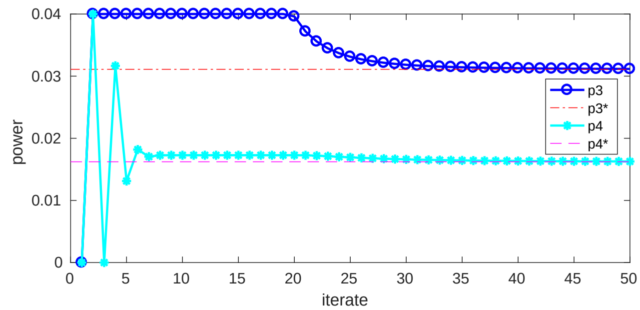


(b) Power marginals for link 3 and 4

Fig. 1. Convergence behavior of Algorithm 2 with $p_i^{max} = 10^4$ for all links i.



(a) Power marginals for link 1 and 2



(b) Power marginals for link 3 and 4

Fig. 2. Convergence behavior of Algorithm 2 with $p_i^{max} = 0.04$ for all links i.

A. Deterministic and Time-Invariant channel

Consider a deterministic and time-invariant channel of four links with the following parameters.

$$G = \begin{bmatrix} 6 & 1 & 2 & 1 \\ 1 & 6 & 1 & 2 \\ 2 & 1 & 6 & 1 \\ 1 & 2 & 1 & 6 \end{bmatrix}, \eta = \begin{bmatrix} 0.1 \\ 0.2 \\ 0.3 \\ 0.1 \end{bmatrix}$$

This channel, with parameters (G, η), is feasible. Further, the optimal power vector is

$$\mathbf{p}^* = \begin{bmatrix} 0.016815 \\ 0.023363 \\ 0.031101 \\ 0.016220 \end{bmatrix}$$

Figure 1 depicts the convergence behavior of the channel (G, η). Specifically, Figure 1(a) shows the individual power iterates for the first and second links. Figure 1(b) shows the individual power iterates for the third and forth links. Here, we set $y_i^0 = 0$ and $p_i^{max} = 10^4$ for all links: the latter means that the effective maximum allowable power is infinity because the individual power iterates never exceed this upper bound and hence there is no need for downward projection (upward projection when the power iterate drops below 0 still occurs). Figure 1 indicates that power iterate in dual averaging (Algorithm 2) converges to the optimal power vector $\mathbf{p}^*$.

Figure 2 depicts the convergence behavior of the same channel (G, η) with $p_i^{max} = 0.04$ for all links i. This

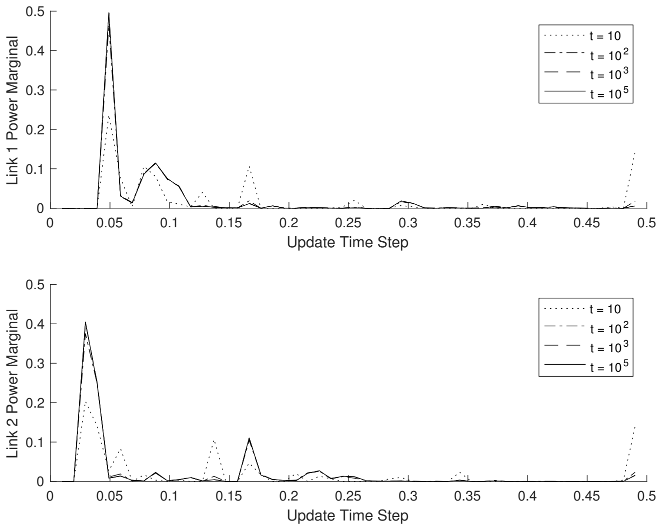


Fig. 3. Convergence to the steady-state distribution for the Foschini-Miljanic power control algorithm. Distributions are shown at several increasing time steps. The power distributions converge to the steady-state distribution (thick black curve) for the Foschini-Miljanic algorithm.

maximum power bound is intentionally set to be low to trigger downward projection when the power iterates exceed this value. Figure 2 demonstrates the convergence of dual averaging to the power vector $\mathbf{p}^*$ in this case as well.

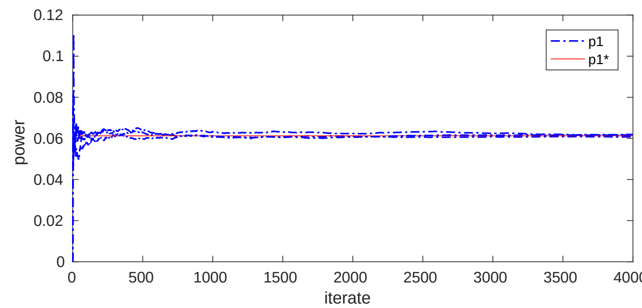
B. Stochastic and Time-Varying Channel: Bernoulli IID Environment

We now proceed to investigate the more interesting case where the channel is stochastic and time-varying. We consider a simple two-link case where the channel environment is **iid** according to the following distribution over (G, η).

