Stochastic Optimization and Its Applications in Multihop Wireless MIMO Networks

Yih-Hao Lin, Tara Javidi, Rene L. Cruz and Laurence B. Milstein

Department of Electrical and Computer Engineering

University of California San Diego

La Jolla, CA 92093-0407

Email: yhlin@qualcomm.com, tara@ece.ucsd.edu, rcruz@ucsd.edu, milstein@ece.ucsd.edu

Abstract

In this paper, an algorithm is devised to solve a general mathematical problem of optimizing an expected average cost criterion subject to constraints on both 1) instantaneous resource allocation decisions and 2) the long-term average resource allocation requirements. The construction of the algorithm does not require knowledge about the statistics of the underlying process, yet the performance of this algorithm, under mild mixing conditions on the underlying stochastic process, is shown to converge asymptotically to the optimal under causal observations. This general result is applied to the case of joint power-rate allocation in MIMO wireless multi-hop networks, in which nodes can form broadcast clusters in an attempt to take advantage of dirty paper coding. A distributed implementation of the algorithm is presented for the applied case. The numerical results shows that using dirty paper coding within each broadcast cluster provides a significant improvement over traditional link-based techniques.

I. Introduction

In many wireless networking applications, the design of a network necessitates optimizing an expected network performance criterion, such as the sum power of the network, while satisfying a set of long-term average QoS measures, such as the average end-end rate, in the face of instantaneous and stochastic variation in availability and quality of the wireless channel. In all such settings, the design problem can be formulated as a constrained optimization problem with an expected average cost criterion, e.g., total power, with two types of constraints: a set of constraints on the expected average of certain quantities, e.g., flow rates, and a set of instantaneous constraints which are functions of the underlying stochastic process, e.g., the channel conditions. Motivated by such applications, this paper studies the general mathematical problem of optimizing an expected average cost criterion subject to constraints on both 1) instantaneous resource allocation decisions (a function of the instantaneous realization of the underlying stochastic process) and 2) the long-term average resource allocation requirements. We are ultimately interested in solutions to problems which do not rely on the availability of a precise statistical model (joint distribution) of the underlying stochastic process.

The contributions of the paper are two fold. First, to solve the problem, we use stochastic approximation to devise an algorithm, whose construction does not require knowledge about the statistics of the underlying process. The performance of the algorithm, under mild mixing conditions on the underlying stochastic process, is shown to converge asymptotically to the optimal under the instantaneous and long-term average resource allocation constraints. The second contribution of the paper is applying this general result to the case of joint power-rate allocation in MIMO wireless multi-hop networks. Here we consider a MIMO network, where nodes can form broadcast clusters in an attempt to take advantage of dirty paper coding (DPC). We use the proposed algorithm to devise a cross-layer optimized resource allocation scheme. In this setting, we identify network scenarios under which the use of DPC provides a significant improvement over more traditional and link-based techniques.

Our work in the paper is strongly motivated by [1], [2], [10], [16] and [17], all of which address various aspects of maximizing the expected network utility, which is a strictly concave function of allocated rate vectors, subject to stochastically varying constrained sets. The problem we study in this paper has a close conceptual tie to these papers, but provides a distinct problem formulation: in our setting, the objective function (power) may or may not be explicitly written as a strictly concave function of control variables of interest (power and rate). We would like to point out, though, that similar to findings in [17], our proposed algorithm — based on a stochastic approximation/subgradient approach — is ”very parsimonious, and naturally decomposes to become a decentralized algorithm in special cases when different network elements can make their own control decisions independently.” In the interest of brevity, we do not include these naturally decentralized solutions in this paper; instead, we refer the interested reader to [13] [14].

Our work can also be viewed as a generalization of the problem studied in [11], in which the authors studied the opportunistic downlink scheduling problem under time-varying channels. They proposed a two stage off-line algorithm, applying the duality technique and the stochastic subgradient method. The first stage (warm-up) generates a sequence of dual variables which converges to a dual problem optimal point. In each time slot, the scheduler observes the channel state and updates the dual variable accordingly, by adding a scaled “subgradient-like” vector to it. In the second stage, the scheduler applies the optimal dual variables, obtained from the first stage, to the Lagrangian relaxation of the primal problem, and decides the transmit power and rates with regard to the instantaneous channel quality.

Notice that the algorithm presented in [11] is an off-line algorithm, because the sequence generated in the first stage of the algorithm can, in general, take a long time to converge. Instead, we propose an online algorithm which generates sequences of dual and primal variables concurrently, while still maintaining the optimality of the algorithm. Another important point we would like to make is with regard to the proof of optimality for the proposed algorithm. The optimality proof given in [11] depends on an important assumption: the conditional expectation of the proposed “subgradient-like” vector, given the past history, forms a subgradient of the dual function. Strictly speaking, this may or may not be true, because the conditional expectation of the “subgradient-like” vector is, in general, not a subgradient of the dual function. Some degree of dependency between the instantaneous channel condition and the scheduler’s previous decisions is expected to remain. Such a subtle effect was not captured in [11]. In contrast to [11], we apply an alternative approach, not using the subgradient assumptions in [11] directly, to show that the sequence of dual variables generated in each recursive step asymptotically converges in the expectation of the norm to the set of maximizer(s) of the dual problem (Theorem 3).

We would also like to point out that our application of interest coincides with the problem studied by [15], with the main difference being our approaches and methodologies: Our work relies on constrained convex optimization and stochastic approximation frameworks to minimize the expected sum power, subject to a constraints on the minimum long-term average rate at the flow level, while [15] uses a Lyapunov-type penalty function as the cost of violating the QoS requirements. The Lyapunov drift analysis is, then, applied to prove stability, and hence admissibility and optimality. Working within a stochastic approximation framework, our formulation allows for more general arrival and channel processes. More precisely, while Lyapunov techniques usually rely on independent renewals over time, the convergence of the stochastic approximation algorithm only requires mild mixing conditions. In addition, channel and arrival processes in our work do not require finite support as required in [15].

This paper is organized as follows. We begin with the notation and some mathematical preliminary results which will be used later in the paper, and present the problem formulation of interest and our main results in Section III. In the first part of Section III, we present the problem formulation as a convex optimization and also present the appropriate dual formulation to achieve a lower bound on the optimal performance. In Subsection III-B, motivated by the sub-gradient projection approach, we introduce a class of control policies, which we call dual controllers. The main results of the paper, i.e. the asymptotic admissibility and optimality of the proposed controller, are given in Subsection III-C. Section IV provides an analysis of the problem by outlining the proof of admissibility and optimality theorems stated in Subsection III-C. The detailed proof of these theorems are completed in the appendices. In Section V, we specialize the proposed solution to a MIMO wireless adhoc network context. In Section VI, we present numerical examples demonstrating the performance of the proposed algorithm using a small network of 15 nodes. Finally, we conclude this work with a discussion of promising future research topics in Section VII.

II. Preliminaries and Notation

For completeness, in this section, we provide definitions and facts frequently used in convex optimization and stochastic approximation. These are used for establishing our result later in the body of the paper. Those readers familiar with the terms and facts below can skip this section.

***Definition** 1:* A stochastic process $\{\xi^n\}$ is called *exogenous* to process Z^n if its future evolution, given the past history of Z^n, depends only on the past history of ξ^n. More precisely, let $\{\mathcal{F}_n\}$ be a sequence of nondecreasing σ-algebras, where $\mathcal{F}_n$ measures $\{Z^{j-1}, \xi^j; j \leq n\}$, then an exogenous process satisfies

$$P(\xi^k \mid \mathcal{F}_n) = P(\xi^k \mid \xi^j, j = 0, 1, \ldots, n) \text{ for } k \geq n. \tag{1}$$

***Definition** 2:* Let $\mathcal{B}_m^n$ be the σ-algebra generated by the random variables $\{\xi^m, \xi^{m+1}, \ldots, \xi^n\}$. $\{\xi^k\}$ is called a ϕ-mixing process [9, p.p. 356] (uniform mixing process [6, p.p. 345]) if for

$$\phi_k \triangleq \sup_i \sup_{A \in \mathcal{B}_0^{i+k}, B \in \mathcal{B}_0^i} |\, P\{A|B\} - P\{A\} \,|\,, \tag{2}$$

we have $\lim_k \phi_k = 0$.

***Fact** 1:* [7, Chapter 6, Lemma 4]

Let ξ^n be ϕ-mixing, and let g^n be measurable on $\mathcal{F}_n^\infty$ with $|g^n| \leq K$, then

$$\Big| \mathbb{E}[\, g^{n+k} \,|\mathcal{B}_0^n \,] - \mathbb{E}[\, g^{n+k} \,] \Big| \leq 2K\phi_k. \tag{3}$$

***Definition** 3:* Let $N_\delta(x)$ denote a δ-neighborhood of x. A set-valued function $\mathcal{B}(\cdot)$ is said to be *upper-semi-continuous* if it satisfies

$$\bigcap_{\delta > 0} \text{co} \left[\bigcup_{y \in N_\delta(x)} \mathcal{B}(y) \right] = \mathcal{B}(x), \tag{4}$$

where $\text{co}(A)$ denotes the convex hull of the set A.

***Definition** 4:* Let P_A denote the measure on the Borel sets of $\mathbb{R}^k$ determined by a random variable A. A sequence of random variables A_n is said to converge weakly to a random variable A if

$$EF(A_n) = \int F(x) dP_n(x) \rightarrow EF(A) = \int F(x) dP_A(x) \tag{5}$$

for each bounded and continuous real-valued function $F(\cdot)$ on $\mathbb{R}^k$. We use the following notation to denote the weak convergence:

$$A_n \Rightarrow A. \tag{6}$$

***Definition** 5:* Given a concave function $f : \mathbb{R}^n \mapsto \mathbb{R}$, we say that a vector $d \in \mathbb{R}^n$ is a *subgradient* of f at a point $x \in \mathbb{R}^n$ if

$$f(z) \leq f(x) + (z-x)'d, \quad \forall z \in \mathbb{R}^n. \tag{7}$$

We close the section with a word on our notations. Given a sequence of vector $\{\vec{v}^n, n = 1,2,3,\ldots\}$, where $\vec{v}^n = (v_1^n, v_2^n, \ldots, v_r^n)$, then $\limsup_n \vec{v}^n$ is defined element by element, i.e.,

$$\limsup_n \vec{v}^n \triangleq (\limsup_n v_1^n, \limsup_n v_2^n, \ldots, \limsup_n v_r^n).$$

$\mathbb{R}^{r,+}$ denotes the nonnegative r-dimensional real valued Euclidean space.

III. Stochastic Optimization

A. Problem Formulation and Model

Consider a system with a countable-state underlying stochastic process denoted by $\{\xi^n;\ n = 0,1,2,\ldots\}$, where the probability distribution of $\{\xi^n\}$ is unknown, but the realization $\{\xi^n(\omega)\}$ can be perfectly observed by the controller/scheduler at time n. At any time n, the controller /scheduler can select a set of control actions called control variables $\{(X^n, \vec{Y}^n) |\ X^n \in \mathbb{R}, \vec{Y}^n \in \mathbb{R}^r, n = 1,2,\ldots\}$, utilizing the observation $\{\xi^k(\omega);\ k = 0,1,2,\ldots,n\}$. Given a realization $\xi^n(\omega)$, the choice of control variables $(X^n, \vec{Y}^n)$ is restricted to the set $\mathcal{D}(\xi^n(\omega)) \subset \mathbb{R} \times \mathbb{R}^r$ determined by $\xi^n(\omega)$. The scalar value X^n is the performance measure of interest, such as network power consumption, to be optimized, while the appropriate choice of $\vec{Y}^n$ guarantees an acceptable level of performance in terms of other measures of interest, such as minimum end-end vector of flow rates. In other words, we are interested in selecting admissible control variables $(X^n, \vec{Y}^n) \in \mathcal{D}(\xi^n(\omega))$ in order to minimize the long-term average of the expected value of X^n, while guaranteeing that the long-term average of the expected value of $\vec{Y}^n$ satisfies the constraint (9). Before we proceed with the precise formulation, we need to establish the following definition of a causal policy as a sequence of mappings which is used by the controller to determine the control actions.

***Definition** 6:* A causal (possibly randomized) policy $\pi = \{\pi_0, \pi_1, \ldots\}$ is defined to be a sequence of mappings

$$\pi_n : \{X^{j-1}, \vec{Y}^{j-1}, \xi^j; j \leq n\} \rightarrow \mathcal{P}(\mathcal{D}(\xi^n(\omega)),$$

where $\mathcal{P}(\mathcal{D}(\xi^n(\omega))$ is a probability measure on the set $\mathcal{D}(\xi^n(\omega))$. The set of all causal (possibly randomized) policies is denoted by Γ.

In other words, the control policy considered in this paper could be a randomized policy, meaning that, given realization $\{\xi^n(\omega)\}$, the vector of controlled variables $(X^n_\pi, \vec{Y}^n_\pi)$ forms a sequence of random vectors. All expectations, $\mathbb{E}\{\cdot\}$, are taken over the probability space constructed by $\{\xi^n\}$, as well as the selected policy π (possibly randomized).

With the notion of policy defined above, we are ready to introduce the stochastic optimization in the following general form:

***Problem** 1 (Stochastic Optimization):*

$$X^* = \inf_{\pi\in\Gamma} \limsup_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{X_\pi^n\} \tag{8}$$

subject to

$$\limsup_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{\vec{Y_\pi^n}\} \le 0. \tag{9}$$

and

$$\text{for all } n: \quad (X_\pi^n, \vec{Y_\pi^n}) \in \mathcal{D}(\xi^n), \tag{10}$$

where X_π^n is a scalar and $\vec{Y_\pi^n} \in \mathbb{R}^r$. Here, we added script π to $\{X_\pi^n\}$ and $\{\vec{Y_\pi^n}\}$ to distinguish the control actions taken by different policies, i.e., $\{X_\pi^n\}$ and $\{\vec{Y_\pi^n}\}$ denote the sequence of controlled variables chosen by the causal[1] policy $\pi \in \Gamma$. Note that (10) is usually referred to as the feasibility constraint, while (9), referred to as admissibility constraint[2], guarantees a long-term average performance. More precisely,

***Definition** 7:* A policy $\pi \in \Gamma$ is called an *admissible* control policy if it chooses control variables $(X_\pi^n, \vec{Y_\pi^n})$ from the set $\mathcal{D}(\xi^n(\omega))$ such that $\limsup_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{\vec{Y_\pi^n}\} \le 0$.

We aim to find admissible randomized control policies achieving the optimal value of Problem 1 under the following technical assumptions:

(A.1) $\mathcal{D}(\xi^n(\omega))$ is compact for each n and ω. In addition, the elements in $\mathcal{D}(\xi^n(\omega))$ are bounded

[1]Although we assume the causal policies in our model, the results and the associated proofs in this paper are independent of the causality. In other words, if we consider a broader class of policies which allows controlling the system based on both the future and the past events of $\{\xi^n\}$, the optimal performance will not change.

