University of Pennsylvania
ScholarlyCommons

Departmental Papers (ESE)

Department of Electrical & Systems Engineering

11-17-2010

Ergodic Stochastic Optimization Algorithms for Wireless Communication and Networking

Alejandro Ribeiro
University of Pennsylvania, aribeiro@seas.upenn.edu

Suggested Citation:
Ribeiro, A. (2010). "Ergodc Stochastic Optimization Algorithms for Wireless Communication and Networking." *IEEE Transactions on Signal Processing* Vol. 58(12). pp. 6369 - 6386

©2010 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/ese_papers/578
For more information, please contact repository@pobox.upenn.edu.

Ergodic Stochastic Optimization Algorithms for Wireless Communication and Networking

Abstract
Ergodic stochastic optimization (ESO) algorithms are proposed to solve resource allocation problems that involve a random state and where optimality criteria are expressed in terms of long term averages. A policy that observes the state and decides on a resource allocation is proposed and shown to almost surely satisfy problem constraints and optimality criteria. Salient features of ESO algorithms are that they do not require access to the state's probability distribution, that they can handle nonconvex constraints in the resource allocation variables, and that convergence to optimal operating points holds almost surely. The proposed algorithm is applied to determine operating points of an orthogonal frequency division multiplexing broadcast channel that maximize a given rate utility.

Keywords
Adaptive signal processing, cross-layer design, gradient methods, OFDM, optimization, wireless communications, wireless networks

Disciplines
Electrical and Computer Engineering | Engineering

Comments
Suggested Citation:
Ribeiro, A. (2010). "Ergodc Stochastic Optimization Algorithms for Wireless Communication and Networking." *IEEE Transactions on Signal Processing* Vol. 58(12). pp. 6369 - 6386

©2010 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

This journal article is available at ScholarlyCommons: http://repository.upenn.edu/ese_papers/578

Ergodic Stochastic Optimization Algorithms for Wireless Communication and Networking

Alejandro Ribeiro, *Member, IEEE*

***Abstract*—Ergodic stochastic optimization (ESO) algorithms are proposed to solve resource allocation problems that involve a random state and where optimality criteria are expressed in terms of long term averages. A policy that observes the state and decides on a resource allocation is proposed and shown to almost surely satisfy problem constraints and optimality criteria. Salient features of ESO algorithms are that they do not require access to the state's probability distribution, that they can handle non-convex constraints in the resource allocation variables, and that convergence to optimal operating points holds almost surely. The proposed algorithm is applied to determine operating points of an orthogonal frequency division multiplexing broadcast channel that maximize a given rate utility.**

***Index Terms*—Adaptive signal processing, cross-layer design, gradient methods, OFDM, optimization, wireless communications, wireless networks.**

I. Introduction

THIS paper develops ergodic stochastic optimization (ESO) algorithms to solve problems that involve a time varying random state $\mathbf{h}(t)$ with probability distribution function (pdf) $m(\mathbf{h})$, a resource allocation function $\mathbf{p}(t)$ and an ergodic limit variable $\mathbf{x} := \lim_{t\to\infty}(1/t)\sum_{u=1}^{t}\mathbf{x}(u)$. The goal is to design an adaptive algorithm that observes $\mathbf{h}(t)$ to determine $\mathbf{p}(t)$ and $\mathbf{x}(t)$ without knowledge of the state's distribution $m(\mathbf{h})$ in order to satisfy given problem constraints and optimality criteria. Problem constraints restrict instantaneous values $\mathbf{p}(t)$ and $\mathbf{x}(t)$ as well as ergodic limits $\mathbf{x}$. Optimality criteria, however, depend only on the ergodic average $\mathbf{x}$. Constraints are convex in $\mathbf{x}$ and $\mathbf{x}(t)$ but need not be convex in $\mathbf{p}(t)$.

A. Motivation and Context

Problems with the characteristics described above are common in signal processing, most notably in the context of optimal wireless communications and networking. In this case $\mathbf{h}(t)$ denotes time varying fading channels, $\mathbf{p}(t)$ instantaneous power allocations and $\mathbf{x}(t)$ includes communication rates and other variables specific to the problem. If the operation time scale is much larger than the communication time scale, perceived quality of service is reasonably captured as a function of ergodic limits $\mathbf{x}$. Particular examples where this type of problem formulation might arise include optimization of orthogonal frequency division multiplexing (OFDM), [3], beamforming, [4], [5] and cognitive radio, [6], [7]. In OFDM, for example, terminals observe current channel realizations $\mathbf{h}(t)$ to determine power allocation across frequencies and users. Power allocation is restricted by instantaneous constraints, e.g., at most one terminal may transmit on each frequency, as well as ergodic average constraints, e.g., satisfy a given power budget. Optimality criteria, however, are most often formulated in terms of ergodic capacities. Similar interplays between instantaneous resource allocations and ergodic rates can be observed in beamforming and cognitive radio problems. Preceding comments also extend naturally to wireless networks where on top of power allocation and link capacities, routes and admission control variables are also part of the optimization space. Applications of ESO algorithms in this context are of particular interest because optimization is an important tool to analyze and design wireline and wireless communication networks, e.g., [8]–[16].

The proposed ESO algorithm uses *stochastic* subgradient descent on the dual function. Subgradient descent algorithms were developed for minimizing nondifferential convex functions and are commonly used to minimize dual functions which are always convex and often nondifferentiable. Subgradient descent with constant step sizes does not necessarily converge to the solution of the optimization problem considered, see e.g., [17], but this is not a significant drawback because it is possible to either use decreasing step sizes or, if this is not desirable, optimal solutions can be recovered from the time average of iterates. When the function to be minimized involves a random component it is possible to devise a stochastic counterpart, e.g., [18]–[20]. Similar convergence results, i.e., convergence with decreasing step sizes or through ergodic averages of iterates can be established using stochastic approximation tools [21]. Derivation and analysis of these algorithms is analogous to the transformation of a gradient descent algorithm to minimize mean squared errors into the least mean square (LMS) algorithm, e.g., [22, Ch. 4].

Implementation of dual subgradient descent yields, as a byproduct, a sequence of primal iterates. Do these primal iterates approximate optimal primal variables? Not always. When using deterministic or stochastic subgradient descent on the dual function this is true only when the problem Lagrangian is strictly concave with respect to primals. This condition is not satisfied if some variables appear only in linear constraints and linear terms of the optimization objective. Although this restriction might seem minor, non strictly concave Lagrangians

Manuscript received September 06, 2009; accepted June 09, 2010. Date of publication July 12, 2010; date of current version November 17, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Martin Schubert. This work is supported by ARO P-57920-NS, and NSF CAREER CCF-0952867.

The author is with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: aribeiro@seas.upenn.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2010.2057247

1053-587X/$26.00 © 2010 IEEE

do appear frequently—e.g., wireless networking problems are typically not strictly concave with respect to routing variables [23]. To overcome this limitation in deterministic subgradient descent, the use of ergodic averages of primal iterates has been proposed and shown to approximate optimal primal variables [24], [25]. Notice the remarkable coincidence of this property with the problem studied in this paper. The use of ergodic limits is necessary for convergence and the problems we purport to study include ergodic limits as part of their formulations.

While this much is known in a deterministic setting, convergence results for primal variables in stochastic subgradient descent are mostly restricted to convergence in mean for problems with strict convexity [20]. This paper shows that ergodic limits of primal iterates obtained from the implementation of a stochastic subgradient descent algorithm converge almost surely to the solution of the given optimization problem. With respect to existing work on stochastic subgradient descent the contributions of this work are: (i) we allow for non strictly concave Lagrangians, in which case convergence in the primal domain is not a simple consequence of convergence in the dual domain; (ii) we allow for nonconvex constraints in the resource allocation variables, in which case it is necessary to overcome the challenge that lack of duality gap cannot be ruled out *a priori*; and (iii) we prove almost sure convergence of the ergodic averages of primal iterates, which is stronger and more difficult to establish than convergence in expected value. These three properties are important in signal processing applications.