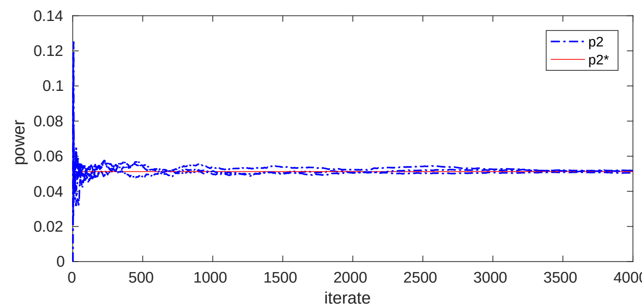
$$(G, \eta) = \begin{cases} (G_1, \eta_1), & w.p. \quad 0.25 \\ (G_2, \eta_2), & w.p. \quad 0.75, \end{cases}$$

where $G_1 = \left[\begin{smallmatrix} 1 & 3 \\ 3 & 1 \end{smallmatrix}\right]$, $G_2 = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right]$, $\eta_1 = \left[\begin{smallmatrix} 0.1 \\ 0.2 \end{smallmatrix}\right]$ and $\eta_2 = \left[\begin{smallmatrix} 0.15 \\ 0.05 \end{smallmatrix}\right]$.

This is a good simple example to demonstrate the superiority of the proposed dual averaging power control algorithm (Algorithm 3), which has two aspects. The first aspect is stability. In this case, as shown in Figure 3, the Foschini-Miljanic (FM) algorithm only converges to a distribution (the thick black curve is the steady-state distribution), while dual averaging converges almost surely to a constant power vector (Figure 4). The second aspect is performance guarantee. Specifically, the constant power vector that the dual averaging power iterates converge to (almost surely) is the optimal power vector of the average channel $(\bar{G}, \bar{\eta})$, where $\bar{G} = 0.25G_1 + 0.75G_2 = \left[\begin{smallmatrix} 1.75 & 1.50 \\ 1.50 & 1.75 \end{smallmatrix}\right]$ and $\bar{\eta} = 0.25\eta_1 + 0.75\eta_2 = \left[\begin{smallmatrix} 0.1375 \\ 0.0875 \end{smallmatrix}\right]$. This average $(\bar{G}, \bar{\eta})$ channel is feasible and its corresponding optimal power vector is $\mathbf{p}^* = \left[\begin{smallmatrix} 0.06125 \\ 0.05125 \end{smallmatrix}\right]$. This indicates that if the dual averaging algorithm were deployed in such channels, then the noisy environment is the equivalent to the corresponding average channel that is deterministic and time-invariant, a property that is absent in FM.

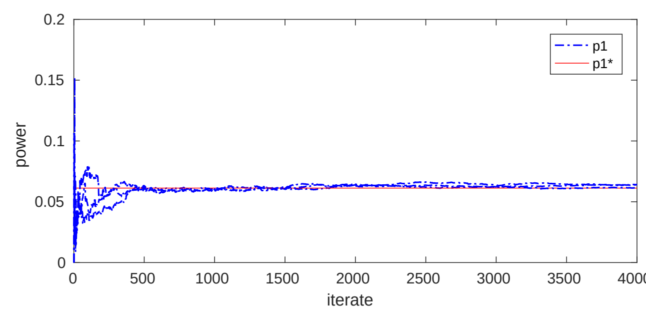


(a) Power marginals for link 1

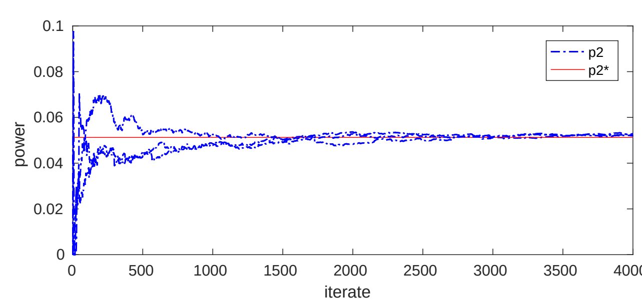


(b) Power marginals for link 2

Fig. 4. Almost Sure Convergence of Dual Averaging for the Bernoulli IID environment. In each subfigure, all the sample traces representing different trajectories of the power iterates converge to the same optimal power vector $\mathbf{p}^*$.

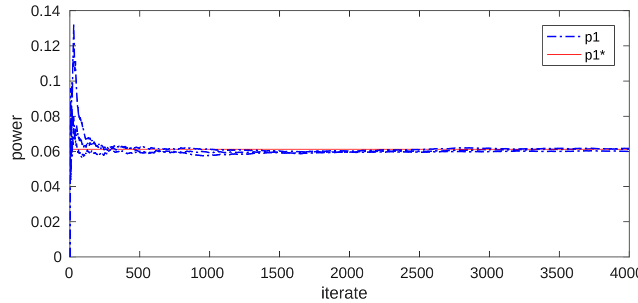


(a) Power marginals for link 1

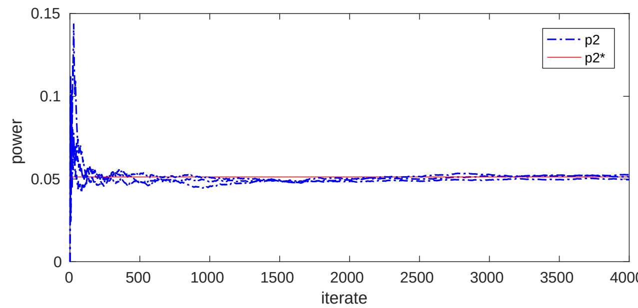


(b) Power marginals for link 2

Fig. 5. Almost Sure Convergence of Dual Averaging for the Gaussian IID environment.