[2]The presentation of Problem 1 has been chosen to simplify the analysis. Although constraint (9) appears to be decoupled in each entry of $(\mathbb{E}\{Y_{\pi,1}^n\}, \mathbb{E}\{Y_{\pi,2}^n\}, \ldots, \mathbb{E}\{Y_{\pi,r}^n\})$, the formulation actually covers more general linear inequality constraint. This is done by introducing auxiliary variables with additional constraints to redefine $\mathcal{D}(\xi^n)$. An example containing weighted sum linear inequality constraint can be found in Section V.

in norm uniformly in n and ω, i.e., we assume there exists a constant B_1 such that $\forall\ (x, \vec{y}) \in \mathcal{D}(\xi^n(\omega))$, we have $\|x\| \leq B_1$ and $\|\vec{y}\| \leq B_1$, irrespective of n and ω.

(A.2) $\{\xi^n\}$ is stationary and ϕ-mixing (uniform mixing), has a finite support, and is an exogenous process to causal control actions (control actions which utilize only the causal observations).

The above assumptions enable a dual approach to the above constraint optimization problem in which, for a given $\vec{\theta} \in \mathbb{R}^{r,+}$, we define:

$$V(\vec{\theta}) \triangleq \inf_{\pi\in\Gamma} \limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{X_\pi^n + \vec{\theta} \cdot \vec{Y_\pi^n}\}. \tag{11}$$

On the other hand,

$$\begin{aligned} V(\vec{\theta}) &= \inf_{\pi\in\Gamma} \limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{X_\pi^n + \vec{\theta} \cdot \vec{Y_\pi^n}\} \\ &= \limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{\min_{(X^n,\vec{Y}^n)\in\mathcal{D}(\xi^n(\omega))} X^n + \vec{\theta} \cdot \vec{Y}^n\} \\ &= \limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{V(\vec{\theta}, \xi^n)\} \end{aligned} \tag{12}$$

whereby denoting

$$V(\vec{\theta}, \xi^n) \triangleq \min_{(X^n,\vec{Y}^n)\in\mathcal{D}(\xi^n)} X^n + \vec{\theta} \cdot \vec{Y}^n, \tag{13}$$

we have slightly abused the notation $V(\cdot)$ to emphasize that $V(\theta, \xi^n)$ is an analog of $V(\theta)$ in a per-time-instant sense. Note that the second equality in (12) holds because of the existence of a causal policy whose control decision at time n follows the solution to

$$\min_{(X^n,\vec{Y}^n)\in\mathcal{D}(\xi^n(\omega))} X^n + \vec{\theta} \cdot \vec{Y}^n.$$

Applying the duality concepts in optimization theory, we obtain the following bound:

***Proposition** 1:* For all $\vec{\theta} \in \mathbb{R}^{r,+}$, $V(\vec{\theta})$ is a lower bound for the solution to the stochastic optimization problem, i.e., $V(\vec{\theta}) \leq X^*$.

The proof of this proposition is given in Appendix A.

B. The Dual Controller and Subgradient Projection

Next, in light of Proposition 1, and motivated by the subgradient projection methods used in solving convex optimization, we propose a recursive algorithm which asymptotically solves Problem 1. Let us restrict our attention to a class of stationary control policies which at time n make decisions, based on the instantaneous observation of $\xi^n(\omega)$ and a (possibly time varying) referenced parameter θ^n:

Definition *8:* Given a sequence of reference variables θ^n, the *dual control policy* π^* defines the control actions at time n $(= 0, 1, 2, \ldots)$ in the following manner. For a sequence of time dependent referenced parameters $\{\theta^n,\ n = 0, 1, 2, \ldots\}$, the policy π^* selects the random vector $(X^n_{\pi^*,\theta^n}, \vec{Y}^n_{\pi^*,\theta^n})$ uniformly over the compact set of minimizers of $V(\theta_n, \xi^n)$, i.e.,

$$(X^n_{\pi^*,\theta^n}, \vec{Y}^n_{\pi^*,\theta^n}) \in \arg \min_{(x,\vec{y}) \in \mathcal{D}(\xi^n)} x + \vec{\theta}^n \cdot \vec{y}. \tag{14}$$

If the referenced parameter θ is time independent , π^* is a stationary policy:

$$(X^n_{\pi^*,\theta}, \vec{Y}^n_{\pi^*,\theta}) \in \arg \min_{(x,\vec{y}) \in \mathcal{D}(\xi^n)} x + \theta \cdot \vec{y}. \tag{15}$$

When there is a tie in (14) or (15), it is broken arbitrarily.

We will show in subsequent sections that this restriction, which is motivated by the dual formulation above[3], does not cause a loss of optimality when reference values are generated carefully, e.g., according to a subgradient projection scheme, and under the following technical assumption:

(A.3) The maximizer set of $V(\cdot)$ over $\mathbb{R}^{r,+}$, denoted by Θ, is compact. In other words, there exists $0 < K_u < \infty$, such that $\forall \vec{\theta}^* = (\theta^*_1, \ldots, \theta^*_r)$ for which $V(\vec{\theta}^*) = \max_{\vec{\theta} \geq 0} V(\vec{\theta})$, then

[3] θ plays the role of a noisy estimate of the dual variables in convex optimization Problem 1 [5]. The details can be found in [12].

$\theta_i^* \in [0, K_u)$ for $i = 1, \ldots, r$. We denote the compact set by

$$H \triangleq \{\theta \in \mathbb{R}^{r,+} |\ 0 \leq \theta_i \leq K_u\}. \tag{16}$$

Note that assumption (A.3) holds when constraint (9) contains an interior admissible point. Moreover, if such a point is known in advance, a bound, as denoted by K_u, can be explicitly derived. This result is summarized in the following proposition, whose proof is given in Appendix B.

Proposition *2:* Θ is compact if and only if there exists an admissible policy π such that $\limsup_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \mathbb{E}\{\vec{Y}_\pi^n\} < 0$. Moreover, if $\limsup_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \mathbb{E}\{\vec{Y}_\pi^n\} < -\vec{\delta}$ for some $\vec{\delta} = (\delta, \delta, \ldots, \delta)$ with $\delta > 0$, the i-th components of $\vec{\theta}^*$ is bounded above by $(B_1 - X^*)/\delta$ for each $\vec{\theta}^* \in \Theta$.

To complete the dual control policy introduced in *Definition* 8, the reference variable $\vec{\theta}^n$ is updated via the recursive algorithm below:

$$\vec{\theta}^{(n+1)} = \Pi_H[\vec{\theta}^n + \epsilon \vec{Y}^n_{\pi^*,\theta^n}], \tag{17}$$

where $\epsilon > 0$ is the step size and Π_H is the projection onto the set H.

The asymptotic behavior of recursive algorithms given by (14) and (17) determines the performance of dual control policy π^*, and it can be analyzed using stochastic approximation [8]. Note that the evolution of $\{\theta^n\}$ in general depends on the step size ϵ. To emphasize this point, in the rest of this paper, we add subscript ϵ to θ^n as needed.

C. Overview of the Main Results

The main contribution of our work is summarized by the following theorems establishing the asymptotic admissibility and the optimality of dual control policy, given the recursive algorithm (17).

Theorem *1 (Asymptotic Admissibility):* Under Assumptions (A.1)-(A.3), and given the sequence of reference parameters obtained via recursive algorithm (17), the dual control policy π^* is asymptotically admissible, i.e.,

$$\lim_{\epsilon\to 0}\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{\vec{Y}^n_{\pi^*,\vec{\theta}^n}\}\le 0.$$

Theorem *2 (Asymptotic Optimality):* Under Assumptions (A1)-(A3), and given the sequence of reference parameters obtained via recursive algorithm (17), the dual control policy π^* is asymptotically optimal, i.e., for $\forall\gamma>0$

$$\limsup_{\epsilon\to 0}\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{X^n_{\pi^*,\vec{\theta}^n}\}\le X^*+\gamma. \tag{18}$$

IV. Analysis of Problem 1: Proof Outline for Theorems 1 and 2

In this section, we provide an outline of the proof for admissibility and optimality of the subgradient dual controller. The detailed proof are given in Appendix C.

The first step in proving the result is to establish the asymptotic convergence of the sequence of reference parameters, $\{\theta^n_\epsilon\}_n$.

Step1: The Asymptotic Behavior of θ^n

In Appendix C, we establish the following theorem:

Theorem *3:* Let Θ be the set of maximizers of $V(\cdot)$ over H, and assume Assumptions (A.1)-(A.3) hold. Given $\delta>0$, there exists an $\hat{\epsilon}>0$ such that, for any $\epsilon<\hat{\epsilon}$,

$$\mathbb{E}\{\text{ distance}(\,\theta^n_\epsilon,\Theta\,)\,\}<\delta \text{ for all but a finite number of times in } n, \tag{19}$$

where distance(θ,A) denotes the distance from point θ to set A.

Step 2: Asymptotic Admissibility of the Dual Control Policy

We define the compensation terms from below $\hat{Z}^n \in \mathbb{R}^{r,+}$ and above $\check{Z}^n \in \mathbb{R}^{r,+}$ by rewriting the recursion (17) as

$$\vec{\theta}^{(n+1)} = \vec{\theta}^n + \epsilon \; (\vec{Y}^n_{\pi^*,\vec{\theta}^n} + \hat{Z}^n - \check{Z}^n). \tag{20}$$

Equation (20) can be rewritten as

$$\vec{\theta}^{(n+1)} - \vec{\theta}^n = \epsilon \; (\vec{Y}^n_{\pi^*,\vec{\theta}^n} + \hat{Z}^n - \check{Z}^n). \tag{21}$$

We sum both sides of (21) from $n = 0$ to $n = N - 1$. The total of the left hand side of (20) is a telescoping sum, therefore, we have

$$\vec{\theta}^{(N)} - \vec{\theta}^0 = \epsilon \sum_{n=0}^{N-1} (\vec{Y}^n_{\pi^*,\vec{\theta}^n} + \hat{Z}^n - \check{Z}^n) \tag{22}$$

Take the expectation on both sides of (22) and dividing them by ϵN, we have:

$$\frac{\mathbb{E}\{\vec{\theta}^{(N)} - \vec{\theta}^0\}}{\epsilon N} = \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{(\vec{Y}^n_{\pi^*,\vec{\theta}^n} + \hat{Z}^n - \check{Z}^n)\} \overset{(a)}{\geq} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{(\vec{Y}^n_{\pi^*,\vec{\theta}^n} - \check{Z}^n)\}. \tag{23}$$

The inequality (a) in (23) results from the fact $\hat{Z}^n \geq 0$. Upon taking $\limsup_\epsilon \limsup_N$ on both sides of (23), we get

$$\limsup_{\epsilon} \limsup_{N} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{(\vec{Y}^n_{\pi^*,\vec{\theta}^n} - \check{Z}^n)\} \leq \limsup_{\epsilon} \limsup_{N} \frac{\mathbb{E}\{\vec{\theta}^{(N)} - \vec{\theta}^0\}}{\epsilon N} \overset{(b)}{=} 0, \tag{24}$$

in which the equality (b) holds because $\|\vec{\theta}^n\|$ is bounded by rK_u, where r is the dimension of $\vec{\theta}^n$.

The following lemma complete the last piece of our proof for the admissibility of dual control policy.

***Lemma** 1:*

$$\limsup_{\epsilon} \limsup_{N} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{\check{Z}^n\} = 0. \tag{25}$$

Proof: See Lemmas 13 and 14 in Appendix C. ■

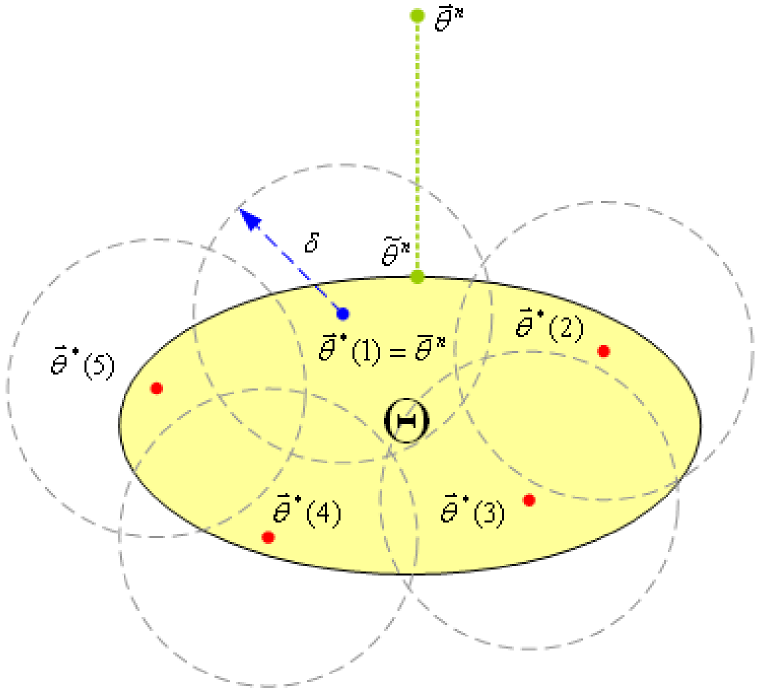

Fig. 1. Projected Point and Partition Points.

Step 3: Asymptotic Optimality

In order to establish asymptotic optimality of our proposed scheme, one needs to study the limiting behavior of $\limsup_N \frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{V(\vec{\theta}^n,\xi^n)\}$. In particular, we show that it is possible to bound the performance of the algorithm close to X^*. To do so, we use Theorem 3 characterizing the limit point for process $\{\theta_\epsilon^n\}$.

Assume (A.3) holds. Since Θ is bounded by a compact set A, given $\delta > 0$, one can find a finite set of points $\{\vec{\theta}^*(1), \vec{\theta}^*(2), \ldots, \vec{\theta}^*(r_\delta)\}$ in Θ such that $\Theta \subset \bigcup_{i=1}^{r_\delta} N_\delta(\vec{\theta}^*(i))$, where $N_\delta(\vec{\theta}) = \{x \in \mathbb{R}^r \mid \|x - \vec{\theta}\| \le \delta\}$.