In the context of wireless networking, the work on stochastic network optimization of [14] and [15] has similar features to ESO in that it can operate without knowledge of the fading pdf and handle constraints that are not convex in $\mathbf{p}(t)$. The goal in stochastic network optimization is to determine operating points that maximize a network utility subject to queue stability constraints. To guarantee queue stability, it is necessary to find operating variables whose ergodic average is inside a capacity region that may be defined in terms of nonconvex constraints. To find such optimal operating points, stochastic network optimization relies on the backpressure algorithm, [26]. Performance analysis is undertaken through the use of Lyapunov drifts to establish that the average path across realizations of the state process $\mathbf{h}(t)$ comes close to a solution of the optimization problem. In ESO, the nonconvex constraints may appear as part of the definition of a capacity region, as in, e.g., [2]; or, they may appear in a different form, as in, e.g., [27]. When the problem formulations coincide, as in [2], ESO yields a different algorithm because backpressure does not play a role in the solution. Convergence properties of ESO and stochastic network optimization differ too because the analysis of ESO's behavior is based on supermartingales and ergodic theory, which allows for almost sure performance claims as opposed to claims on the average across ensembles.

B. Organization and Results

The paper starts in Section II by introducing the optimization problem whose solution determines optimal resource allocations and ergodic limits. Problem constraints are assumed convex in the ergodic limits but not necessarily so with respect to the resource allocation. The problem's objective is a concave function of ergodic limits only. The ESO algorithm and the main result of the paper, regarding convergence of resource allocation and ergodic limit sequences, is also introduced here in Theorem 1. It is claimed that: (i) resource allocation and ergodic limit sequences almost surely satisfy problem constraints in an ergodic sense; and (ii) the ergodic limit sequence is almost surely close to optimal.

Salient features of the ESO algorithm are that it does not require access to the state's pdf, that it can handle nonconvex constraints in the resource allocation variables, and that convergence to optimal operating points holds almost surely. To exemplify these characteristics we apply the ESO algorithm to determine the optimal operating point of an OFDM broadcast channel in Section II-A. A common access point (AP) administers a set of frequency bands and a given power budget to communicate with a set of nodes. Power and frequency allocations adapt to instantaneous channel realizations in order to maximize a given rate utility. The problem is complicated by the fact that at most one terminal can occupy a frequency band—yielding an integer constraint—and that the function that maps powers to communication rates may not be concave. Despite these complications, the ESO algorithm can be used to find an optimal operating point without requiring knowledge of the channels' pdf. The example serves as illustration of how ESO can be used to solve a nonconvex optimization problem with thousands of variables with reasonable computational cost. Numerical results of this example are presented in Section II-B. See also [2] for the application of ESO to general wireless networking problems and [27] for an application to random access channels.

The proofs of almost sure feasibility and optimality are respectively presented in Sections III and IV. They differ from typical stochastic approximation arguments in that while ESO descends on the dual domain we are interested in convergence in the primal domain. We start discussing convergence in the dual domain in Section III introducing a weak convergence result stated in Theorem 2. The proof of Theorem 2 is typical of stochastic approximation problems and is thereby relegated to the Appendix. Almost sure feasibility follows as a consequence of dual convergence and is presented in Section III. The proof of almost sure near optimality is presented in Section IV. We show that the expected value of the ergodic limit is a point with small optimality gap and use this result to prove almost sure convergence. This latter part of the proof is in Section IV-A. It is worth anticipating that the proof's argument is more involved than the prevalent ergodicity assumption. If ESO iterates $\mathbf{x}(t)$ were realizations of an ergodic process all ergodic limits would be equal, therefore equal to their expected value, and almost sure optimality would follow trivially from average optimality. However, the process with realizations $\mathbf{x}(t)$ might or might not be ergodic and it is therefore necessary to prove that even if ergodic limits $\mathbf{x}$ are different for different sequences $\mathbf{x}(t)$ they are all close to optimal. Concluding remarks are given in Section V.

Notation: Boldface denotes vectors $\mathbf{x}$ and vector valued functions $\mathbf{f}(\mathbf{x})$. Vector inequalities $\mathbf{x} \lesseqgtr \mathbf{y}$ are meant to imply $x_j \lesseqgtr y_j$ for all components x_j and y_j of $\mathbf{x}$ and $\mathbf{y}$, while $\mathbf{x} \lesseqgtr z$ with z a scalar implies $x_j \lesseqgtr z$ for all x_j. The pdf of a random variable $\mathbf{x}$ evaluated at $\mathbf{x} = \mathbf{x}_0$ is denoted as $m_{\mathbf{x}}(\mathbf{x}_0)$. When not leading to confusion we use the same symbol to denote values

and random variables in which case we drop the subindex to write $m_{\mathbf{x}}(\mathbf{x}) = m(\mathbf{x})$. The operator $\mathbb{E}_{m(\mathbf{x})}(\cdot)$ denotes expected value with respect to the pdf $m(\mathbf{x})$ of the random variable $\mathbf{x}$. If this pdf is fixed we write $\mathbb{E}_{m(\mathbf{x})}(\cdot) = \mathbb{E}_{\mathbf{x}}(\cdot)$. For a sequence with elements $\mathbf{x}(t)$ we use $\mathbf{x}(t_1 : t_2)$ when referring to all elements between times t_1 and t_2, $\mathbf{x}(t : \infty)$ for all elements at times greater than t, and $\mathbf{x}(\mathbb{N})$ when referring to the whole sequence. We use $\mathbf{X}(\mathbb{N})$ to denote a stochastic process with realizations $\mathbf{x}(\mathbb{N})$. The operator $[\cdot]^+$ denotes projection to the positive orthant.

II. Ergodic Stochastic Optimization Algorithm

Consider problems involving a random state $\mathbf{h} \in \mathcal{H}$ with probability distribution $m_{\mathbf{h}}(\mathbf{h}) = m(\mathbf{h})$, a resource allocation $\mathbf{p}(\mathbf{h})$ associated with state realization $\mathbf{h}$ and having pdf $m_{\mathbf{p}(\mathbf{h})}[\mathbf{p}(\mathbf{h})] = m(\mathbf{p}(\mathbf{h}))$, and an ergodic limit variable $\mathbf{x}$. The goal is to determine resource allocations and ergodic limits that are optimal in the sense of solving the optimization problem

$$\begin{aligned} P = &\max f_0(\mathbf{x}) \\ &\text{s.t.} \quad \mathbf{x} \leq \mathbb{E}_{\mathbf{h}}\Big[\mathbb{E}_{m(\mathbf{p}(\mathbf{h}))}\Big(\mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h})\Big)\Big], \quad \mathbf{f}_2(\mathbf{x}) \geq \mathbf{0} \\ &\qquad \mathbf{x} \in \mathcal{X},\ \{m(\mathbf{p}(\mathbf{h})) : \mathbf{p}(\mathbf{h}) \in \mathcal{P}(\mathbf{h})\}_{\mathbf{h}\in\mathcal{H}}. \end{aligned} \tag{1}$$

The optimization in (1) is with respect to ergodic limits $\mathbf{x}$ and probability distributions $m(\mathbf{p}(\mathbf{h}))$ for all $h \in \mathcal{H}$. We emphasize that the expected value is taken with respect to the pdfs $m(\mathbf{h})$ of the state $\mathbf{h}$ and $m(\mathbf{p}(\mathbf{h}))$ of the resource allocation $\mathbf{p}(\mathbf{h})$ and that while $m(\mathbf{h})$ is fixed the pdfs $m(\mathbf{p}(\mathbf{h}))$ are part of the optimization space. See Section II-A for a particular example.