(a) Power marginals for link 1



(b) Power marginals for link 2

Fig. 6. Almost Sure Convergence of Dual Averaging 3 for the stationary Markovian environment.

C. More Stochastic and Time-Varying Channels: Log-normal IID and Markovian Environments

Here we consider two more stochastic and time-varying two-link channels and demonstrate the desired properties of the dual averaging power control algorithm. In each of the two scenarios, FM suffers from the same drawbacks as in the previous discussion in the Bernoulli IID environment. Consequently, we will omit the corresponding plots for FM due to space limitation.

The first scenario is Log-normal IID environment. Specifically, every entry of G and η is sampled **iid** from a Log-normal distribution with $\sigma = 1$. The μ parameter for each entry is selected such that $\mathbb{E}[G] = \bar{G}, \mathbb{E}[\eta] = \bar{\eta}$ where $(\bar{G}, \bar{\eta})$ are the same as in V-B. Figure 5 demonstrates almost sure convergence of dual averaging (Algorithm 3) for this channel environment is to $\mathbf{p}^*$. The important point to note here is that the almost sure convergence of dual averaging is distribution-free: the specific distribution of the underlying channel environment does not influence the final power vector that dual averaging converges to, so long as the mean is the same.

The second scenario is a Markovian environment, where the channel environment, $s_t = (G_t, \eta_t)$, abides by a Markov chain from the state space $\mathcal{S} = \{s_1, s_2\}$ where $s_i = (G_i, \eta_i)$. The pairs (G_i, η_i), for $i \in \{1, 2\}$, are the same as in Subsection V-B. The transition probabilities of the Markov chains are as follows:

$$\mathbb{P}(s_{t+1}|s_t = s_1) = \begin{cases} \frac{2}{5}, & s_{t+1} = s_1 \\ \frac{3}{5}, & s_{t+1} = s_2 \end{cases}$$

$$\mathbb{P}(s_{t+1}|s_t = s_2) = \begin{cases} \frac{1}{5}, & s_{t+1} = s_1 \\ \frac{4}{5}, & s_{t+1} = s_2 \end{cases}$$

The system is initialized with the invariant distribution, $(\pi(s_1), \pi(s_2)) = (\frac{1}{4}, \frac{3}{4})$. As a result the chain will be stationary and ergodic.

Figure 6 describes the convergence behavior of dual averaging (Algorithm 3) in this channel, where the sample paths are plotted for each link. Here again, the power iterates converge almost surely to $\mathbf{p}^*$. This example indicates that not only is convergence distribution-free, but also the iid assumption, an important condition in [15], is not needed either.

References

[1] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

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

[3] G. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *Vehicular Technology, IEEE Transactions on*, vol. 42, no. 4, pp. 641–646, Nov 1993.

[4] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1341–1347, 1996.

[5] D. Mitra, "An asynchronous distributed algorithm for power control in cellular radio systems," in *Wireless and Mobile Communications*. Springer, 1994, pp. 177–186.

[6] S. Ulukus and R. D. Yates, "Stochastic power control for cellular radio systems," *Communications, IEEE Transactions on*, vol. 46, no. 6, pp. 784–798, 1998.

[7] D. Famolari, N. Mandayam, D. Goodman, and V. Shah, "A new framework for power control in wireless data networks: Games, utility, and pricing," 1999.

[8] Z. Zhou and N. Bambos, "Wireless communications games in fixed and random environments," in *2015 54th IEEE Conference on Decision and Control (CDC)*, Dec 2015, pp. 1637–1642.

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

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

[11] P. C. Weeraddana, M. Codreanu, and M. Latva-aho, *Weighted sum-rate maximization in wireless networks: A review*. Now Publishers Inc., 2012.

[12] C. W. Tan, "Wireless network optimization by perron-frobenius theory," in *Information Sciences and Systems (CISS), 2014 48th Annual Conference on*. IEEE, 2014, pp. 1–6.

[13] T. Holliday, N. Bambos, P. Glynn, and A. Goldsmith, "Distributed power control for time varying wireless networks: Optimality and convergence," in *in Proceedings: Allerton Conference on Communications, Control, and Computing*, 2003.

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

[15] Z. Zhou, D. Miller, P. Glynn, and N. Bambos, "A stochastic stability characterization of the foschini-miljanic algorithm in random wireless networks," in *2016 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2016.

[16] A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro, "Robust stochastic approximation approach to stochastic programming," *SIAM Journal on optimization*, vol. 19, no. 4, pp. 1574–1609, 2009.