For notational simplicity, let $\vec{\alpha}^n \triangleq \Pi_\Theta(\vec{\theta}^n)$ denote the projection of $\vec{\theta}^n$ onto the solution set Θ of Problem 2.5. In the proof, we need to identify in which neighborhood, $N_\delta(\vec{\theta}^*(i))$, $\vec{\theta}^n$ lies. To this end, we denote $\vec{\beta}^n$ as the point in $\{\vec{\theta}^*(1), \ldots, \vec{\theta}^*(r_\delta)\}$ that is closest to $\vec{\alpha}^n$. Figure 1 illustrates an example in which the point $\vec{\theta}^n$ is projected onto the set Θ at $\vec{\alpha}^n$. The set Θ is covered by $\bigcup_{i=1}^5 N_\delta(\vec{\theta}^*(i))$, and $\vec{\beta}^n = \vec{\theta}(1)$ is the point in $\{\vec{\theta}^*(1), \ldots, \vec{\theta}^*(5)\}$ which is closest to $\vec{\alpha}^n$.

To prove the asymptotic optimality, we use the following three auxiliary lemmas whose proofs are given in Appendix D.

***Lemma** 2:*

$$\frac{1}{L}\sum_{n=mL}^{mL+L-1} \mathbb{E}\{X^n_{\pi^*,\vec{\beta}^{mL}}\} + \frac{1}{L}\sum_{n=mL}^{mL+L-1} \mathbb{E}\{\vec{\beta}^{mL} \cdot \vec{Y}^n_{\pi^*,\vec{\beta}^{mL}}\} \le X^* + \frac{1}{L}\sum_{n=0}^{L-1} 2B_1(B_2+1)\phi_n,$$

where $B_2 = \max\{\ \|\vec{\theta}^*(1)\|, \ldots, \|\vec{\theta}^*(r_\delta)\|\}$.

***Lemma** 3:* There exists $L' > 0$ such that, for all $L > L'$,

$$\frac{1}{L}\sum_{n=mL}^{mL+L-1} \mathbb{E}\{V(\vec{\theta}^n, \xi^n)\} \le \quad X^* + 2B_1(B_2+1)\delta + \epsilon\frac{L+1}{2}B_1^2 + \mathbb{E}\{\text{distance}(\vec{\theta}^{mL}, \Theta)\}B_1 + \delta B_1.$$

***Lemma** 4:* Let X^* be the optimal solution of Problem 1. Then

$$\limsup_N \frac{1}{N}\sum_{n=0}^{N-1} \mathbb{E}\{V(\vec{\theta}^n, \xi^n)\} = \limsup_M \frac{1}{M}\sum_{m=0}^{M} \frac{1}{L}\sum_{n=mL}^{mL+L-1} \mathbb{E}\{V(\vec{\theta}^n, \xi^n)\}$$
$$\le \quad X^* + 2B_1(B_2+1)\delta + \epsilon\frac{L+1}{2}B_1^2 + \limsup_M \frac{1}{M}\sum_{m=0}^{M} \mathbb{E}\{\text{distance}(\vec{\theta}^{mL}, \Theta)\}B_1 + \delta B_1.$$

Taking the square of both sides of (17), we have

$$\|\vec{\theta}^{n+1}\|^2 - \|\vec{\theta}^n\|^2 \le 2\epsilon\vec{\theta}^n \cdot \vec{Y}^n_{\pi^*,\vec{\theta}^n} + \epsilon^2\|\vec{Y}^n_{\pi^*,\vec{\theta}^n}\|^2. \tag{26}$$

Summing both sides of (26) from $n = 0$ to $N-1$ and dividing them by $\epsilon N/2$, we have

$$\frac{1}{2\epsilon N}\mathbb{E}\{\|\ \vec{\theta}^{N+1}\ \|^2\} - \frac{1}{\epsilon N}\mathbb{E}\{\|\ \vec{\theta}^0\ \|^2\} \le \frac{1}{2N}\sum_{n=0}^{N-1} \mathbb{E}\{\vec{\theta}^n \cdot \vec{Y}^n_{\pi^*,\vec{\theta}^n}\} + \epsilon\frac{1}{2N}\sum_{n=0}^{N-1} \mathbb{E}\{\|\ \vec{Y}^n_{\pi^*,\vec{\theta}^n}\ \|^2\}. \tag{27}$$

Note that the second term on the right-hand side of (27) is bounded by ϵB_1. As a result, we have the following lower bound:

$$\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{\vec{\theta}^n\cdot\vec{Y}^n_{\pi^*,\vec{\theta}^n}\}\geq -\epsilon B_1^2/2-\frac{1}{\epsilon N}\mathbb{E}\{\|\ \vec{\theta}^0\ \|^2\}. \tag{28}$$

Now, combining (28) and the definition of $V(\vec{\theta}^n,\xi^n)$ in (13), we have

$$\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{X^n_{\pi^*,\vec{\theta}^n}\}\leq \epsilon B_1^2/2+\frac{1}{\epsilon N}\mathbb{E}\{\|\ \vec{\theta}^0\ \|^2\}+\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{V(\vec{\theta}^n,\xi^n)\}. \tag{29}$$

Combining (29) with Lemma 3 and 4 results in

$$\begin{aligned}\limsup_N\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{X^n_{\pi^*,\vec{\theta}^n}\} \quad\leq\quad & \limsup_N\frac{\mathbb{E}\{\|\ \vec{\theta}^0\ \|^2\}}{\epsilon N}+X^*+\epsilon\frac{L+2}{2}B_1^2+B_1(2B_2+3)\delta\\ +\quad & \limsup_N\frac{1}{N}\sum_{m=0}^{N}\mathbb{E}\{\text{distance}(\vec{\theta}^0,\Theta)\}B_1.\end{aligned}$$

But, from Theorem 3, we have that, $\forall\delta$, there exists an $\hat{\epsilon}(\delta)$ such that $\forall\epsilon<\hat{\epsilon}(\delta)$ and m large enough, the following inequality holds:

$$\mathbb{E}\{\text{distance}(\vec{\theta}^{mL}_\epsilon,\Theta)\}<\delta. \tag{30}$$

This leads to the following inequality:

$$\limsup_N\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{X^n_{\pi^*,\vec{\theta}^n}\}\leq X^*+\epsilon\frac{L+2}{2}B_1^2+B_1(2B_2+3)\delta+\delta B_1. \tag{31}$$

Now let $B_1(2B_2+4)\delta=\gamma$, while letting $\epsilon<\hat{\epsilon}(\delta)$ go to 0. This yields the result of Theorem 2, namely,

$$\limsup_\epsilon\limsup_N\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{X^n_{\pi^*,\vec{\theta}^n}\}\leq X^*+\gamma. \tag{32}$$

In other words, we have shown that to establish asymptotic optimality of our recursive algorithm (17) and the policy π^* (i.e., Theorem 2), all we need is to establish the validity of Theorem 3 as well as Lemmas 1-4. The proof of the theorem and the lemmas are given in Appendices A and B.

V. Case Study

As an application of stochastic optimization, we study a power efficient scheduling problem in a multihop MIMO wireless network.

A. Network Description

Consider a MIMO system of M nodes, in which each node is equipped with n_t transmit antennas and n_r receive antennas. We identify each node with a unique integer index taken from $\{1, 2, \ldots, M\}$. Let $\mathcal{N}$ denote the set of all nodes in the network. A MIMO link denotes a logical connection from one node to another node over the RF channels between multiple antennas at the transmitter and the receiver, i.e., each MIMO link corresponds to an $n_t \times n_r$ channel matrix. The i^{th} $(i = 1, 2, ..., L)$ link is denoted by $l_i \triangleq (a_i, b_i)$, where a_i and b_i denote the indices of the transmitter and receiver of the i^{th} link, respectively. Therefore, we are able to enumerate the link using the subscript i, or specify the group of links that originate from the same transmit node (a_i) or end at the same receive node (b_i). For simplicity, the terms "MIMO link" and "link" are used interchangeably. Let $\mathcal{L}$ denote the set of links in the network. In multihop wireless systems, data are transferred from their source node to their destination nodes through either single or multiple paths supported by the links. The abstraction of this end-to-end connection is called a flow. A total of J flows in the system are assumed. The j^{th} flow is denoted by $f_j \triangleq (s_j, d_j)$, where s_j and d_j indicate the source and destination nodes of the j^{th} flow, respectively. Let $\mathcal{J}$ denote the set of all flows.

System time is divided into equally spaced unit intervals called time slots denoted by n $(= 0, 1, 2, \ldots)$. At time slot n, the channel matrix from node a to node b is described by a random matrix $\mathbf{H}^{(n)}((a, b)) \in \mathbb{C}^{n_t \times n_r}$, where the (i, j)-th element in $\mathbf{H}^{(n)}((a, b))$ denotes the channel gain from the i^{th} transmit antenna of node a to the j^{th} receive antenna of node b. The channel matrix of link l_i is $\mathbf{H}^{(n)}(l_i)$. Note that we slightly abuse the notation, so that $\mathbf{H}^{(n)}(l_i)$ and $\mathbf{H}^{(n)}((a_i, b_i))$ represent the same entity. We assume that $\mathbf{H}^{(n)}(l_i)$ and $\mathbf{H}^{(n)}(l_j)$ are mutually independent if $i \neq j$. Furthermore, we assume for any link l_i, the random sequence of channel matrices $\{\mathbf{H}^{(n)}(l_i),\ n \geq 0\}$ is stationary and uniformly mixing [6, p.p. 345] (ϕ-mixing [9, p.p. 356]). The global channel state of the network is denoted by $\xi^n = (\Pi^{(n)}(1), \Pi^{(n)}(2), \ldots, \Pi^{(n)}(M))$, where $\Pi^{(n)}(a) = \{\mathbf{H}^{(n)}((a, b)) : b \in \{1, 2, \ldots, M\}\}$.

At any given time n, and with the full knowledge of the global channel state $(\Pi^{(n)}(1), \ldots, \Pi^{(n)}(M))$, a controller (in general, centralized) is responsible to allocate the following:

1) <u>System Resource</u>:

1) $\vec{P}^n = [P^n(1), P^n(2), \ldots, P^n(M)]$: The array of the transmit powers P^n_m of node m $(= 1, 2, \ldots, M)$ at time n.
2) $\vec{R}^n = [R^n(l_1), R^n(l_2), \ldots, R^n(l_L)]$: The array of the transmit link rates $R^n(l_i) \geq 0$ on link l_i at time n.
3) $\mathbf{C}^n = \{C^n(f_j, l_i);\ j = 1, \ldots, J,\ i = 1, \ldots, L\}$: The array of the transmit rates over link l_i belonging to flow f_j at time n.

   Note that $C^n(f_j, l_i) \geq 0$ and $\sum_{j=1}^{J} C^n(f_j, l_i) \leq R^n(l_i)$.

To ensure reliable transfer of information over the time varying wireless medium, the above quantities have to be chosen according to the capacity region of the network. In general, this is a non-trivial question, as it is closely related to the information theoretic capacity of ad-hoc networks, which is an open problem. Instead of solving the open problem of the capacity of a wireless ad-hoc network, we take a divide-and-conquer approach. We intentionally partition the system into manageable subsystems, and enforce some rules to operate each subsystem so that the mutual influence between subsystems is negligible. Admittedly, in doing this, we may sacrifice optimality, depending on the choice of system subblocks and the manner in which subblocks are decoupled. The challenge in identifying and achieving capacity in a given ad-hoc network is in handling interference. One simple and elegant capacity result recently studied in MIMO communication is that of MIMO broadcast (MIMO-BC) when dirty paper coding is used. In this paper, leveraging the BC, we decompose the system into MIMO-BC subsystems, and assume the use of ad hoc rules to avoid interference between each BC unit. In this way, we partially remove the effect of cochannel interference from different BC subsystems, but keep some degree of freedom to manage the cochannel interference within each BC subsystem. Even though the choice of ad-hoc rules via which the MIMO-BC subsystems are decoupled is beyond the scope of this paper, the above MIMO-BC decomposition will satisfy the following:

Assumptions on the Decomposition

(D.1) For each node a, in each time slot, there exists a set of nodes $\mathcal{N}_{BC}(a)$ denoting a MIMO broadcast subsystem in which node a is the transmitter.

(D.2) In time slot n, for $b \notin \mathcal{N}_{BC}(a)$, the interference caused by node a at b is negligible. This is achieved by orthagonalizing the transmissions across MIMO-BC subsystems. We assume that $\mathbf{H}^{(n)}((a,b)) \approx 0$ for $b \notin \mathcal{N}_{BC}(a)$.

(D.3) If $c \in \mathcal{N}_{BC}(a) \bigcap \mathcal{N}_{BC}(b)$, the signals from node a and node b are orthogonal at node c.

The MIMO-BC decomposition along with assumptions (D.1)-(D.3) create a grouping of nodes in the network $\{\mathcal{N}_{BC}(1), \mathcal{N}_{BC}(2), \ldots, \mathcal{N}_{BC}(N)\}$. We refer to this decomposition as the network topology at time n. For notational simplicity, we denote the set of all outgoing links at node m by $\mathcal{E}(m)$, and the set of all incoming links at node m by $\mathcal{F}(m)$. With this particular decomposition, and the existing result on the capacity of MIM-BC, we are ready to describe the allocation constraints (QoS and PHY).

2) Resource and QoS Constraints:

C1− (**Physical Layer Constraint I**)

$(\vec{P}^n, \mathbf{C}^n)$ is said to be feasible under the channel realization $(\Pi^{(n)}(1), \ldots, \Pi^{(n)}(M))$ if and only if, for each node m, there exists a link rate schedule $\{R^n(l_i) \mid i = 1, \ldots, L;\ l_i \in \mathcal{E}(m)\}$ such that, for any node m,

$$\sum_{j=1}^{J} C^n(f_j, l_i) \leq R^n(l_i) \ \ \forall\ l_i \in \mathcal{E}(a),\ i = 1, \ldots, L \tag{33}$$

and

$$\{R^n(l_i) \mid i = 1, \ldots, L;\ l_i \in \mathcal{E}(m)\} \in \ \mathcal{C}_{\text{BC}}(P_a^n, \Pi^{(n)}(m)), \tag{34}$$

where $\mathcal{C}_{\text{BC}}(P_a^n, \Pi^{(n)}(m))$ is the feasible link rate region of the MIMO Gaussian broadcast channels formed by node m and all the links originating at node m. In theory, the capacity region of the MIMO gaussian broadcast channels [18] can be characterized by the total power transmitted P_m^n and the channel states $\Pi^{(n)}(m)$.

C2− (**Physical Layer Constraint II**)

The peak transmit power of node m is limited to $P_m^{\max}$, that is,

$0 \leq P_m^n \leq P_m^{\max}$.