The problem structure in (1) originates in optimal resource allocation problems with infinite time horizons allowing performance characterization through ergodic limits. System operation is affected by a random state with realizations $\mathbf{h}(t)$. In response to the observed state $\mathbf{h}(t)$, a resource allocation variable $\mathbf{p}(t) \in \mathcal{P}(\mathbf{h}(t))$ measuring how many units of a certain resource are allocated at time t is determined. Allocation of $\mathbf{p}(t)$ units of resource when the random state is $\mathbf{h}(t)$, results in the production of $\mathbf{f}_1(\mathbf{p}(t);\mathbf{h}(t))$ units of goods. In the same time slot t consumption is determined by $\mathbf{x}(t) \in \mathcal{X}$ variables. Consumption cannot exceed production, but if long time horizons are of interest, instead of imposing such restriction for every t it suffices to constraint the limits of the time averages, i.e.

$$\lim_{t\to\infty} \frac{1}{t}\sum_{u=1}^{t} \mathbf{x}(u) \leq \lim_{t\to\infty} \frac{1}{t}\sum_{u=1}^{t} \mathbf{f}_1(\mathbf{p}(u);\mathbf{h}(u)). \tag{2}$$

The first constraint in (1) follows upon defining the ergodic limit $\mathbf{x} := \lim_{t\to\infty}(1/t)\sum_{u=1}^{t}\mathbf{x}(u)$ and assuming ergodicity in the limit on the right-hand side (RHS) of (2). Notice that equal state realizations $\mathbf{h}(t_1) = \mathbf{h}(t_2)$ can be associated with different resource allocations $\mathbf{p}(t_1) \neq \mathbf{p}(t_2)$. That is why the expected value in (1) is taken with respect to the pdfs $m(\mathbf{h})$ and $m(\mathbf{p}(\mathbf{h}))$ and the optimization is with respect to probability distributions $m(\mathbf{p}(\mathbf{h})) \in \mathcal{P}(\mathbf{h})$ not values $\mathbf{p}(\mathbf{h}) \in \mathcal{P}(\mathbf{h})$. The constraint $\mathbf{f}_2(\mathbf{x}) \geq \mathbf{0}$ imposes further restrictions on the ergodic average $\mathbf{x}$.

Functions $f_0(\mathbf{x})$ and $\mathbf{f}_2(\mathbf{x})$ in (1) are concave with respect to their argument $\mathbf{x}$. The family of functions $\mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h})$ is parameterized by the random state $\mathbf{h}$ and, different from $f_0(\mathbf{x})$ and $\mathbf{f}_2(\mathbf{x})$, is not necessarily concave with respect to the resource allocation $\mathbf{p}(\mathbf{h})$. The sole requirement for the functions $\mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h})$ is that they be finite for finite argument, i.e., for every bounded vector of resources $\mathbf{p}(\mathbf{h}) < \infty$, the vector function $\mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h}) < \infty$ is also bounded. The set $\mathcal{X}$ to which the ergodic limits $\mathbf{x}$ are constrained is compact and convex, while the set $\mathcal{P}(\mathbf{h})$ constraining resource allocation *values* $\mathbf{p}(\mathbf{h})$ is compact but not necessarily convex. Notice that the set $\mathcal{P}(\mathbf{h})$ constrains the resource allocation $\mathbf{p}(\mathbf{h})$ on a per-state basis, i.e., there exists a set $\mathcal{P}(\mathbf{h})$ for each random state realization $\mathbf{h}$.

Because of the distributions $m(\mathbf{p}(\mathbf{h}))$, there is an infinite number of variables in the primal domain. Observe though, that there is a finite number of inequality constraints. Thus, the dual problem contains a finite number of variables hinting that the problem is likely more tractable in the dual space. Define then dual variables $\boldsymbol{\lambda}_1 \geq \mathbf{0}$ associated with the constraint $\mathbf{x} \leq \mathbb{E}_{\mathbf{h}}[\mathbb{E}_{m(\mathbf{p}(\mathbf{h}))}(\mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h}))]$ and $\boldsymbol{\lambda}_2 \geq \mathbf{0}$ associated with $\mathbf{f}_2(\mathbf{x}) \geq \mathbf{0}$. Using these definitions the Lagrangian for the optimization problem in (1) is written as

$$\begin{aligned} &\mathcal{L}[\boldsymbol{\lambda}, \mathbf{x}, m(\mathbf{p}(\mathbf{h}))] \\ &= f_0(\mathbf{x}) + \boldsymbol{\lambda}_1^T\Big[\mathbb{E}_{\mathbf{h}}\Big[\mathbb{E}_{m(\mathbf{p}(\mathbf{h}))}\Big(\mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h})\Big)\Big] - \mathbf{x}\Big] + \boldsymbol{\lambda}_2^T\mathbf{f}_2(\mathbf{x}) \\ &= f_0(\mathbf{x}) - \boldsymbol{\lambda}_1^T\mathbf{x} + \boldsymbol{\lambda}_2^T\mathbf{f}_2(\mathbf{x}) + \mathbb{E}_{\mathbf{h}}\Big[\mathbb{E}_{m(\mathbf{p}(\mathbf{h}))}\Big(\boldsymbol{\lambda}_1^T\mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h})\Big)\Big] \end{aligned} \tag{3}$$

where we defined the aggregate dual variable $\boldsymbol{\lambda} := [\boldsymbol{\lambda}_1^T, \boldsymbol{\lambda}_2^T]^T$ and reordered terms to obtain the second equality. The dual function is defined as the maximum of the Lagrangian over the set of feasible ergodic limits $\mathbf{x} \in \mathcal{X}$ and probability distributions $m(\mathbf{p}(\mathbf{h}))$ in the set of feasible powers $\mathbf{p}(\mathbf{h}) \in \mathcal{P}(\mathbf{h})$, i.e.

$$\begin{aligned} g(\boldsymbol{\lambda}) := \max \quad &\mathcal{L}[\boldsymbol{\lambda}, \mathbf{x}, m(\mathbf{p}(\mathbf{h}))] \\ \text{s.t.} \quad &\mathbf{x} \in \mathcal{X}, \quad \{m(\mathbf{p}(\mathbf{h})) : \mathbf{p}(\mathbf{h}) \in \mathcal{P}(\mathbf{h})\}_{\mathbf{h}\in\mathcal{H}} \end{aligned} \tag{4}$$

The dual problem is defined as the minimization of the dual function over all positive dual variables, i.e.,

$$D = \min_{\boldsymbol{\lambda}\geq\mathbf{0}} g(\boldsymbol{\lambda}). \tag{5}$$

Introduce now a discrete time index t and consider a state stochastic process $\mathbf{H}(\mathbb{N})$ with realizations $\mathbf{h}(\mathbb{N})$ having values $\mathbf{h}(t)$ identically and independently distributed (i.i.d.) according to $m(\mathbf{h})$. The ESO algorithm consists of iterative application of the following steps.

(S1) Primal Iteration: Given multipliers $\boldsymbol{\lambda}(t)$ find primal variables $\mathbf{x}(t) \in \mathcal{X}$ and $\mathbf{p}(t) \in \mathcal{P}(\mathbf{h}(t))$ such that

$$\mathbf{x}(t) = \mathbf{x}(\boldsymbol{\lambda}(t)) = \arg\max_{\mathbf{x}\in\mathcal{X}} f_0(\mathbf{x}) - \boldsymbol{\lambda}_1^T(t)\mathbf{x} + \boldsymbol{\lambda}_2^T(t)\mathbf{f}_2(\mathbf{x}) \tag{6}$$

$$\mathbf{p}(t) = \mathbf{p}(\mathbf{h}(t), \boldsymbol{\lambda}(t)) = \arg\max_{\mathbf{p}(\mathbf{h}(t))\in\mathcal{P}(\mathbf{h}(t))} \boldsymbol{\lambda}_1^T(t)\mathbf{f}_1(\mathbf{p}(\mathbf{h}(t));\mathbf{h}(t)). \tag{7}$$

(S2) Dual Stochastic Subgradients: Define the stochastic subgradient $\hat{\mathbf{s}}(t) = \hat{\mathbf{s}}(\mathbf{h}(t), \boldsymbol{\lambda}(t)) = [\hat{\mathbf{s}}_1^T(t), \hat{\mathbf{s}}_2^T(t)]^T$ of the dual function with components

$$\hat{\mathbf{s}}_1(t) := \mathbf{f}_1(\mathbf{p}(t);\mathbf{h}(t)) - \mathbf{x}(t), \quad \hat{\mathbf{s}}_2(t) := \mathbf{f}_2(\mathbf{x}(t)). \tag{8}$$

(S3) Dual Iteration: The algorithm's iteration is completed by an update in the dual domain moderated by a predetermined step size ϵ along the direction $-\hat{\mathbf{s}}(t)$

$$\begin{aligned}\boldsymbol{\lambda}(t+1) &= \left[\boldsymbol{\lambda}(t) - \epsilon\hat{\mathbf{s}}(t)\right]^+ \\ &= \begin{bmatrix}\boldsymbol{\lambda}_1(t) - \epsilon(\mathbf{f}_1(\mathbf{p}(t);\mathbf{h}(t)) - \mathbf{x}(t)) \\ \boldsymbol{\lambda}_2(t) - \epsilon\mathbf{f}_2(\mathbf{x})\end{bmatrix}^+. \quad (9)\end{aligned}$$

Resource allocation iterates $\mathbf{p}(t)$ are functions of dual iterates $\boldsymbol{\lambda}(t)$ and state realizations $\mathbf{h}(t)$. Since $\mathbf{h}(t)$ is random, so is $\mathbf{p}(t)$ implying that the sequence $\mathbf{p}(\mathbb{N})$ is a realization of a stochastic process $\mathbf{P}(\mathbb{N})$. In fact, all other sequences $\mathbf{x}(\mathbb{N})$, $\boldsymbol{\lambda}(\mathbb{N})$, and $\hat{\mathbf{s}}(\mathbb{N})$ are realizations of random processes respectively denoted as $\mathbf{X}(\mathbb{N})$, $\boldsymbol{\Lambda}(\mathbb{N})$, and $\hat{\mathbf{S}}(\mathbb{N})$, as can be seen by noting that $\mathbf{x}(t+1)$ depends on $\boldsymbol{\lambda}(t+1)$ which depends on $\hat{\mathbf{s}}(t)$ which depends on (random) $\mathbf{p}(t)$. This chain of relations further shows that with $\boldsymbol{\lambda}(t)$ given, $\mathbf{x}(t)$ is determined and the probability distributions of $\mathbf{p}(t)$, $\hat{\mathbf{s}}(t)$ and $\boldsymbol{\lambda}(t+1)$ depend only on the probability distribution of $\mathbf{h}(t)$. Consequently the process $\boldsymbol{\Lambda}(\mathbb{N})$ is Markov with time invariant transition probabilities, while processes $\mathbf{X}(\mathbb{N})$, $\mathbf{P}(\mathbb{N})$ and $\hat{\mathbf{S}}(\mathbb{N})$ are hidden Markov. Henceforth, the operator $\mathbb{E}[\cdot]$, without subindex, is used to denote expected value with respect to a probability space measuring, at least, $\mathbf{H}(\mathbb{N})$, $\mathbf{X}(\mathbb{N})$, $\mathbf{P}(\mathbb{N})$, $\boldsymbol{\Lambda}(\mathbb{N})$, and $\hat{\mathbf{S}}(\mathbb{N})$.

Solving an optimization problem like (1) entails finding the optimal value P and optimal arguments $\mathbf{x}^*$ and $m^*(\mathbf{p}(\mathbf{h}))$ such that the constraints in (1) are satisfied and $P = f_0(\mathbf{x}^*)$. Heeding the connection with ergodic constraints we adopt a different definition of solution. Our goal is not to find $\mathbf{x}^*$ and $m^*(\mathbf{p}(\mathbf{h}))$ but to show that sequences $\mathbf{x}(\mathbb{N})$ and $\mathbf{p}(\mathbb{N})$ generated by the ESO algorithm (S1)–(S3) satisfy (2) with the ergodic limit $\mathbf{x}$ of the $\mathbf{x}(\mathbb{N})$ sequence further satisfying $\mathbf{f}_2(\mathbf{x}) \geq \mathbf{0}$ and $P \approx f_0(\mathbf{x})$. Because the algorithm is stochastic, these results will be established in probability. Specifically, we will prove the following theorem.

Theorem 1: Consider the optimization problem in (1) and sequences $\mathbf{x}(\mathbb{N})$ and $\mathbf{p}(\mathbb{N})$ generated by the ESO algorithm defined by (6)–(9). Let $\hat{S}^2 \geq \mathbb{E}[\|\hat{\mathbf{s}}(t)\|^2 \,|\, \boldsymbol{\lambda}(t)]$ be a bound on the second moment of the norm of the stochastic subgradients $\hat{\mathbf{s}}(t)$ and assume that there exists strictly feasible $\mathbf{x}_0 \in \mathcal{X}$ and $m_0(\mathbf{p}(\mathbf{h}))$ such that $\mathbb{E}_{\mathbf{h}}[\mathbb{E}_{m_0(\mathbf{p}(\mathbf{h}))}(\mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h}))] - \mathbf{x}_0 > C$ and $\mathbf{f}_2(\mathbf{x}_0) > C$ for some strictly positive constant $C > 0$. Then

i) **Almost sure feasibility**. Sequences $\mathbf{x}(\mathbb{N})$ and $\mathbf{p}(\mathbb{N})$ are feasible with probability 1, i.e.

$$\lim_{t\to\infty}\frac{1}{t}\sum_{u=1}^{t}\mathbf{x}(u) \leq \lim_{t\to\infty}\frac{1}{t}\sum_{u=1}^{t}\mathbf{f}_1(\mathbf{p}(u);\mathbf{h}(u)) \quad \text{a.s.}, \tag{10}$$

$$\mathbf{f}_2\left[\lim_{t\to\infty}\frac{1}{t}\sum_{u=1}^{t}\mathbf{x}(u)\right] \geq \mathbf{0} \quad \text{a.s.} \tag{11}$$

ii) **Almost sure near optimality**. The ergodic average of $f_0(\mathbf{x}(u))$ almost surely converges to a value with optimality gap smaller than $\epsilon\hat{S}^2/2$, i.e.

$$P - f_0\left[\lim_{t\to\infty}\frac{1}{t}\sum_{u=1}^{t}\mathbf{x}(u)\right] \leq \frac{\epsilon\hat{S}^2}{2} \quad \text{a.s.} \tag{12}$$

It is important to elaborate on what (10)–(12) imply in terms of finding a solution of (1). The ergodic limit $\mathbf{x} := (1/t)\sum_{u=1}^{t}\mathbf{x}(u)$ satisfies the constraints in (1) and the objective function evaluated at $\mathbf{x}$ is within $\epsilon\hat{S}^2/2$ of optimal. Since $\mathcal{X}$ and $\mathcal{P}(\mathbf{h})$ are compact sets it follows that the bound $\hat{S}^2$ is finite. Therefore, reducing ϵ it is possible to make $\mathbf{f}_0(\mathbf{x})$ arbitrarily close to P and as a consequence $\mathbf{x}$ arbitrarily close to some optimal argument $\mathbf{x}^*$. The optimal resource allocation distribution $m^*(\mathbf{p}(\mathbf{h}))$, however, is not computed by the ESO algorithm. Rather, (10) implies that, asymptotically, the ESO algorithm is drawing resource allocation realizations $\mathbf{p}(t)$ from a resource allocation distribution that is close to the optimal $m^*(\mathbf{p}(\mathbf{h}))$. This is not a drawback in practice because realizations $\mathbf{p}(t)$ are sufficient for implementation. In that sense, (6)–(9) yield an optimal resource allocation policy, i.e., allocate $\mathbf{p}(t)$ units at time t, that supports optimal consumption $\mathbf{x}$ in an ergodic sense.