C3− (**Queue Stability**)

The queue buildup at each relaying node m is stable, i.e., the long-term average of information bit rate of flow f_j entering node m, given that node m is neither the source nor the destination for flow f_j, equals the long-term average of the information bit rate of flow f_j leaving node m. This requirement is satisfied by fulfilling the following equation:

$$\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\left\{\sum_{l\in\mathcal{E}(m)}\mathbb{E}(C^n(f_j,l_i))-\sum_{l\in\mathcal{F}(m)}\mathbb{E}(C^n(f_j,l_i))\right\}\leq 0. \tag{35}$$

C4− (**Minimum Flow Rate Constraint**)

Each flow f_j $(j=1,\ldots,J)$ is guaranteed a long-term average rate equal to the target rate $v(f_j)$. That is,

$$\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\left\{\sum_{l\in\mathcal{E}(m)}\mathbb{E}(C^n(f_j,l_i))-\sum_{l\in\mathcal{F}(m)}\mathbb{E}(C^n(f_j,l_i))\right\}\leq\begin{cases}v(f_j), & \text{if } m=d_j\\ -v(f_j), & \text{if } m=s_j\end{cases} \tag{36}$$

Note that (C3) and (C4) can be combined: For any node m and any flow f_j,

$$\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\left\{\sum_{l\in\mathcal{F}(m)}\mathbb{E}(C^n(f_j,l_i))-\sum_{l\in\mathcal{E}(m)}\mathbb{E}(C^n(f_j,l_i))\right\}\leq\nu_j(m), \tag{37}$$

where

$$\nu_j(m)=\begin{cases}v(f_j), & \text{if } m=d_j \text{ (destination node of the } j^{th} \text{ flow)}\\ -v(f_j), & \text{if } m=s_j \text{ (source node of the } j^{th} \text{ flow)}\\ 0, & \text{otherwise.}\end{cases} \tag{38}$$

3) Objective: The objective of this work is to develop a locally centralized cross-layer scheduling algorithm which minimizes the average transmit power consumption while maintaining minimum end-to-end throughput requests for time-varying channels. In other words, given the set of control policies that satisfy (**C1**)-(**C4**), we are interested in the one which minimizes the long-term average transmit power, $\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\sum_{m=1}^{M}\mathbb{E}(P_m^n)$.

B. Stochastic Optimization Formulation

The network resource allocation problem discussed above can be mathematically formulated as a stochastic optimization problem of the form studied in previous sections.

Primal problem (P-MIMO-BC)

$$\min \limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \sum_{m=1}^{M} \mathbb{E}(P_m^n)$$

subject to

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \left\{ \sum_{l\in\mathcal{F}(m)} \mathbb{E}(C^n(f_j,l_i)) - \sum_{l\in\mathcal{E}(m)} \mathbb{E}(C^n(f_j,l_i)) \right\} \leq \nu_j(m)$$

and

$$\text{for all } n: \quad \begin{cases} C^n(f_j,l_i) \geq 0, & \text{for } f_j \in \mathcal{J}, \text{ and } l_i \in \mathcal{L} \\ \sum_{j=1}^{J} C^n(f_j,l_i) \leq R^n(l_i), & \text{for } l_i \in \mathcal{L} \\ 0 \leq P_m^n \leq P_m^{\max}, & \text{for } m \in \mathcal{N} \\ \{R^n(l_i) \mid l_i \in \mathcal{E}(m)\} \in \mathcal{C}_{\text{BC}}(P_a^n, \Pi^{(n)}(m)), & \text{for } m \in \mathcal{N} \end{cases} \tag{39}$$

1) Asymptotically Optimal (Iterative) Control Policy: We use the result obtained in previous sections to provide an iterative procedure for solving Problem (P-MIMO-BC).

Define $\mathcal{D}_{\text{BC}}(\xi^n)$ as the feasible region of the primal problem (P) defined by the set of constraints given in (39). To decide the control actions and given the reference parameters $\vec{\theta}$, we need to solve the following minimization for policy π^*:

$$\min_{\mathcal{D}_{\text{BC}}(\xi^n)} \sum_{m=1}^{M} P_m^n + \sum_{m=1}^{M} \sum_{i=1}^{L} \sum_{j=1}^{J} \left\{ \sum_{l\in\mathcal{F}(m)} \mathbb{E}(C^n(f_j,l_i)) - \sum_{l\in\mathcal{E}(m)} \mathbb{E}(C^n(f_j,l_i)) - \nu_j(m) \right\}. \tag{40}$$

After rearranging some terms, (40) can be rewritten as:

$$\min_{\mathcal{D}_{\text{BC}}(\xi^n)} \sum_{m=1}^{M} P_m^n + \sum_{m=1}^{M} \sum_{i=1}^{L} \sum_{j=1}^{J} (\theta_j(b_i) - \theta_j(m)) I_{\{l_i\in\mathcal{E}(m)\}} C^n(f_j,l_i) + \sigma_0, \tag{41}$$

where $\sigma_0 = -\sum_{j=1}^{J} \sum_{m=1}^{M} \theta_j(m)\nu_j(m)$.

Next, we briefly explain the procedure to implement the outer loop, i.e., recursive algorithm (17) as follows.

$$\theta_j^{n+1}(m) = \theta_j^n(m) + \epsilon \; \Pi_H \left[\sum_{l_i \in \mathcal{F}(m)} C^{*n}(f_j, l_i) - \sum_{l_i \in \mathcal{E}(m)} C^{*n}(f_j, l_i) + \nu_j(m) \right], \tag{42}$$

where the superscript '$*$' denotes the optimal controlled variables under policy π^*.

We note that, because of (A.1)-(A.3), $\mathcal{D}(\xi^n) = \mathcal{D}(\Pi^{(n)}(1), \ldots, \Pi^{(n)}(M))$ can be written in a product form. Therefore, the solution to the inner optimization above can be obtained by solving two M separated optimization subproblems at each node m for each subsystem $\mathcal{N}_{BC}(m)$. Furthermore, in (42), the updates can be accomplished locally at each node a by exchanging the values of $C^{*n}(f_j, l_i)$ with its predecessor and successor nodes in the directed graph $G(E, V)$ which describes the system topology. Therefore, the overall algorithm can be implemented in a distributed way by exchanging certain parameters with neighboring nodes, namely the updated dual variables and the optimal flow rates to the inner optimization. We do not discuss this reduction in complexity and/or distributed implementation of the schemes in this paper. However, we refer the interested reader to [14] [13].

VI. Numerical Example

In this section, we use the developed algorithms to identify the benefits of multi-user detection in MIMO ad-hoc networks.

A. Virtual Geographical Cells

In this subsection, similar to [15], we look into an example in which the geographic plane is virtually partitioned into isolated cells. Exploiting the information of its instantaneous position, each node is aware of which cell it belongs to. This can be done with the aid of the *global positioning system* (GPS). Each cell is assigned a frequency band. Two cells that are adjacent to each other are assigned non-overlapping frequency bands. Nodes in the same cell can transmit using the frequency band assigned to that cell. However, in each cell, the associated frequency band is used exclusively by only one node to send data at any given time. In other words, within each cell, the transmissions are scheduled in a TDMA fashion. In Section VI-B, we briefly discuss the impact of such constraints. We define the neighboring cells of node a as the set consisting of the cell where node a is present and the cells adjacent to it. For each transmit node, its potential one-hop receivers lie in this set of neighboring cells. In other words, for each node a, the set $\mathcal{N}_{BC}(a)$ consists of all other nodes in the cell to which a belongs and all nodes in its adjacent cells. In practice, all the transmissions that take place outside a given node's neighboring cells cause interference. However, we assume here that with appropriate frequency planning, the cochannel interference is negligible. We assume that the geographic cell partition requires K non-overlapping frequency bands to avoid neighboring cells using the same frequency band[4]. To implement this system, each node is equipped with K transceivers tuned to K frequency bands. If a node is selected to transmit, it uses one frequency band to send data, and the remaining $(K-1)$ frequency bands to receive the data sent from the $(K-1)$ transmitters in its neighboring cells (except its own cell). If a node is not selected to transmit, it listens to all K frequency bands to receive signals from the transmitters in its neighboring cells.

We note that, with this approach, we lose frequency efficiency by a factor of K; however, we can implement our scheduling algorithm in a decentralized manner.

[4]For example, for the hexagon cell, this requires 7 frequency bands.

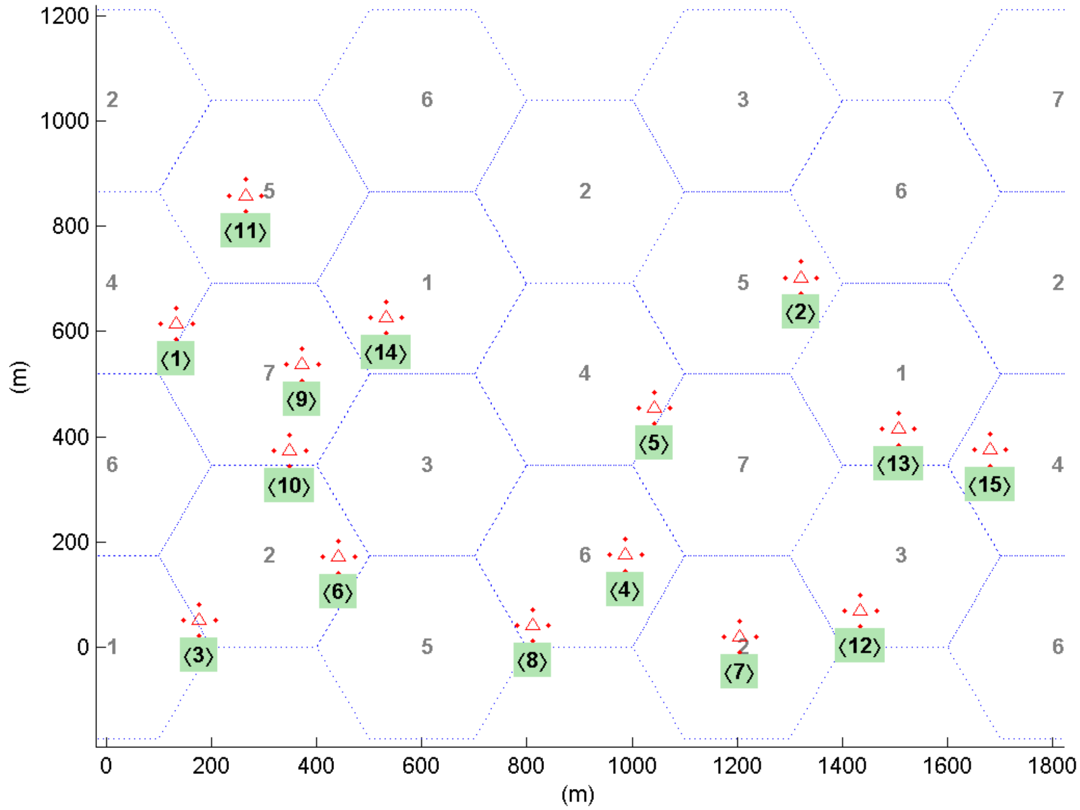

Fig. 2. System Topology with Partition of Hexagonal Cells.

B. Numerical Results

We consider a MIMO ad-hoc network with 15 nodes, and the network topology is depicted in Figure 2. In the center of each cell we label seven non overlapping frequency bands from 1 to 7. Cells labeled with the same number operate in the same frequency band. The small triangles in Figure 2 represent the starting locations of nodes. Nodes are identified by their index. Each node is equipped with four antennas, represented by four dots around the triangle in Figure 2. The peak transmit power of each node is limited to 1Watt. We set the pathloss exponent to 4, the available bandwidth to 1MHz, and the carrier frequency to 2.4GHz. We assume additive white Gaussian noise, which contributes 5.7e-16Watts received noise power at each receiving node. The nodes are initially placed in the cells as illustrated in Figure 2. In every time slot, each node independently moves a units in a direction u relative to its initial position, where a is a uniform random variable distributed over an interval between 0 and the length of the cell's side, and u is a two dimensional unit vector with phase uniformly distributed over $[0\ \pi]$. Note that the mobility model used in this example was chosen for illustration purpose, which may or may not capture the real systems' behavior.

To model the time-varying behavior, we let each node randomly move around the neighboring cells of the cell where it is present when the simulation starts. This model captures variations in the channel gains, while keeping the system topology quasi-static. To model the effect of multipath, we assume each channel is subjected to five multipaths reflected from randomly positioned scatterers. We consider two network traffic scenarios, each with four end-to-end flows. The simulations are run for equal flow demands at 3Mbps, 4Mbps and 5Mbps. Three MIMO schemes are considered in order to capture the impact of multi-user detection: 1) *MVDR beamforming* (BF), 2) point-to-point *spatial multiplexing* (SM), and 3) *dirty paper coding* (DP) for MIMO broadcast channels.

In the first example, there are four end-to-end flow commodities: node 1 to node 15, node 3 to node 2, node 11 to node 7, and node 9 to node 12. In this setting, the overall traffic has a left to right flow.

The average power consumption under our algorithms is given in Figure 3. When the flow rate demand is low, the three MIMO techniques (BF, SM, DP) perform approximately the same. This suggests that TDM is near optimal at low flow rate demands. However, as the flow rate demand increases, spatial multiplexing and dirty paper coding outperform the beamforming scheme, while spatial multiplexing remains comparable to dirty paper coding.

In the second example, we reverse the directions of two flows in the first example, i.e., the flows are between nodes 1 and 15, nodes 3 and 2, nodes 7 and 11, and nodes 12 and 9. In this case, two flows send data from the right to the left of the network, while the remaining two send data from the left to the right. The total power consumption versus the requested throughput under different MIMO techniques is plotted in Figure 4. Here we see that at high rates, SM outperforms BF. In this example, this is because SM can use up to four eigen-modes of the MIMO channel to transmit data, while BF uses only the best one. Further, the dirty paper coding scheme significantly outperforms both SM and BF. This is due to the fact that flows traverse across the network in opposite directions. More specifically, while beamforming and spatial multiplexing are limited to only point-to-point link transmission, the dirty paper coding scheme allows transmission to take place in both directions concurrently, resulting in significant improvements over BF and SM.

VII. Discussion

In this paper, we have developed distributed algorithms which minimize the average transmit power of the system, while maintaining end-to-end flow rates, if feasible. The numerical results show that the benefit of multi-user transmission schemes is significant when the average flow rate demand is high.