Recall that $\boldsymbol{\Lambda}(\mathbb{N})$ and primal iterates processes $\mathbf{X}(\mathbb{N})$ and $\mathbf{P}(\mathbb{N})$ are respectively Markov and hidden Markov processes with time invariant transition probabilities. As a consequence, all the ergodic limits in (10)–(12) exist, a fact that is exploited in subsequent proofs. Note also, that the limits in (10)–(12) might be different for different state sequences $\mathbf{h}(\mathbb{N})$. The claims in Theorem 1 are on the probability distributions of these limits. The proof of Theorem 1 and corresponding discussions are presented in Sections III and IV after the following remarks and the discussion of an example in Sections II-A and B.

Remark 1: The problem formulation in (1) makes what seems an arbitrary distinction between constraints $\mathbf{f}_2(\mathbf{x}) \geq \mathbf{0}$ and $\mathbf{x} \in \mathcal{X}$. While one is expressed as a function inequality and the other as a set inclusion both are convex contraints in the ergodic limit $\mathbf{x}$. Despite this similarity they are intended to model different constraint modalities. The constraint $\mathbf{f}_2(\mathbf{x}) \geq \mathbf{0}$ is incorporated into the Lagrangian in (3) and becomes a maximization objective in the primal ESO iteration in (6). As a consequence, it is satisfied in an ergodic sense. Ergodic limits of $\mathbf{x}(\mathbb{N})$ sequences satisfy $\mathbf{f}_2(\mathbf{x}) \geq \mathbf{0}$ but individual variables $\mathbf{x}(t)$ might or might not satisfy $\mathbf{f}_2(\mathbf{x}(t)) \geq \mathbf{0}$. The constraint $\mathbf{x} \in \mathcal{X}$ is not incorporated into $\mathcal{L}[\boldsymbol{\lambda}, \mathbf{x}, m(\mathbf{p}(\mathbf{h}))]$ and is an implicit constraint in the primal ESO iteration in (6). It is thus satisfied for all times, i.e., $\mathbf{x}(t) \in \mathcal{X}$, for all t. This is an important distinction in applications, e.g., transmitted power in wireless communications must comply with both, ergodic and instantaneous power constraints.

Remark 2: As mentioned in the introduction particular cases of the optimization problem in (1) that appear in wireless communication and networking can be solved through stochastic network optimization algorithms [14], [15]. These algorithms use backpressure to handle nonconvex constraints that appear in the definition of a capacity region. In ESO the nonconvex constraints may appear as part of the definition of a capacity region, as in, e.g., [2]; or they may appear in a different form, as in, e.g., [27]. When the problem formulations coincide, as in [2], the proposed algorithm is slightly different because backpressure does not play a role in the solution here. Furthermore, ESO and stochastic network optimization deal differently with nonconvex constraints $\mathbf{x} \leq \mathbb{E}_{\mathbf{h}}[\mathbb{E}_{m(\mathbf{p}(\mathbf{h}))}(\mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h}))]$ and $\mathbf{p}(\mathbf{h}) \in \mathcal{P}(\mathbf{h})$. In stochastic network optimization, the nonconvex constraints are eliminated through the introduction of a capacity region

and are left implicit in the definition of the dual function. In ESO, the nonconvex constraints are incorporated in the definition of the dual function. The treatment of the convex constraints $\mathbf{f}_2(\mathbf{x}) \geq \mathbf{0}$ is also different in [14] and [15] necessitating the introduction of auxiliary variables.

Convergence properties of ESO and stochastic network optimization differ too. Results in stochastic network optimization establish negative Lyapunov drifts to prove that the expected value of ergodic limits satisfy problem constraints and that the expected value of the time average of iterates yields a utility that is close to optimal [14], [15]. Here we use supermartingales and ergodic theory, which allows for almost sure performance claims. The feasibility results for ESO iterates claim that constraints are almost surely and exactly satisfied by ergodic limits [cf. (10) and (11)], while ergodic averages almost surely converge to a near optimal point [cf. (12)]. Almost sure convergence is an important guarantee in practice.

A. Example: Wireless Broadcast Channel

As an application of the ESO algorithm consider a wireless broadcast channel using OFDM. A common AP administers a group of frequency tones $\mathcal{F}$ and an average power budget P_0. The goal is to design an algorithm that allocates power and frequency to maximize a given ergodic rate utility metric. Specifically, consider J nodes $\{N_i\}_{i=1}^J$ served by the given AP. At time t the AP observes fading channels $h_i^f(t)$ for all frequencies $f \in \mathcal{F}$ and nodes N_i. Depending on the values of these fading channels it decides on frequency allocation $a_i^f(t) \in \{0,1\}$ and power allocation $p_i^f(t) \in [0, p_{\rm m}]$. Variable $a_i^f(t) = 1$ if and only if frequency f is allocated to node N_i at time t. If $a_i^f(t) = 1$, the power allocated for such communication is $p_i^f(t)$. Since no more than one communication can utilize a given frequency, at most one $a_i^f(t)$ can be different from zero for given f and t. To capture this constraint define the vector $\mathbf{a}^f(t) = [a_{1f}(t), \ldots, a_{Jf}(t)]^T$ and require $\mathbf{a}^f(t) \in \mathcal{A}$ with the set $\mathcal{A}$ defined as

$$\mathcal{A} := \{\mathbf{a} = [a_1, \ldots a_J]^T : a_j \in \{0,1\}, \quad \mathbf{a}^T\mathbf{1} \leq 1\}. \tag{13}$$

With channel realization $h_i^f(t)$ and power allocation $p_i^f(t)$, $C(p_i^f(t); h_i^f(t))$ units of information are delivered to N_i. The function $C(p_i^f(t); h_i^f(t))$ mapping channels and powers to transmission rates depends on the type of modulation and codes used. As an example, consider adaptive modulation and coding (AMC) that relies on a set of L communication modes. The l-th mode supports a rate α_l and is used when the signal to noise ratio (SNR) at the receiver end is between β_{l-1} and β_l. Normalizing channels $h_i^f(t)$ so that noise power is $\sigma_i^f(t) = 1$, the communication rate function $C(p_i^f(t); h_i^f(t))$ for AMC can be written as

$$C\left(p_i^f(t); h_i^f(t)\right) = \sum_{l=1}^{L} \alpha_l \mathbb{I}\left(\beta_l \leq h_i^f(t) p_i^f(t) < \beta_{l+1}\right). \tag{14}$$

The amount of information $C(p_i^f(t); h_i^f(t))$ in (14) is delivered to N_i only if $a_i^f(t) = 1$. The information delivered to N_i at time t over all frequencies f is therefore given by $\sum_{f\in\mathcal{F}} a_i^f(t) C(p_i^f(t); h_i^f(t))$. The AMC's rate function in (14) is an example, the following discussion holds true for this or any other function taking finite values for finite argument.

While $\sum_{f\in\mathcal{F}} a_i^f(t) C(p_i^f(t); h_i^f(t))$ units of information are *delivered* by the AP, $c_i(t) \in [0, c_{\rm m}]$ units of information are *accepted* for delivery and queued to await transmission. To guarantee delivery of packets it suffices to ensure stability of information queues which in turn can be guaranteed by requiring

$$\begin{aligned} c_i &:= \lim_{t\to\infty} \frac{1}{t} \sum_{u=1}^{t} c_i(u) \\ &\leq \lim_{t\to\infty} \frac{1}{t} \sum_{u=1}^{t} \left[\sum_{f\in\mathcal{F}} a_i^f(u) C\left(p_i^f(u); h_i^f(u)\right) \right] \end{aligned} \tag{15}$$

where we have defined the ergodic limit $c_i := \lim_{t\to\infty}(1/t)\sum_{u=1}^{t} c_i(u)$. Similarly, the amount of power consumed at time t is the sum of powers used on all frequencies for communication with all terminals, i.e., $\sum_{f\in\mathcal{F}} a_i^f(t) p_i^f(t)$. This cannot exceed allocated power P_0 thus yielding the constraint

$$P_0 \geq \lim_{t\to\infty} \frac{1}{t} \sum_{u=1}^{t} \left[\sum_{f\in\mathcal{F}} a_i^f(u) p_i^f(u) \right]. \tag{16}$$

Assuming the time scale of communication to be much smaller than the time scale of operation, perceived quality of service is determined by the ergodic limit c_i. Assigning value $U_i(c_i)$ to ergodic rate c_i the goal is to determine sequences of frequency allocations $a_i^f(\mathbb{N})$, powers $p_i^f(\mathbb{N})$ and rates $c_i(\mathbb{N})$ such that: (i) the ergodic limits of $c_i(\mathbb{N})$ sequences maximize a sum utility $\sum_{i=1}^{J} U_i(c_i)$; (ii) constraints in (15) and (16) are satisfied; and (iii) instantaneous frequency allocations are feasible, i.e., $\mathbf{a}^f(t) \in \mathcal{A}$ for $\mathcal{A}$ in (13). The ESO algorithm solves problems of this form.