We emphasis that, in this work, we treat the MIMO broadcast decomposition and neighbor associations as given, in order to derive cross-layer optimal solutions. The cost-benefit analysis of such decomposition, on the other hand, remains largely untouched. In other words, optimizing the decomposition scheme itself is an interesting area of future research. It will also be interesting to see how to adaptively change the decomposition with respect to the channel states and the

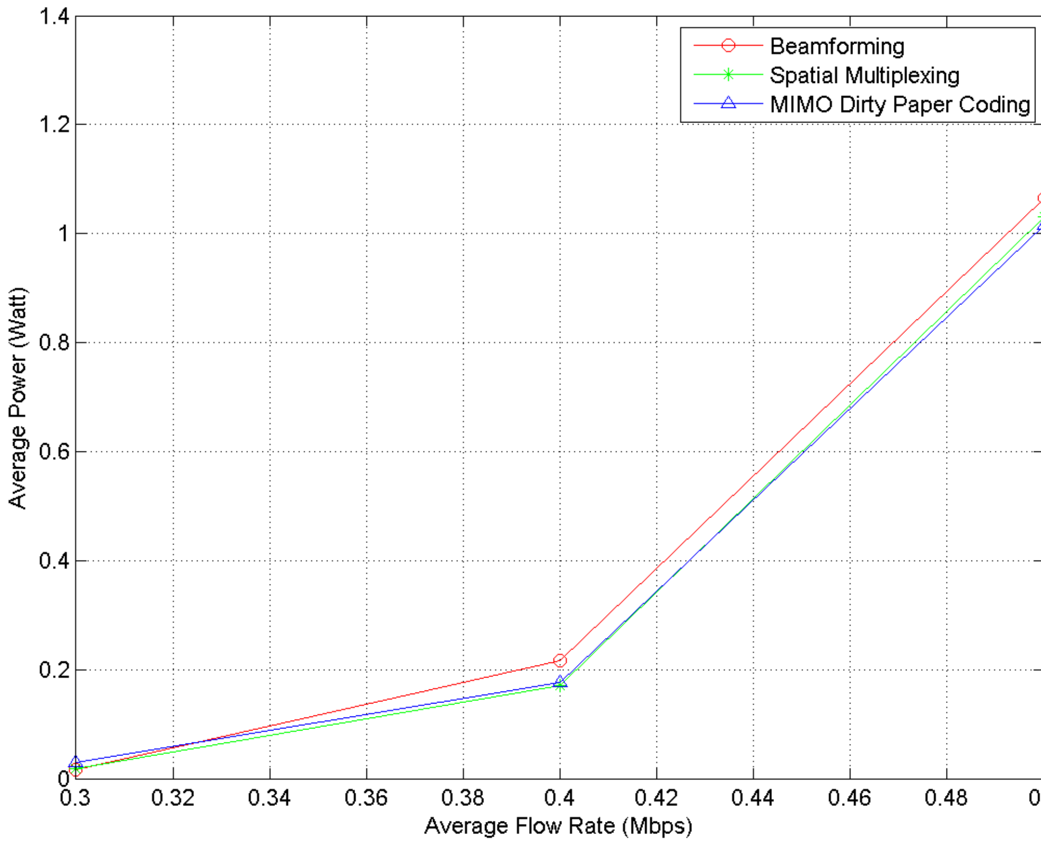

Fig. 3. Average Transmit Power Consumptions of Four Flows (Case 1).

queue backlogs.

In our setting, the information flows are passed over links in a decode-store-forward fashion. Incorporating ideas such as cooperative relaying, as well as network coding, are other interesting and challenging topics for future research.

REFERENCES

[1] R. Agrawal and V. Subramanian. Optimality of certain channel-aware scheduling policies. In *Proc. 40th Annu. Allerton Conf. Communications, Control, and Computing*, pages 1532–1541, 2002.

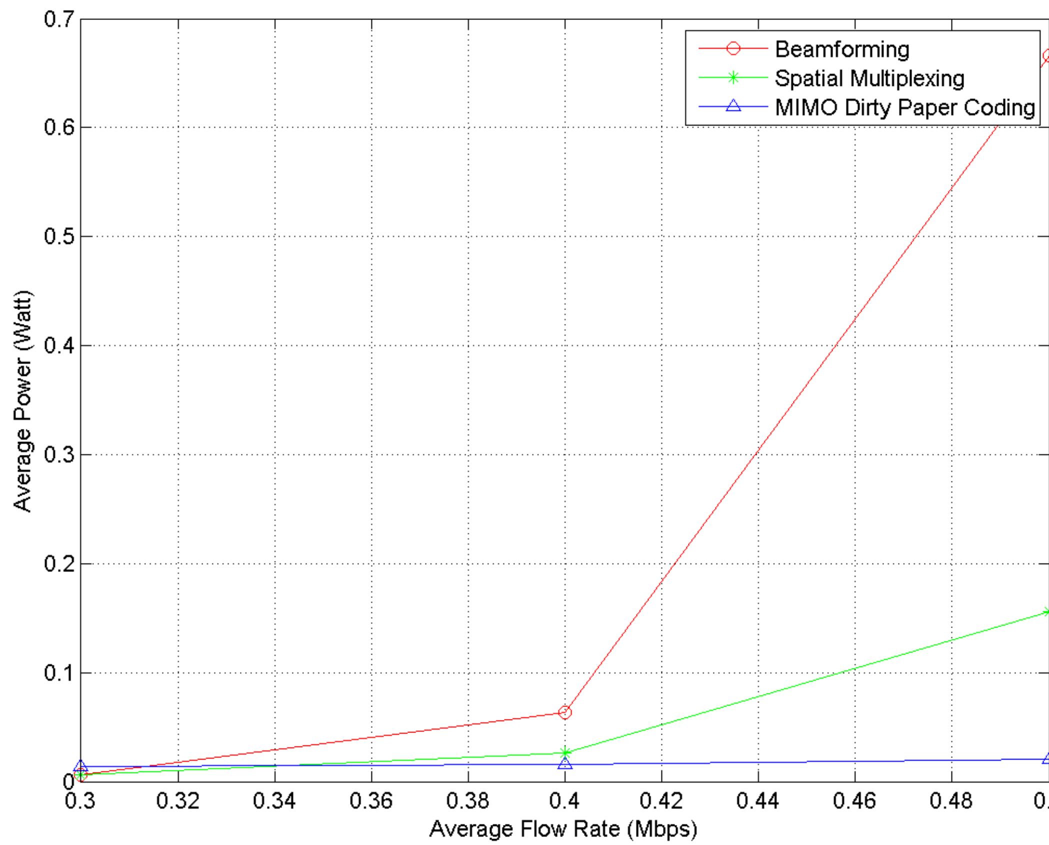

Fig. 4. Transmit Power Consumptions of Four Flows (Case 2)

[2] Rajeev Agrawal, Vijay Subramanian, and Randall A. Berry. Joint scheduling and resource allocation in cdma systems.

[3] Jean-Pierre Aubin and Arrigo Cellina. *Differential Inclusions: Set-Valued Maps and Viability Theory (Grundlehren Der Mathematischen Wissenschaften, Vol 264)*. Springer, 7 1984.

[4] Dimitri Bertsekas. *Convex Analysis and Optimization*. Athena Scientific, 4 2003.

[5] Dimitri P. Bertsekas. *Nonlinear Programming*. Athena Scientific, 2nd edition, 9 1999.

[6] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence (Wiley Series in Probability and Statistics)*. Wiley-Interscience, 2rev ed edition, 9 2005.

[7] Harold J. Kushner. *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic Systems Theory (Signal Processing, Optimization, and Control)*. The MIT Press, 4 1984.

[8] Harold J. Kushner and Dean S. Clark. *Stochastic Approximation Methods for Constrained and Unconstrained Systems (Applied Mathematical Sciences, Vol. 26)*. Springer, 8 1978.

[9] Harold J. Kushner and G. George Yin. *Stochastic Approximation and Recursive Algorithms and Applications (Stochastic Modelling and Applied Probability)*. Springer, 2 edition, 7 2003.

[10] H.J. Kushner and P.A. Whiting. Convergence of proportional-fair sharing algorithms under general conditions. *Wireless Communications, IEEE Transactions on*, 3(4):1250–1259, July 2004.

[11] J.-W. Lee, R.R. Mazumdar, and N.B. Shroff. Opportunistic power scheduling for dynamic multi-server wireless systems. *Wireless Communications, IEEE Transactions on*, 5(6):1506–1515, June 2006.

[12] Yih-Hao Lin. *Stochastic Optimization and Its Applicaiton in Scheduling over Random Time Varying Channels*. PhD thesis, University of California, San Diego, 2007.

[13] Yih-Hao Lin and Rene L. Cruz. Opportunistic link scheduling, power control, and routing for multi-hop wireless networks over time varying channels. In *Proceedings of Annual Allerton Conference on Communication, Control, and Computing, 2005.*, 2005.

[14] Yih-Hao Lin and R.L. Cruz. Power control and scheduling for interfering links. In *Information Theory Workshop, 2004. IEEE*, pages 288–291, 24-29 Oct. 2004.

[15] M.J. Neely. Energy optimal control for time-varying wireless networks. *Information Theory, IEEE Transactions on*, 52(7):2915–2934, July 2006.

[16] Alexander L. Stolyar. On the asymptotic optimality of the gradient scheduling algorithm for multiuser throughput allocation. *Oper. Res.*, 53(1):12–25, 2005.

[17] Alexander L. Stolyar. Greedy primal-dual algorithm for dynamic resource allocation in complex networks. *Queueing Syst. Theory Appl.*, 54(3):203–220, 2006.

[18] S. Vishwanath, N. Jindal, and A. Goldsmith. Duality, achievable rates, and sum-rate capacity of gaussian mimo broadcast channels. *Information Theory, IEEE Transactions on*, 49(10):2658–2668, Oct. 2003.

Appendix A

A. Proof of Proposition 1

***Proposition** 1:* For all $\vec{\theta} \in \mathbb{R}^{r,+}$, $V(\vec{\theta})$ is a lower bound for the solution to the stochastic optimization problem, i.e., $V(\vec{\theta}) \leq X^*$.

Proof:

Given $\vec{\theta} \in \mathbb{R}^{r,+}$, the optimal value of the following optimization with respect to $\vec{\theta}$ is referred to as the dual function of (8) with respect to $\vec{\theta}$:

$$\varphi(\vec{\theta}) := \inf_{\pi \in \Gamma} \left\{ \limsup_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{X_\pi^n\} + \vec{\theta} \cdot \limsup_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{\vec{Y_\pi^n}\} \right\} \tag{A-1}$$

where Γ is the collection of randomized admissible policies. Now, from

$$\limsup_n a_n + \limsup_n b_n \geq \limsup_n (a_n + b_n), \tag{A-2}$$

we have

$$\varphi(\vec{\theta}) \geq \inf_{\pi \in \Gamma} \limsup_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{X_\pi^n + \vec{\theta} \cdot \vec{Y_\pi^n}\} = V(\vec{\theta}). \tag{A-3}$$

On the other hand, by weak duality, the value of the dual function over $\vec{\theta} \geq 0$ is always smaller than the optimal value of the primal problem (Problem 1). Therefore,

$$V(\vec{\theta}) \leq X^*, \text{ for } \vec{\theta} \geq 0,$$

and, hence, the assertion of the proposition. ■

Appendix B

A. Boundedness of the Dual Optimal Set Θ

Lemma *5:* $V(\vec{\theta})$ and the dual function of Problem 1 are equivalent. Specifically,

$$V(\vec{\theta}) = \inf_{\pi \in \Gamma} \left\{ \limsup_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{X_\pi^n\} + \vec{\theta} \cdot \limsup_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{\cdot \vec{Y_\pi^n}\} \right\}$$

Proof:

$$
\begin{aligned}
&\inf_{\pi\in\Gamma}\left\{\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{X_\pi^n\}+\vec{\theta}\cdot\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{\cdot\vec{Y_\pi^n}\}\right\}\\
\geq\ &\inf_{\pi\in\Gamma}\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{X_\pi^n+\vec{\theta}\cdot\vec{Y_\pi^n}\}\\
=\ &\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{V(\vec{\theta},\xi^n)\}\\
\overset{(a)}{=}\ &\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{V(\vec{\theta},\xi^0)\}\\
=\ &\mathbb{E}\{V(\vec{\theta},\xi^0)\}\\
=\ &\mathbb{E}\{X_{\pi^*}^0\}+\vec{\theta}\cdot\mathbb{E}\{\vec{Y}_{\pi^*}^0\}\\
\overset{(b)}{=}\ &\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{X_{\pi^*}^n\}+\vec{\theta}\cdot\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{\vec{Y}_{\pi^*}^n\}\\
\geq\ &\inf_{\pi\in\Gamma}\left\{\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{X_\pi^n\}+\vec{\theta}\cdot\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{\cdot\vec{Y_\pi^n}\}\right\},
\end{aligned}
$$

where π^* is the stationary policy defined in (15) and the equalities (a) and (b) above follows from the stationarity of $\{\xi^n\}$. Now noticing that $V(\vec{\theta})=\inf_{\pi\in\Gamma}\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\{X_\pi^n+\vec{\theta}\cdot\vec{Y_\pi^n}\}$, we have the assertion of the lemma. ■

Fact *2:* (Nonlinear Farkas' Lemma, [4] Proposition 3.5.4, p.p. 204) Let C be a nonempty convex subset of $\mathbb{R}^n$, and let $f: C\to\mathbb{R}$ and $g_j: C\to\mathbb{R}$, $j=1,\dots,r$, be convex functions. Consider the set F given by

$$F=\{x\in C\mid \vec{g}(x)\leq 0\}\,,$$

where $\vec{g}(x)=(g_1(x),\dots,g_r(x))$, and assume that

$$f(x)\geq 0, \forall x\in F$$

Consider the subset Q^* of $\mathbb{R}^r$ given by

$$Q^* = \left\{ \vec{\theta} \;\middle|\; \vec{\theta} \geq 0, f(x) + \vec{\theta} \cdot \vec{g}(x) \geq 0, \ \forall x \in C \right\}$$

Then Q^* is nonempty and compact if and only if there exists a vector $\bar{x} \in C$ such that

$$g_j(\bar{x}) < 0, \forall j = 1, \ldots, r.$$

***Remark** 1:* The equivalence between Problem 1 and a convex optimization was established in [12], which infers the use of Nonlinear Fakars' Lemma is legitimate. As a simple application of Lemma 2, it follows that Θ is compact if and only if there exists an admissible policy π satisfying $\limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{\vec{Y_\pi^n}\} < 0$ strictly. This is summarize in the following proposition.

Proposition 2: Θ is compact if and only if there exists an admissible policy π such that

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{\vec{Y_\pi^n}\} < 0.$$

Moreover, if $\limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{\vec{Y_\pi^n}\} < -\vec{\delta}$ for some $\vec{\delta} = (\delta, \delta, \ldots, \delta)$ with $\delta > 0$, we have $\sum_{i=1}^{r} \theta_i^* \leq (B_1 - X^*)/\delta$ for each $\vec{\theta}^* = (\theta_1^*, \ldots, \theta_r^*) \in \Theta$, where B_1 is a constant defined in (A.1).

Proof: The equivalence between the compactness of Θ and the interior admissible solution follows from Fact 2. The second statement is justified as follows:

$$\begin{aligned}
B_1 - \delta \sum_{i=1}^{r} \theta_i^* &\geq \limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{X_\pi^n\} - \delta \sum_{i=1}^{r} \theta_i^* \\
&\geq \limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{X_\pi^n\} + \vec{\theta}^* \cdot \limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{\cdot \vec{Y_\pi^n}\} \\
&\geq V(\vec{\theta}^*) \\
&\overset{(*)}{=} X^*,
\end{aligned}$$

where the equality (*) follows from Lemma 5. ■

APPENDIX C

A. Review: Stochastic Approximation Algorithms

The generic stochastic approximation algorithms consist of three major components:

1) *Step Size*: $\epsilon > 0$,
2) *State Variable*: $\theta_\epsilon^n = (\theta_{\epsilon,1}^n, \cdots, \theta_{\epsilon,r}^n) \in \mathbb{R}^r$, which is updated in each iteration,
3) *Search Direction*: $Y_\epsilon^n = (Y_{\epsilon,1}^n, \cdots, Y_{\epsilon,r}^n) \in \mathbb{R}^r$; Y_ϵ^n is a stochastic process that depends on an exogenous process $\{\xi_\epsilon^n\}$.

Note that the variables are subscripted with step size ϵ to emphasize the fact that the selection of these variables may depend on the step size that we take.

The evolution of variable θ_ϵ^n is determined by the projected stochastic difference equation given below:

$$\theta_\epsilon^{n+1} = \Pi_H \left(\theta_\epsilon^n + \epsilon Y_\epsilon^n\right). \tag{C-1}$$

Define the reflection term $Z_\epsilon^n = (Z_{\epsilon,1}^n, \cdots, Z_{\epsilon,r}^n) \in \mathbb{R}^r$ by rewriting (C-1) as

$$\theta_\epsilon^{n+1} = \theta_\epsilon^n + \epsilon Y_\epsilon^n + \epsilon Z_\epsilon^n. \tag{C-2}$$

Note that, since the constraint set H is restricted to $\mathbb{R}^{+,r}$, θ_ϵ^n is component-wise nonnegative. Stochastic approximation is concerned with the asymptotic behavior ($n \to \infty$) of $\{(\theta_\epsilon^n, Z_\epsilon^n),\ n = 0, 1, 2, \ldots\}$ for a sufficiently small step size. The asymptotic behavior of the above stochastic sequences is studied via continuous-time interpolated processes as follows:

We are particularly interested in the evolution of $\{\theta_\epsilon^n\}$ in the following two extreme cases. First, how does the sequence $\{\theta_\epsilon^n\}$ behave as the step size ϵ goes to zero? Second, does the sequence converge as n goes to infinity? These questions can be answered by looking into the interpolated process of sequences $\{\theta_\epsilon^n\}$, $\{Y_\epsilon^n\}$ and $\{Z_\epsilon^n\}$. The interpolated process of a sequence with respect to step size ϵ is the "continuous-time" interpolations[5] of the sequence with the interpolation interval equal to ϵ. Following [7], the interpolations $\theta_\epsilon(t)$, $Y_\epsilon(t)$ and $Z_\epsilon(t)$ are defined as follows:

$$\theta_\epsilon(t) = \begin{cases} \theta_\epsilon^0 & \text{for } t < 0 \\ \\ \theta_\epsilon^n & \text{on the time interval } [n\epsilon, n\epsilon + \epsilon) \end{cases} \tag{C-3}$$

$$Y_\epsilon(t) = \begin{cases} \epsilon \sum_{i=0}^{\lfloor t/\epsilon \rfloor - 1} Y_\epsilon^i, & \text{where } t \geq 0 \\ \\ 0 & \text{for } t < 0 \end{cases} \tag{C-4}$$

$$Z_\epsilon(t) = \begin{cases} \epsilon \sum_{i=0}^{\lfloor t/\epsilon \rfloor - 1} Z_\epsilon^i, & \text{where } t \geq 0 \\ \\ 0 & \text{for } t < 0 \end{cases} \tag{C-5}$$

where $\lfloor x \rfloor$ denotes the integer part[6] of x. An example of $\theta_\epsilon(t)$ is illustrated in Figure 5.

In this paper, we use Fact 3 stated below, whose original statement and proof are given in [7].

***Fact** 3:* [9, Chapter 8, Theorem 2.5] Let $\{\xi_\epsilon^n;\ n \geq 0\}$ be random variables over a certain complete and separable metric space Ξ; $\{\mathcal{F}_n^\epsilon\}$ be a sequence of nondecreasing σ-algebras, where $\mathcal{F}_n^\epsilon$ measures $\{\theta_\epsilon^j, Y_\epsilon^{j-1}, \xi_\epsilon^j; j \leq n\}$; and $\mathbb{E}_n^\epsilon$ be the expectation conditioned on $\mathcal{F}_n^\epsilon$. Assume

(A-C.1) $\{Y_\epsilon^n; \epsilon, n\}$ is uniformly integrable,

(A-C.2) there are measurable functions $g_\epsilon^n(\cdot)$ such that

$$\mathbb{E}_n^\epsilon Y_\epsilon^n = g_\epsilon^n(\theta_\epsilon^n, \xi_\epsilon^n), \tag{C-6}$$

[5]The term "continuous-time" is not the real "time," but a virtual continuous index used to describe the asymptotic behavior of the interpolated process.

[6]The integer part of a real number is defined to be the largest integer less than or equal to it.

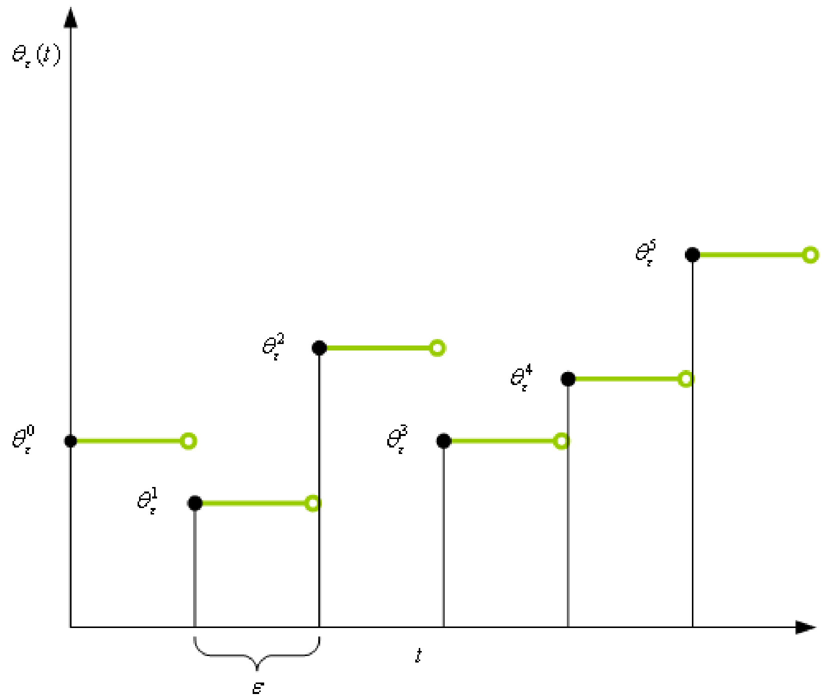

Fig. 5. Interpolated Process $\theta_\epsilon(t)$

(A-C.3) for each $\delta > 0$ there is a compact set $A_\delta \subset \Xi$ such that

$$\inf_{n,\epsilon} P\{\xi_\epsilon^n \in A_\delta\} \geq 1 - \delta, \tag{C-7}$$

(A-C.4) for each θ, the sequences

$$\{g_\epsilon^n(\theta_\epsilon^n, \xi_\epsilon^n);\, \epsilon, n\}\,, \quad \{g_\epsilon^n(\theta, \xi_\epsilon^n);\, \epsilon, n\} \tag{C-8}$$

are uniformly integrable, and

(A-C.5) there is a set-valued function[7] $\mathcal{G}(\cdot)$ that is upper semi-continuous, and for each compact set A, and any sequence α, α_i^n, satisfying

$$\lim_{n,m\to\infty} \sup_{n\leq i\leq n+m} |\alpha_n^i - \alpha| = 0, \tag{C-9}$$

[7]In our work, a set-valued function is a mapping from $\mathbb{R}^r$ to the subsets of $\mathbb{R}^r$.

we have

$$\lim_{n,m} \text{distance} \left[\frac{1}{m} \sum_{i=n}^{n+m-1} \mathbb{E}_n^{\epsilon} g^i(\alpha_n^i, \xi^i), \mathcal{G}(\alpha) \right] I_{\{\xi^n \in A\}} = 0 \quad \text{w.p.1.} \tag{C-10}$$

Then, for any nondecreasing sequence of integers q_ϵ, and for each subsequence of $\{\theta_\epsilon(\epsilon q_\epsilon + \cdot) Z_\epsilon(\epsilon q_\epsilon + \cdot)\}$, $\epsilon > 0$, there exists a further subsequence and a process $(\theta(\cdot), Z(\cdot))$ such that

$$(\theta_\epsilon(\epsilon q_\epsilon + \cdot), Z_\epsilon(\epsilon q_\epsilon + \cdot)) \Rightarrow (\theta(\cdot), Z(\cdot)) \tag{C-11}$$

as $\epsilon \to 0$ through the convergent subsequence, where $\theta(t)$ and $Z(t)$ have Lipschitz continuous paths with probability one.

Fact *4:* Consider the setting for Fact 3. The limit function, $(\theta(\cdot), Z(\cdot))$, satisfies the following differential inclusion: there is an integrable $z(\cdot)$ such that

$$Z(t) = \int_0^t z(s)\, ds, \text{ where } z(t) \in -C(\theta(t)) \text{ for almost all } t, \omega \tag{C-12}$$

and

$$\dot{\theta} \in \mathcal{G}(\theta) + z, \text{ for almost all } t, \omega., \tag{C-13}$$

where the set-valued function $C(x)$ is defined for $x \in H$ (the constraint set defined for the recursive algorithm in (C-1)) as follows. If $x \in H^o$, the interior of H, $C(x)$ contains only the zero element; while if $x \in \partial H$, the boundary of H, $C(x)$ is the infinite convex cone generated by the outer normals at x of the faces on which x lies.

Note that since an "infinitesimal" change in x does not increase the number of active constraints, $C(\cdot)$ is upper-semi-continuous.

In brief, Facts 3 and 4 provide a scenario in which the updated sequence generated by the stochastic approximation algorithm follows the trajectory of the solution to a projected differential inclusion. We use this to 1) study the asymptotic behavior of θ_ϵ^n, and 2) characterize the performance of our proposed iterative algorithm via the study of the corresponding differential inclusion.

B. Stochastic Approximations and Subdifferentials

We apply Fact 3 to our case, where ϵ, $\vec{\theta}^n$ and $\vec{Y}^n_{\pi^*,\theta^n}$ in (14) and (17) form the step size, state variables, and search directions, respectively. In this appendix, we show that these processes satisfy conditions (A-C.1)-(A-C.4).

We note that Conditions (A-C.1), (A-C.3), and(A-C.4), directly follow from Assumptions (A.1) and (A.2). Define $\mathbb{E}^{\epsilon}_n Y^n_\epsilon \triangleq g^n_\epsilon(\alpha, \xi^n_\epsilon)$. By construction of $(X^n_{\pi^*,\alpha}, \vec{Y}^n_{\pi^*,\alpha})$ and the finiteness of support set for ξ^n_ϵ, we have the measurability of function $g^n_\epsilon(\alpha, \xi^n_\epsilon)$, i.e., Condition (A-C.2). What remains is to show that $g^n_\epsilon(\alpha, \xi^n_\epsilon)$ satisfies Condition (A-C.5).

We first note that $g^n_\epsilon(\alpha, \xi^n_\epsilon)$ is a subgradient of $V(\vec{\theta}, \xi^n)$. Recall that from the construction given in (13), it is clear that $V(\vec{\theta}, \xi^n)$ is concave in θ. The following lemma is easy to establish:

Lemma *6:* $g^n_\epsilon(\theta, \xi)$ is a subgradient of $V(\vec{\theta}, \xi)$ at $\theta \in \mathbb{R}^r$ and $\limsup_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \mathbb{E}\{Y^n_{\pi^*,\theta}\}$ is a subgradient of $V(\vec{\theta})$.

Proof: The first part of the lemma is a simple consequence of (13) and the definition of a subgradient. Essentially, using linearity and the equality (12) for $V(\vec{\theta})$, $\mathbb{E}\{V(\vec{\theta}, \xi^n)\}$ and $V(\vec{\theta})$ are concave functions, too. According to [5, Proposition B. 25], this implies that the long-term average, $\limsup_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \mathbb{E}\{Y^n_{\pi^*,\theta}\}$, is a subgradient of $V(\vec{\theta})$. ■

The following lemma establishes the validity of (A-C.5) for subgradient processes:

Lemma *7:* Let $\{\xi^n_\epsilon\}$ be a stationary ϕ-mixing stochastic process and $g^n_\epsilon(\alpha, \xi^n_\epsilon)$ be a subgradient of $V(\alpha, \xi^n_\epsilon)$ at α. Assume that $\exists B > 0$ such that $\|g^n_\epsilon(\alpha, \xi)\| \leq B$ for all α, n, ϵ, ξ. For any collection of vectors α_i^n and vector α satisfying

$$\lim_{n,m\to\infty} \sup_{n\le i\le n+m} |\alpha_n^i - \alpha| = 0, \tag{C-14}$$

we have

$$\lim_{n,m} \frac{1}{m} \sum_{i=n}^{n+m-1} \mathbb{E}_n^\epsilon g_\epsilon(\alpha_n^i, \xi_\epsilon^i) \in \partial\, \mathbb{E}^\epsilon V(\alpha, \xi_\epsilon^0), \tag{C-15}$$

where $\mathbb{E}^\epsilon$ denotes the unconditional expectation, and $\partial\, \mathbb{E}^\epsilon V(\alpha, \xi_\epsilon^0)$ denotes the subdifferential of $\mathbb{E}^\epsilon\{V(\alpha, \xi_\epsilon^0)\}$ at α.