Let $\mathbf{h}$ aggregate channel variables h_i^f for all i and f and define frequency and power allocations $a_i^f(\mathbf{h})$ and $p_i^f(\mathbf{h})$. Also, let $\mathbf{a}(\mathbf{h})$ aggregate all $a_i^f(\mathbf{h})$ and $\mathbf{p}(\mathbf{h})$ represent all $p_i^f(\mathbf{h})$. Define then the optimization problem

$$\begin{aligned} \max \quad & \sum_{i=1}^{J} U_i(c_i) \\ \text{s.t.} \quad & c_i \leq \mathbb{E}_{\mathbf{h}} \left[\mathbb{E}_{m(\mathbf{a}(\mathbf{h}),\mathbf{p}(\mathbf{h}))} \left(\sum_{f\in\mathcal{F}} a_i^f(\mathbf{h}) C\left(p_i^f(h); h_i^f\right) \right) \right] \\ & P_0 \geq \mathbb{E}_{\mathbf{h}} \left[\mathbb{E}_{m(\mathbf{a}(\mathbf{h}),\mathbf{p}(\mathbf{h}))} \left(\sum_{i=1}^{J} \sum_{f\in\mathcal{F}} a_i^f(\mathbf{h}) p_i^f(\mathbf{h}) \right) \right] \\ & \mathbf{a}^f(\mathbf{h}) \in \mathcal{A},\ 0 \leq p_i^f(\mathbf{h}) \leq p_{\rm m},\ 0 \leq c_i \leq c_{\rm m}. \end{aligned} \tag{17}$$

The optimization in (17) is with respect to ergodic limits $c_i \in [0, c_{\rm m}]$ and probability distributions $m(\mathbf{a}(\mathbf{h}), \mathbf{p}(\mathbf{h}))$ restricted to $\mathbf{a}^f(t) \in \mathcal{A}$ and $0 \leq p_i^f(\mathbf{h}) \leq p_{\rm m}$. This optimization problem is of the form in (1). Introducing multipliers λ_i associated with the capacity constraint and μ associated with the power

constraint, the primal iteration of the ESO algorithm is [cf. (6) and (7)]

$$c_i(t) = \underset{0 \le c_i \le c_{\mathrm{m}}}{\arg\max}\, U_i(c_i) - \lambda_i(t) c_i \tag{18}$$

$$a_i^f(t), p_i^f(t) = \underset{\substack{0 \le p_i^f(\mathbf{h}(t)) \le p_{\mathrm{m}} \\ \mathbf{a}^f(\mathbf{h}(t)) \in \mathcal{A}}}{\arg\max}\; a_i^f(\mathbf{h}(t)) \Big[\lambda_i(t) C(p_i^f(\mathbf{h}(t)); h_i^f(t)) - \mu(t) p_i^f(\mathbf{h}(t))\Big]. \tag{19}$$

Since for fixed f at most one variable $a_i^f(t)$ can be 1 in (19), determination of $a_i^f(t)$ and $p_i^f(t)$ can be separated further. Powers $p_i^f(t)$ that maximize $\lambda_i(t)C(p_i^f(\mathbf{h}(t)); h_i^f(t)) - \mu(t)p_i^f(\mathbf{h}(t))$ are first computed for all i and frequency f is then allocated to the maximum of $\lambda_i(t)C(p_i^f(t); h_i^f(t)) - \mu(t)p_i^f(t)$. Formally, compute

$$p_i^f(t) = \underset{0 \le p_i^f(\mathbf{h}(t)) \le p_{\mathrm{m}}}{\arg\max}\; \lambda_i(t) C\Big(p_i^f(\mathbf{h}(t)); h_i^f(t)\Big) - \mu(t) p_i^f(\mathbf{h}(t)) \tag{20}$$

$$i^f(t) = \underset{i}{\arg\max}\; \lambda_i(t) C\left(p_i^f(t); h_i^f(t)\right) - \mu(t) p_i^f(t) \tag{21}$$

and set $a_{i^f(t)}^f(t) = 1$ and $a_i^f(t) = 0$ for all other $i \neq i^f(t)$. Notice that while the maximization in (20) involves the non concave discontinuous function $C(p_i^f(\mathbf{h}(t)); h_i^f(t))$, it is nonetheless simple to solve as the maximand depends on a single variable $p_i^f(\mathbf{h}(t))$. The ESO algorithm for optimal resource allocation in broadcast channels is completed with an iteration in the dual domain [cf. (8) and (9)]

$$\lambda_i(t+1) = \left[\lambda_i(t) - \epsilon\left[\sum_{f \in \mathcal{F}} a_i^f(t) C\left(p_i^f(t); h_i^f(t)\right) - c_i(t)\right]\right]^+$$

$$\mu(t+1) = \left[\mu(t) - \epsilon\left[P_0 - \sum_{i=1}^{J}\sum_{f \in \mathcal{F}} a_i^f(t) p_i^f(t)\right]\right]^+. \tag{22}$$

As per Theorem 1 iterative application of (18) and (20)–(22) yields sequences $c_i(\mathbb{N})$, $a_i^f(\mathbb{N})$ and $p_i^f(\mathbb{N})$ such that: (i) the sum utility for the ergodic limits of $c_i(\mathbb{N})$ is almost surely within a small constant of optimal; (ii) the ergodic constraints in (15) and (16) are almost surely satisfied; and (iii) instantaneous frequency values of $a_i^f(t)$ are feasible, i.e., $\mathbf{a}^f(t) \in \mathcal{A}$. The stated goal is then satisfied with probability 1. This result is true despite the presence of the nonconvex integer constraint $\mathbf{a}^f(t) \in \mathcal{A}$, the nonconcave function $C(p_i^f(t); h_i^f(t))$, lack of access to the channel's probability distribution and the infinite dimensionality of the optimization problem.

B. Numerical Results

The ESO algorithm for optimal resource allocation in broadcast channels defined by (18) and (20)–(22) is simulated for a system with $J = 16$ nodes using 3 frequency tones for communication. Three AMC modes corresponding to capacities 1, 2, and 3 bits/s/Hz are used with transitions at SINR 1, 3, and 7. Fading channels are generated as i.i.d. Rayleigh with average powers 1 for the first four nodes, i.e., $j = 1, \ldots, 4$, and 2, 3, and 4 for subsequent groups of 4 nodes. Noise power is $N_i^f = 1$ for all frequency bands and average power available is $P_0 = 3$. Rate of packet acceptance is constrained to be $0 \le c_i(t) \le 2$ bits/s/Hz. The optimality criteria is proportional fair scheduling, i.e., $U_i(c_i) = \log(c_i)$ for all i. Steps size is $\epsilon = 0.1$.