Proof:

$$\begin{aligned} &\lim_{n,m} \frac{1}{m} \sum_{i=n}^{n+m-1} \mathbb{E}_n^\epsilon g_\epsilon^i(\alpha_n^i, \xi_\epsilon^i) \\ = \;& \lim_{n,m} \frac{1}{m} \sum_{i=n}^{n+m-1} \mathbb{E}^\epsilon g_\epsilon^i(\alpha_n^i, \xi_\epsilon^i) + \lim_{n,m} \frac{1}{m} \sum_{i=n}^{n+m-1} (\mathbb{E}_n^\epsilon g_\epsilon^i(\alpha_n^i, \xi_\epsilon^i) - \mathbb{E}^\epsilon g_\epsilon^i(\alpha_n^i, \xi_\epsilon^i)) \end{aligned}$$

The second limit in the equality above goes to zero, since

$$\lim_{n,m} \frac{1}{m} \sum_{i=n}^{n+m-1} \left| \mathbb{E}_n^\epsilon g_\epsilon(\alpha_n^i, \xi_\epsilon^i) - \mathbb{E}^\epsilon g_\epsilon(\alpha_n^i, \xi_\epsilon^i) \right| \overset{(a)}{\le} 2K \lim_{n,m} \frac{1}{m} \sum_{i=n}^{n+m-1} \phi_{i-n} \tag{C-16}$$

$$= 2K \lim_{m} \frac{1}{m} \sum_{i=0}^{m-1} \phi_i \tag{C-17}$$

$$\overset{(b)}{=} 0 \tag{C-18}$$

The inequality (a) above follows Fact 1, and the equality (b) holds because the Cesàro mean of a converging sequence converges to the limit of that sequence. The completion of this proof is achieved using the fact:

$$\lim_{n,m} \frac{1}{m} \sum_{i=n}^{n+m-1} \mathbb{E}^\epsilon g_\epsilon^i(\alpha_n^i, \xi_\epsilon^i) \overset{(c)}{=} \lim_{n,m} \frac{1}{m} \sum_{i=n}^{n+m-1} \mathbb{E}^\epsilon g_\epsilon^i(\alpha_n^i, \xi_\epsilon^0) \overset{(d)}{\in} \partial\, \mathbb{E}^\epsilon V(\alpha, \xi_\epsilon^0), \tag{C-19}$$

where (c) follows the stationarity of ξ_ϵ^n and (d) follows the upper-semi-continuous property of the subdifferentials. ■

In light of Conditions (A-C.1)-(A-C.5) for processes $\vec{\theta}_\epsilon^n$ and $\vec{Y}^n_{\pi^*,\theta^n}$, the following lemma is a simple restatement of Facts 3 and 4.

***Lemma** 8:* Assume that assumptions (A.1)-(A.3) hold. Let $\vec{\theta}_\epsilon^n$ and $\vec{Y}^n_{\pi^*,\theta^n}$ be obtained recursively from (14) and (17), while $g^n_\epsilon(\alpha, \xi^n_\epsilon) \triangleq \mathbb{E}^\epsilon_\ltimes\{\vec{Y}^n_{\pi^*,\alpha}\}$ (is shown to be a subgradient of $V(\alpha, \xi^n_\epsilon)$). Let $\mathcal{G}(\theta)$ be equal to the subdifferential of the concave function $V(\theta)$ at θ. Then the interpolated process of the stochastic approximation $\theta_\epsilon(t)$ and its limiting process $\theta(t)$ have the following convergence property: For any nondecreasing sequence of integers q_ϵ, and for each subsequence of $\{\theta_\epsilon(\epsilon q_\epsilon + \cdot)Z_\epsilon(\epsilon q_\epsilon + \cdot)\}$, $\epsilon > 0$, there exists a further subsequence, which converges, and a process $(\theta(\cdot), Z(\cdot))$ such that

$$(\theta_\epsilon(\epsilon q_\epsilon + \cdot), Z_\epsilon(\epsilon q_\epsilon + \cdot)) \Rightarrow (\theta(\cdot), Z(\cdot)) \tag{C-20}$$

as $\epsilon \to 0$ through the convergent subsequence and where $\theta(t)$ and $Z(t)$ have Lipschitz continuous paths with probability one. Furthermore, $\theta(t)$ is a solution of the following projected differential inclusion:

$$\dot{\theta} \in \partial V(\theta) + z, \quad \theta(0) = \theta^0, \ z(t) \in -C(\theta(t)). \tag{C-21}$$

with initial condition $\theta(0) = \theta^0$.

In the next section, we study the properties of the solution to the differential inclusion equation (C-21). We use the uniqueness of the solution as well as its limiting behavior to characterize the limiting behavior of our proposed dual-controller/stochastic-subgradient-projection method.

C. Asymptotic Behavior of the Limiting Process

***Lemma** 9:* The solution to the projected differential inclusion (C-21) is uniquely determined by the initial condition θ^0.

Proof:

Let us first prove the one dimensional case, in which $\theta \in \mathbb{R}^1$. Let $x(t)$ and $y(t)$ be two solutions of (C-21) with initial conditions x^0 and y^0, respectively. By definition, we have

$$x'(t) = g_x(t) + z_x(t) \tag{C-22}$$

$$y'(t) = g_y(t) + z_y(t), \tag{C-23}$$

where $g_x(t) \in \partial V(x(t))$ and $g_y(t) \in \partial V(y(t))$. The derivative of $1/2\|x(t) - y(t)\|^2$ satisfies:

$$\frac{d}{dt}\frac{1}{2}\|x(t) - y(t)\|^2 = (x(t) - y(t)) \cdot (x'(t) - y'(t)) \tag{C-24}$$

$$= (x(t) - y(t)) \cdot (g_x(t) - g_y(t) + z_x(t) - z_y(t)) \tag{C-25}$$

$$\leq 0. \tag{C-26}$$

The inequality (C-26) can be proved using the following two steps. Without loss of generality, we assume that $x(t) < y(t)$ (note that (C-26) holds trivially when $x(t) = y(t)$).

1) By the definition of subgradient, we have: (i) $V(y(t)) - V(x(t) \leq g_x(t)(y(t) - x(t))$, and (ii) $V(x(t)) - V(y(t)) \leq g_y(t)(x(t) - y(t))$. Summing (i) and (ii), we end up with the inequality $(x(t) - y(t)) \cdot (g_x(t) - g_y(t)) \leq 0$
2) Note that, from Assumption (A.3), $0 \leq x(t) < y(t) \leq K_u$. This means that $x(t)$ either lies in the interior of H or $x = 0$. As a result, the compensation term $z_x(t)$ is either 0 (when $x(t)$ is in the interior of H) or positive (when $x(t) = 0$, the drift pushes $x(t)$ away). On the other hand from $0 < y(t) \leq K_u$, we have that the compensation term $z_y(t)$ is either 0 or negative (again, if $y(t) = K_u$, the drift is compensated). In other words, we have $(x(t) - y(t)) < 0$ and $(z_x(t) - z_y(t)) \geq 0$, which gives us the inequality $(x(t) - y(t)) \cdot (z_x(t) - z_y(t)) \leq 0$.

Since $\|x(t) - y(t)\|$ has non-positive derivative, $\|x(t) - y(t)\|$ must be a non-increasing function in t. Comparing with the initial conditions, we have:

$$\|x(t) - y(t)\| \leq \|x^0 - y^0\| \text{ for } t \geq 0. \tag{C-27}$$

Moreover, if $x^0 = y^0 = \theta^0$, the monotonicity of $\|x(t) - y(t)\|$ implies that $x(t) = y(t) \quad \forall t \geq 0$. This proves the uniqueness of the projected differential inclusion.

We now use Assumption (A.3) to reiterate that set H is an n-dimensional interval. In other words, since H is an n-dimensional interval, projecting a vector to H is equivalent to projecting each component of the vector to the corresponding 1-dimensional intervals. Now from above arguments we know that each projected term satisfies (C-26), generalizing the above for the n-dimensional case.

■

One important implication of Lemma 9 is given by the following corollary:

Corollary *1:* Given the initial condition, θ^0, the statement of Lemma 8 can be strengthened as follows. For any nondecreasing sequence of integers q_ϵ, and for each subsequence of $\{\theta_\epsilon(\epsilon q_\epsilon + \cdot)\}$, $\epsilon > 0$, there exists a further subsequence and a process $\theta(\cdot)$ such that for all t

$$\lim_{\epsilon_k \to 0} \mathbb{E}\{||\theta_\epsilon(\epsilon_k q_{\epsilon_k} + t) - \theta(t)||\} = 0. \tag{C-28}$$

Proof: If a random variable converges in distribution to a constant c, then it converges to c in probability. Furthermore, if the random variable is almost surely bounded, it converges to c in any r^{th} mean.

■

Lemma *10:* If $\theta(t)$ is a continuous solution to (C-21) defined in Theorem 9, its derivative $\dot{\theta}(t)$ is right continuous for all but a countable set of t.

Proof: See [12]. ■

***Lemma** 11:* Let $\theta(t)$ be the unique solution of the projected differential inclusion (C-21). The trajectory of $\theta(t)$ converges to a point. That is,

$$\lim_{t\to\infty} \theta(t) = \theta^*, \tag{C-29}$$

where θ^* is a maximizer of $V(\cdot)$ over H.

Proof: Using Lemma 10 and [3, p. 160, Theorem 2].

■

Now we are ready to restate and prove the first main contribution of our work, characterizing the asymptotic behavior of the iterative stochastic approximation based on a subgradient iteration.

***Theorem** 3:* Let Θ be the set of maximizers of $V(\cdot)$ over H. Assuming (A.1)-(A.3), we have the following result:

$$\forall\delta > 0, \exists\hat{\epsilon} > 0, \text{ such that } \forall\epsilon < \hat{\epsilon}, \mathbb{E}\{\text{ distance}(\,\theta_\epsilon^n, \Theta\,)\,\} < \delta \text{ for all but finitely many } n, \tag{C-30}$$

where distance(θ, A) denotes the distance from point θ to set A.

Proof: Let $\theta(t)$ be the weak-limiting process of $\theta_\epsilon(t)$, with $\theta(0) = \theta_\epsilon^0$. Lemma 9 ensures that $\theta(t)$ is unique. Moreover, from Lemma 11, we know that $\lim_{t\to\infty} \theta(t) = \theta^*$ for some $\theta^* \in \Theta$ with probability one. Since $V(\cdot)$ is concave, Θ is a convex set and hence distance$(\cdot, \Theta)$ is a continuous function. This, together with the dominated convergence theorem, implies that $\lim_{t\to\infty} \mathbb{E}\{\text{distance}(\theta(t), \Theta)\} = 0$. In other words,

$$\exists T \text{ such that } \mathbb{E}\{\text{distance}(\theta(t), \Theta)\} \leq \delta/2 \;\text{ for all } t \geq T. \tag{C-31}$$

Next, we show that (C-31) implies (C-30), or equivalently, that (C-30) is a necessary condition for (C-31). We prove this by contradiction, i.e., we show that assuming (C-30), but not (C-31), results in a contradiction. Suppose (C-30) does not hold, i.e., suppose that $\exists\delta > 0$ such that, for

$\forall\sigma$, $\exists\epsilon < \sigma$ such that for $\forall N^*$, $\exists n > N^*$ for which

$$\mathbb{E}\{\,\text{distance}(\,\theta_\epsilon^n, \Theta\,)\,\} \geq \delta. \tag{C-32}$$

In other words, it is possible to pick a sequence $\{\epsilon_k\}$ such that $\epsilon_k \to 0$ and

$$\mathbb{E}\{\,\text{distance}(\,\theta_{\epsilon_k}^n, \Theta\,)\,\} \geq \delta \qquad \text{for infinitely many } n. \tag{C-33}$$

Pick any $t_0 > T$. Define a sequence of integers n_k in the following iterative manner:

- Let $n_1 = \frac{t_0}{\epsilon_1} + 1$,
- For every $k = 2, 3, \ldots$, define n_k such that

$$n_k > n_{k-1} + t_0(\frac{1}{\epsilon_k} - \frac{1}{\epsilon_{k-1}})$$

  and

$$\mathbb{E}\{\,\text{distance}(\,\theta_{\epsilon_k}^{n_k}, \Theta\,)\,\} \geq \delta. \tag{C-34}$$

  From (C-33), we know that n_k exists.

Next, we show that (C-34) and (C-31) cannot hold simultaneously.

Combining (C-34) with (C-31), we have

$$\mathbb{E}\{\,\text{distance}(\,\theta_{\epsilon_k}^{n_k}, \theta(t_0)\,)\,\} \geq \delta/2. \tag{C-35}$$

Now notice that, from the construction of n_k, $q_{\epsilon_k} = n_k - t/\epsilon_k$ is an increasing sequence for which

$$\theta_{\epsilon_k}(\epsilon_k q_{\epsilon_k} + t_0) = \theta_{\epsilon_k}(\epsilon_k n_k) = \theta_{\epsilon_k}^{n_k}.$$

This, together with (C-35), implies that we have found a converging sequence $\{\epsilon_k\}$ and an increasing sequence q_{ϵ_k} such that

$$\mathbb{E}\{\,\text{distance}(\,\theta_{\epsilon_k}(\epsilon_k q_{\epsilon_k} + t_0), \theta(t_0)\,)\,\} \geq \delta/2, \qquad \forall k. \tag{C-36}$$

On the other hand, from Corollary 1, we know that for the converging sequence ϵ_k and the non-decreasing sequence q_{ϵ_k}, there exists a further subsequence ϵ_{k_j} along which

$$\lim_{\epsilon_{k(j)}\to 0} \mathbb{E}\{\, ||\theta_{\epsilon_k}(\epsilon_{k(j)} q_{\epsilon_{k(j)}} + t_0) - \theta(t_0)||\,\} = 0. \tag{C-37}$$

But this is a direct contradiction to (C-36). In other words, it must be true that (C-31) implies (C-30), i.e., for $\forall \delta > 0$, there exists a $\hat{\epsilon} > 0$ such that, for any $\epsilon < \hat{\epsilon}$, $\mathbb{E}\{\, \text{distance}(\, \theta_\epsilon^n, \Theta\,)\,\} < \delta$ for all but a finite number of times in n. ■

D. Admissibility: Asymptotic behavior of Reflections

Lastly, we have two important results for stochastic approximation regarding the reflection terms. By characterizing the asymptotic behavior of reflection terms, the following lemmas establish admissibility of the stochastic subgradient projection algorithm we introduced.

***Lemma** 12:* Assume (A.1)-(A.3) hold. Let $Z(t)$ be the limiting process of the interpolated process defined for the reflection term, and Θ be the set of maximizers of $V(\cdot)$ over H. If $\Theta \subset H^o$ (interior of H) and $\|Y_\epsilon^n\| \leq B_1$ for some constant B_1, there exists B_2 such that $\|Z(t)\| \leq B_2$ for all $t \geq 0$.

Proof:

Since H is compact, it implies that there exist p points, $\{a_1, a_2, \cdots, a_p\}$ in H such that $H \subset \bigcup_{i=1}^p N_{\delta/2}(a_i)$, where $N_{\delta/2}(a_i)$ denotes the neighborhood of a_i of radius $\delta/2$. Consider the limiting process $(\theta(t), Z(t))$ for an arbitrary initial condition. From Lemmas 9-10, we know that $\theta(t) \to \theta^*$, where $\theta^* \in \Theta$. Since every point in H is covered by a neighborhood $N_{\delta/2}(a_i)$, via the triangular inequality and the monotonically decreasing distance between any two solutions of (C-21), we can obtain the equality $\|\theta(t) - \theta^*\| \leq \delta$ for all $t > T'$. This means that $\theta(t)$ locates in the interior of H after time T', and, hence, no reflection occurs after time T'. In other words, $Z(t) = Z(T')$ for $t \geq T'$.