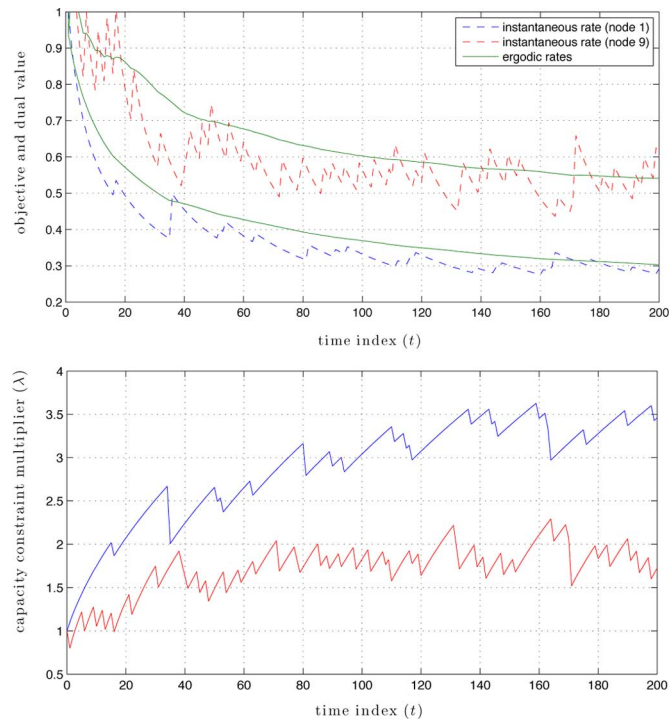

Fig. 1. Evolution of dual variables $\lambda_i(t)$ and corresponding capacities $c_i(t)$ for representative nodes N_1 with average channels $\mathbb{E}[h_{1f}(t)] = 1$ and N_9 with $\mathbb{E}[h_{9f}(t)] = 3$. Ergodic averages $\bar{c}_i(t) = (1/t)\sum_{u=1}^{t} c_i(u)$ also shown. Multipliers $\lambda_i(t)$ and capacities $c_i(t)$ do not converge, but ergodic rates $\bar{c}_i(t)$ do. Multiplier $\lambda_1(t)$ associated with node N_1 is larger than multiplier $\lambda_9(t)$ of node N_9. This improves fairness of resource allocation by increasing the chances of allocating N_1 even when channels $h_{1f}(t)$ are smaller than $h_{9f}(t)$—recall that channel $h_{9f}(t)$ is stronger on average.

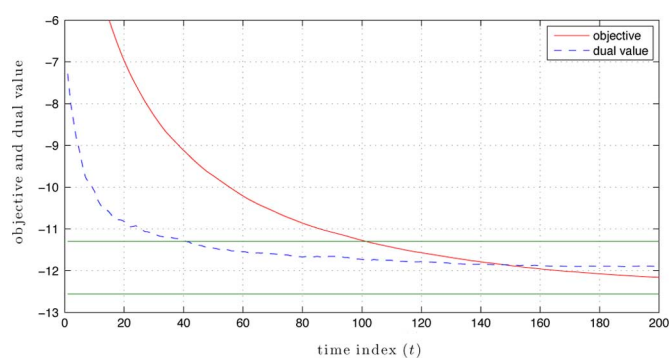

Fig. 2. Objective value $\sum_{i=1}^{J} U_i(\bar{c}_i(t))$ and dual function's value $g(t) := g(\boldsymbol{\lambda}(t), \mu(t))$. Lines showing optimal objective and 90% of optimal objective are also shown. The objective value decreases towards the maximum objective. This is not a contradiction, because variables $\bar{c}_i(t)$ are infeasible but approach feasibility as t grows. Dual function's values upper bound maximum utility. Eventually, the objective value becomes smaller than the dual value.

Fig. 1 shows evolution of dual variables $\lambda_i(t)$ and corresponding capacities $c_i(t)$ for representative nodes N_1 with average channels $\mathbb{E}[h_{1f}(t)] = 1$ and N_9 with $\mathbb{E}[h_{9f}(t)] = 3$. The ergodic average $\bar{C}_i(t) := (1/t)\sum_{u=1}^{t} c_i(u)$ is also shown. Neither multipliers $\lambda_i(t)$ nor capacities $c_i(t)$ converge, but ergodic rates $\bar{c}_i(t)$ do converge. Multiplier $\lambda_1(t)$ associated with node N_1 is larger than multiplier $\lambda_9(t)$ of node N_9. This improves fairness of resource allocation by increasing

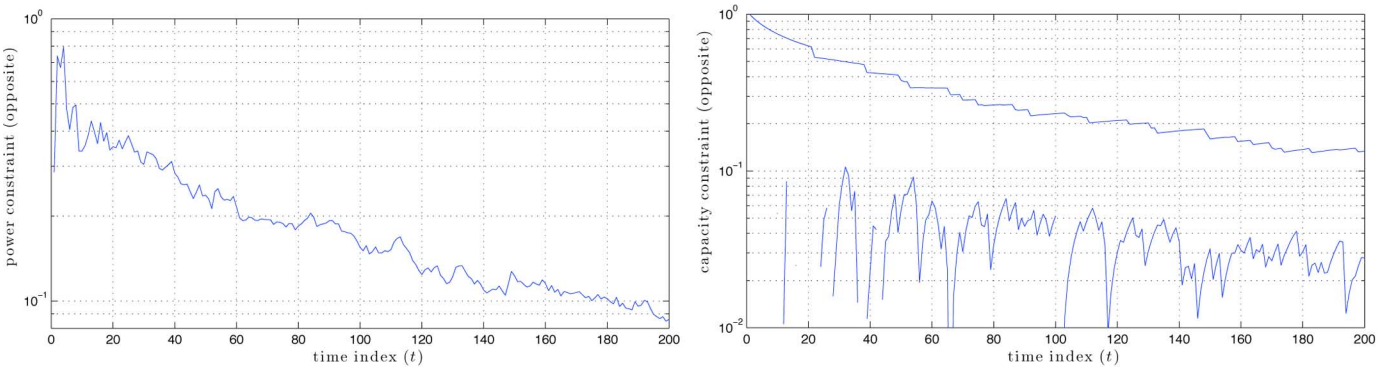

Fig. 3. Power and capacity constraints. Feasibility as time grows is corroborated for the power constraints in (15) (left) and capacity constraint in (16) (right). For the capacity constraint we show the maximum and minimum value of constraint violation.

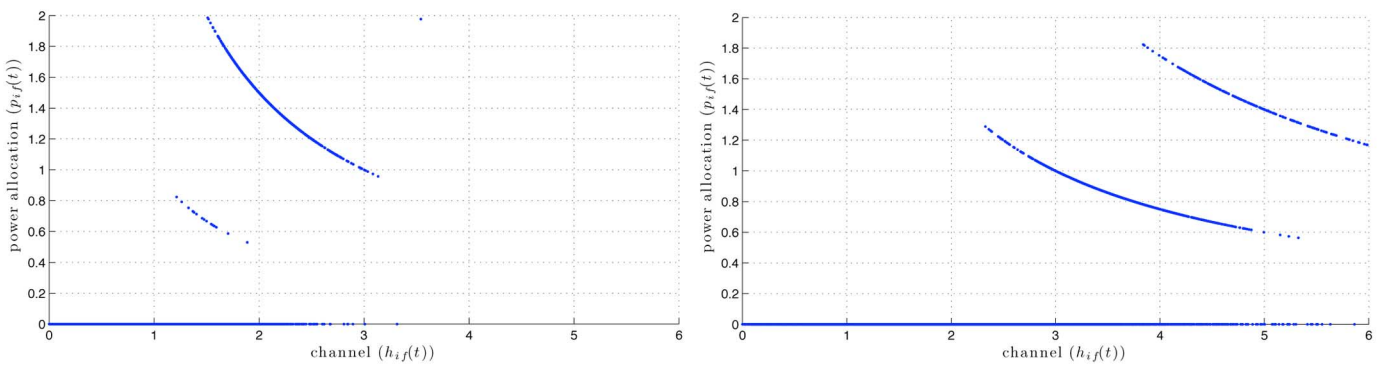

Fig. 4. Power allocations. Power allocated as a function of channel realization is shown for channels with average power $\mathbb{E}[h_i^f(t)] = 1$ (left) and $\mathbb{E}[h_i^f(t)] = 2$ (right). The resulting power allocation is opportunistic in that power is allocated only when channel realizations are above average.