It remains to show that $Z(T')$ is bounded. By the definition of reflection terms, we have that

$$\|Z_\epsilon(t+s) - Z_\epsilon(t)\| \leq \sum_{i=t/\epsilon}^{(t+s)/\epsilon} \epsilon \|Y_\epsilon^i\| \leq (s+\epsilon)B_1. \tag{C-38}$$

Let $\epsilon \to 0$. For each sample path of the limiting process $Z(t)$, we then have

$$\|Z(t+s) - Z(t)\| \leq sB_1 \text{ with probability one.} \tag{C-39}$$

The boundedness of $Z(T')$ follows from this inequality, and the proof is done by setting $B_2 = T'B_1$ ■

***Lemma** 13:* Assume (A.1)-(A.3) hold. Also, assume that the maximizer(s) of $V(\theta)$ constitutes(constitute) a compact set Θ. Given $\forall\delta > 0$ there exists an $\hat{\epsilon} > 0$ such that

$$\limsup_n \frac{\sum_{i=0}^{n-1} \mathbb{E}\{Z_\epsilon^i\}}{n} \leq \delta$$

for all $\epsilon \leq \hat{\epsilon}$.

Proof: The proof is by contradiction. We first prove that $\forall\delta > 0$, there exists $\hat{\epsilon} > 0$ such that $\limsup_t \frac{\mathbb{E}\{Z_\epsilon(t)\}}{t} \leq \delta$ for all $\epsilon \leq \hat{\epsilon}$. Suppose that this is false, i.e., $\exists\delta$ such that $\forall\hat{\epsilon} > 0$ and $\forall T >$, $\exists\epsilon < \hat{\epsilon}$ and $\exists t > T$ such that

$$\frac{\mathbb{E}\{Z_{\epsilon_k}(t)\}}{t} > \delta. \tag{C-40}$$

This implies that there exists a converging sequence $\epsilon_k \to 0$ and a nondecreasing sequence of integers q_k such that

$$\mathbb{E}\{Z_{\epsilon_k}(\epsilon_k q_k + T)\} > \delta T. \tag{C-41}$$

Combining this with Fact 3 implies that $\mathbb{E}\{Z(T)\} > T\delta$, which contradicts Lemma 12 (set $T = \frac{B_2+1}{\delta}$ using the same constant B_2 defined in Lemma 12).

However, by the definition of interpolated process $Z_\epsilon(t)$, we have

$$\left\| \frac{Z_\epsilon(t)}{t} - \frac{\sum_{i=0}^{n-1} Z_\epsilon^i}{n} \right\| \leq B_1/n \quad \text{for } t/\epsilon - 1 < n \leq t/\epsilon, \tag{C-42}$$

where B_1 is the same constant defined in Lemma 12. As a result, given $\delta > 0$, there exists an $\hat{\epsilon} > 0$ such that $\limsup_n \frac{\sum_{i=0}^{n-1} \mathbb{E}\{Z_\epsilon^i\}}{n} \leq \delta$ for all $\epsilon \leq \hat{\epsilon}$, and we have the assertion of the lemma.

■

Alternatively, we can treat these reflection terms according to whether they touch the lower or the upper boundaries of H. We define the reflection term from above $\check{Z}_\epsilon{}^n \geq 0$ and the reflection term from below $\hat{Z}_\epsilon{}^n \geq 0$ by rewriting the recursive algorithm (C-1) as

$$\theta_\epsilon^{n+1} = \theta_\epsilon^n + \epsilon(Y_\epsilon^n + \hat{Z}_\epsilon{}^n - \check{Z}_\epsilon{}^n). \tag{C-43}$$

Also, we define the interpolated processes $\hat{Z}_\epsilon(t)$ and $\check{Z}_\epsilon(t)$ analogously to $Z_\epsilon(t)$, using respectively, $\hat{Z}_\epsilon{}^n$ and $\check{Z}_\epsilon{}^n$, in lieu of Z_ϵ^n. In fact, the proof of the following lemma closely follows that of Lemma 13, and is omitted in the interest of brevity.

***Lemma** 14:* Assume (A.1)-(A.3) hold. Assume, also, that the maximizer(s) of $V(\theta)$ constitutes(constitute) a compact set Θ. Given $\forall\delta > 0$, there exists an $\hat{\epsilon} > 0$ such that

$$\limsup_n \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}\{\check{Z}_\epsilon{}^i\} \leq \delta$$

for all $\epsilon \leq \hat{\epsilon}$.

APPENDIX D

A. Proofs of Lemma 2, 3, and 4

***Lemma** 2:*

$$\frac{1}{L} \sum_{n=mL}^{mL+L-1} \mathbb{E}\{X_{\pi^*, \vec{\beta}^{mL}}^n\} + \frac{1}{L} \sum_{n=mL}^{mL+L-1} \mathbb{E}\{\vec{\beta}^{mL} \cdot \vec{Y}_{\pi^*, \vec{\beta}^{mL}}^n\} \leq X^* + \frac{1}{L} \sum_{n=0}^{L-1} 2B_1(B_2+1)\phi_n,$$

where $B_2 = \max\{ \|\vec{\theta}^*(1)\|, \ldots, \|\vec{\theta}^*(r_\delta)\|\}$.

Proof:

$$\begin{aligned}
&\frac{1}{L}\sum_{n=mL}^{mL+L-1} \mathbb{E}\{X^n_{\pi^*,\vec{\beta}^{mL}} \mid \vec{\beta}^{mL} = \vec{\theta}^*(i) \} + \frac{1}{L}\sum_{n=mL}^{mL+L-1} \mathbb{E}\{\vec{\beta}^{mL}\cdot \vec{Y}^n_{\pi^*,\vec{\beta}^{mL}} \mid \vec{\beta}^{mL} = \vec{\theta}^*(i) \} \\
&= \frac{1}{L}\sum_{n=mL}^{mL+L-1} \mathbb{E}\{X^n_{\pi^*,\vec{\theta}^*(i)} \mid \vec{\beta}^{mL} = \vec{\theta}^*(i) \} \\
&\quad + \frac{1}{L}\sum_{n=mL}^{mL+L-1} \mathbb{E}\{\vec{\theta}^*(i)\cdot \vec{Y}^n_{\pi^*,\vec{\theta}^*(i)} \mid \vec{\beta}^{mL} = \vec{\theta}^*(i) \} \\
&\overset{(a)}{\leq} \frac{1}{L}\sum_{n=mL}^{mL+L-1} \mathbb{E}\{X^n_{\pi^*,\vec{\theta}^*(i)}\} + \frac{1}{L}\sum_{n=mL}^{mL+L-1} 2B_1\phi_{(n-mL)} \qquad \text{(D-1)} \\
&\quad + \frac{1}{L}\sum_{n=mL}^{mL+L-1} \mathbb{E}\{\vec{\theta}^*(i)\cdot \vec{Y}^n_{\pi^*,\vec{\theta}^*(i)}\} + \frac{1}{L}\sum_{n=mL}^{mL+L-1} \|\vec{\theta}^*(i)\| 2B_1\phi_{(n-mL)} \\
&\overset{(b)}{\leq} X^* + \frac{1}{L}\sum_{n=0}^{L-1} 2B_1(B_2+1)\phi_n. \qquad \text{(D-2)}
\end{aligned}$$

The inequality (a) in (D-1) above follows from Lemma 1 under Assumption (A.1) using

$$\|\mathbb{E}\{X^n_{\pi^*,\vec{\theta}^*(i)} \mid \vec{\beta}^{mL} = \vec{\theta}^*(i)\} - \mathbb{E}\{X^n_{\pi^*,\vec{\theta}^*(i)}\}\| \leq 2B_1\phi_{(n-mL)} \qquad \text{(D-3)}$$

and

$$\|\mathbb{E}\{Y^n_{\pi^*,\vec{\theta}^*(i)} \mid \vec{\beta}^{mL} = \vec{\theta}^*(i)\} - \mathbb{E}\{Y^n_{\pi^*,\vec{\theta}^*(i)}\}\| \leq 2B_1\phi_{(n-mL)}. \qquad \text{(D-4)}$$

The inequality (b) in (D-2) is a consequence of the following result due to the stationarity of $\{\xi^n\}$ and the definition of $V(\cdot)$:

$$\frac{1}{L}\sum_{n=mL}^{mL+L-1} \mathbb{E}\{X^n_{\pi^*,\vec{\theta}^*(i)}\} + \frac{1}{L}\sum_{n=mL}^{mL+L-1} \mathbb{E}\{\vec{\theta}^*(i)\cdot \vec{Y}^n_{\pi^*,\vec{\theta}^*(i)}\} = V(\vec{\theta}^*(i)) \leq X^*. \qquad \text{(D-5)}$$

■

***Lemma** 3:* There exists $L' > 0$ such that, for all $L > L'$,

$$\frac{1}{L}\sum_{n=mL}^{mL+L-1} \mathbb{E}\{V(\vec{\theta}^n,\xi^n)\} \leq \quad X^* + 2B_1(B_2+1)\delta + \epsilon\frac{L+1}{2}B_1^2 + \mathbb{E}\{\text{distance}(\vec{\theta}^{mL},\Theta)\}B_1 + \delta B_1$$

Proof: Assume that L is sufficiently large, so that $\frac{1}{L}\sum_{n=0}^{L-1}\phi_n \le \delta$. We have

$$
\begin{aligned}
&\frac{1}{L}\sum_{n=mL}^{mL+L-1}\mathbb{E}\{V(\vec{\theta}^n,\xi^n)\} && \text{(D-6)}\\
&= \frac{1}{L}\sum_{n=mL}^{mL+L-1}\mathbb{E}\{X^n_{\pi^*,\vec{\theta}^n}\}+\frac{1}{L}\sum_{n=mL}^{mL+L-1}\mathbb{E}\{\vec{\theta}^n\cdot\vec{Y}^n_{\pi^*,\vec{\theta}^n}\}\\
&\le \frac{1}{L}\sum_{n=mL}^{mL+L-1}\mathbb{E}\{X^n_{\pi^*,\vec{\beta}^{mL}}\}+\frac{1}{L}\sum_{n=mL}^{mL+L-1}\mathbb{E}\{\vec{\theta}^n\cdot\vec{Y}^n_{\pi^*,\vec{\beta}^{mL}_\epsilon}\}\\
&= \frac{1}{L}\sum_{n=mL}^{mL+L-1}\mathbb{E}\{X^n_{\pi^*,\vec{\beta}^{mL}}\}+\frac{1}{L}\sum_{n=mL}^{mL+L-1}\mathbb{E}\{\vec{\beta}^{mL}\cdot\vec{Y}^n_{\pi^*,\vec{\beta}^{mL}}\}\\
&\quad +\frac{1}{L}\sum_{n=mL}^{mL+L-1}\mathbb{E}\{(\vec{\theta}^n-\vec{\beta}^{mL})\cdot\vec{Y}^n_{\pi^*,\vec{\beta}^{mL}}\}\\
&= \frac{1}{L}\sum_{n=mL}^{mL+L-1}\mathbb{E}\{X^n_{\pi^*,\vec{\beta}^{mL}}\}+\frac{1}{L}\sum_{n=mL}^{mL+L-1}\mathbb{E}\{\vec{\beta}^{mL}\cdot\vec{Y}^n_{\pi^*,\vec{\beta}^{mL}}\}\\
&\quad +\frac{1}{L}\sum_{n=mL}^{mL+L-1}\mathbb{E}\{(\vec{\theta}^n-\vec{\theta}^{mL})\cdot\vec{Y}^n_{\pi^*,\vec{\beta}^{mL}}\}+\frac{1}{L}\sum_{n=mL}^{mL+L-1}\mathbb{E}\{(\vec{\theta}^{mL}-\vec{\beta}^{mL})\cdot\vec{Y}^n_{\pi^*,\vec{\beta}^{mL}}\}\\
&\overset{(c)}{\le} X^*+\frac{1}{L}\sum_{n=0}^{L-1}2B_1(B_2+1)\phi_n+\epsilon\frac{L+1}{2}B_1^2+\mathbb{E}\{\|\vec{\theta}^{mL}-\vec{\beta}^{mL}\|\}B_1, && \text{(D-7)}\\
&\overset{(d)}{\le} X^*+2B_1(B_2+1)\delta+\epsilon\frac{L+1}{2}B_1^2+\mathbb{E}\{\|\vec{\theta}^{mL}-\vec{\alpha}^{mL}\|\}B_1\\
&\quad +\mathbb{E}\{\|\vec{\alpha}^{mL}-\vec{\beta}^{mL}\|\}B_1 && \text{(D-8)}\\
&\overset{(e)}{\le} X^*+2B_1(B_2+1)\delta+\epsilon\frac{L+1}{2}B_1^2+\mathbb{E}\{\text{distance}(\vec{\theta}^{mL},\Theta)\}B_1+\delta B_1. && \text{(D-9)}
\end{aligned}
$$

The inequality (c) in (D-7) results from Lemma 2 and the facts $\|\vec{Y}^n_{\pi^*,\vec{\beta}^{mL}}\|\le B_1$, $\|\vec{Y}^n_{\pi^*,\vec{\theta}^n}\|\le B_1$ and $\|\vec{\theta}^n-\vec{\theta}^{mL}\|\le\epsilon(n-mL)B_1$. The triangular inequality and the assumption $\frac{1}{L}\sum_{n=0}^{L-1}\phi_n\le\delta$ give the inequality (d) in (D-8). By the definition of $\vec{\beta}^{mL}$, we know that $\vec{\beta}^{mL}\in N_\delta(\vec{\alpha}^{mL})$ and $\|\vec{\alpha}^{mL}-\vec{\beta}^{mL}\|\le\delta$. This fact, in turn, gives the inequality (e) in (D-9). ■

***Lemma** 4:* Let X^* be the optimal solution of Problem 1

$$
\begin{aligned}
&\limsup_{N} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}\{V(\vec{\theta}^n, \xi^n)\} \\
&= \limsup_{M} \frac{1}{M} \sum_{m=0}^{M} \frac{1}{L} \sum_{n=mL}^{mL+L-1} \mathbb{E}\{V(\vec{\theta}^n, \xi^n)\} \\
&\leq X^* + 2B_1(B_2+1)\delta + \epsilon \frac{L+1}{2} B_1^2 + \limsup_{M} \frac{1}{M} \sum_{m=0}^{M} \mathbb{E}\{\text{distance}(\vec{\theta}^{mL}, \Theta)\} B_1 \\
&\quad + \delta B_1.
\end{aligned}
$$

Proof: The result follows from Lemma 3. ∎