the chances of allocating N_1 even when channels $h_{1f}(t)$ are smaller than $h_{9f}(t)$—recall that channel $h_{9f}(t)$ is stronger on average. Convergence of the algorithm is ratified by Figs. 2 and 3. Fig. 2 shows evolution of the objective $\sum_{i=1}^{J} U_i(\bar{c}_i(t))$ and the dual function value $g(t) := g(\boldsymbol{\lambda}(t), \mu(t))$. Notice that the objective value is decreasing towards the maximum objective. This is not a contradiction, because variables $\bar{c}_i(t)$ are infeasible but approach feasibility as t grows. The dual function's value is an upper bound on the maximum utility and it can be observed to approach the objective as t grows. Eventually, the objective value becomes smaller than the dual value as expected. Fig. 3 corroborates satisfaction of the constraints in (15) and (16). The amount by which the power constraint (16) is violated is shown in the left. In the right we show the corresponding figure for the capacity constraint in (15). Since there are J of these constraints we show the minimum and maximum violation. All constraints are satisfied as t grows. Resulting power allocations appear in Fig. 4 for a channel with average power $\mathbb{E}[h_i^f(t)] = 1$ and for a channel with $\mathbb{E}[h_i^f(t)] = 2$. Power allocation is opportunistic in that power is allocated only when channel realizations are above average.

III. Convergence of Dual Variables and Almost Sure Feasibility

While the goal of the ESO algorithm is to find feasible optimal primal sequences $\mathbf{p}(\mathbb{N})$ and $\mathbf{x}(\mathbb{N})$, it does so by descending on the convex dual function $g(\boldsymbol{\lambda})$. It is therefore expected that to prove feasibility and optimality of $\mathbf{p}(\mathbb{N})$ and $\mathbf{x}(\mathbb{N})$ it is necessary to characterize convergence properties of the sequence of dual variables $\boldsymbol{\lambda}(\mathbb{N})$. Thus, let us start proving that stochastic subgradients $\hat{\mathbf{s}}(t)$ are, as their name indicates, average descent directions of the convex dual function. This property is formally stated in the following proposition.

Proposition 1: Let $\boldsymbol{\lambda}(t)$ be given and define $g(t) := g(\boldsymbol{\lambda}(t))$. Then, the conditional expected value $\mathbb{E}[\hat{\mathbf{s}}(t) \,|\, \boldsymbol{\lambda}(t)]$ of the stochastic subgradient $\hat{\mathbf{s}}(t)$ in (8) is a subgradient of the dual function. Specifically, for arbitrary $\boldsymbol{\lambda}$

$$\mathbb{E}\left[\hat{\mathbf{s}}^T(t) \,\middle|\, \boldsymbol{\lambda}(t)\right](\boldsymbol{\lambda}(t) - \boldsymbol{\lambda}) \geq g(t) - g(\boldsymbol{\lambda}). \tag{23}$$

Proof: Start noting that the maximization of the Lagrangian in (4) can be simplified. In $\mathcal{L}[\boldsymbol{\lambda}, \mathbf{x}, m(\mathbf{p}(\mathbf{h}))]$ the terms $\mathbf{x}$ and $\mathbf{p}(\mathbf{h})$ appear in separate summands [cf. (3)] implying that the maximizations with respect to $\mathbf{x} \in \mathcal{X}$ and $\{m(\mathbf{p}(\mathbf{h})) : \mathbf{p}(\mathbf{h}) \in \mathcal{P}(\mathbf{h})\}_{\mathbf{h}\in\mathcal{H}}$ in (4) can be carried out as separate maximizations with respect to $\mathbf{x}$ and $m(\mathbf{p}(\mathbf{h}))$ of the corresponding summands. Also, since the expected value operator is linear, to maximize $\mathbb{E}_{\mathbf{h}}[\mathbb{E}_{m(\mathbf{p}(\mathbf{h}))}(\boldsymbol{\lambda}_1^T \mathbf{f}_1(\mathbf{p}(\mathbf{h}); \mathbf{h}))]$ with respect to the family of power allocation distributions $\{m(\mathbf{p}(\mathbf{h}))\}_{\mathbf{h}\in\mathcal{H}}$ it suffices to maximize $\mathbb{E}_{m(\mathbf{p}(\mathbf{h}))}(\boldsymbol{\lambda}_1^T \mathbf{f}_1(\mathbf{p}(\mathbf{h}); \mathbf{h}))$ with respect to each individual distribution $m(\mathbf{p}(\mathbf{h}))$. Thus

$$g(\boldsymbol{\lambda}) = \max_{x\in\mathcal{X}} \left(f_0(\mathbf{x}) - \boldsymbol{\lambda}_1^T \mathbf{x} + \boldsymbol{\lambda}_2^T \mathbf{f}_2(\mathbf{x}) \right) + \mathbb{E}_{\mathbf{h}} \left[\max_{m(\mathbf{p}(\mathbf{h})):\mathbf{p}(\mathbf{h})\in\mathcal{P}(\mathbf{h})} \mathbb{E}_{m(\mathbf{p}(\mathbf{h}))} \left(\boldsymbol{\lambda}_1^T \mathbf{f}_1(\mathbf{p}(\mathbf{h}); \mathbf{h}) \right) \right]. \tag{24}$$

Further note that for any probability measure $m(\mathbf{p}(\mathbf{h}))$ in the set $\mathbf{p}(\mathbf{h}) \in \mathcal{P}(\mathbf{h})$ it holds

$$\mathbb{E}_{m(\mathbf{p}(\mathbf{h}))}\left(\boldsymbol{\lambda}_1^T \mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h})\right) \leq \max_{\mathbf{p}(\mathbf{h})\in\mathcal{P}(\mathbf{h})} \boldsymbol{\lambda}_1^T \mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h}) \quad (25)$$

with equality achieved by assigning probability 1 to any of the arguments $\arg\max_{\mathbf{p}(\mathbf{h})\in\mathcal{P}(\mathbf{h})} \boldsymbol{\lambda}_1^T \mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h})$. This property allows us to write the dual function as

$$g(\boldsymbol{\lambda}) = \max_{\mathbf{x}\in\mathcal{X}} \left(f_0(\mathbf{x}) - \boldsymbol{\lambda}_1^T\mathbf{x} + \boldsymbol{\lambda}_2^T \mathbf{f}_2(\mathbf{x})\right) + \mathbb{E}_{\mathbf{h}}\left[\max_{\mathbf{p}(\mathbf{h})\in\mathcal{P}(\mathbf{h})} \boldsymbol{\lambda}_1^T \mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h})\right]. \quad (26)$$

Consequently, when performing the maximization in (24) it is not necessary to consider all pdfs $m(\mathbf{p}(\mathbf{h}))$ sufficing to maximize $\boldsymbol{\lambda}_1^T \mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h})$ over resource allocations $\mathbf{p}(\mathbf{h}) \in \mathcal{P}(\mathbf{h})$ as is done in (26).

Consider now the value $g(t) := g(\boldsymbol{\lambda}(t))$ of the dual function at time t and compare the expression for the dual function in (26) with the definition of the primal iterates in (6)–(7). Making the substitutions $\boldsymbol{\lambda}_1 \leftrightarrow \boldsymbol{\lambda}_1(t)$, $\boldsymbol{\lambda}_2 \leftrightarrow \boldsymbol{\lambda}_2(t)$ and $\mathbf{h} \leftrightarrow \mathbf{h}(t)$ maximands in (6)–(7) and (26) coincide. We can thus write after reordering terms

$$g(t) = f_0(\mathbf{x}(t)) + \boldsymbol{\lambda}_1^T(t)\left[\mathbb{E}_{\mathbf{h}(t)}\left[\mathbf{f}_1(\mathbf{p}(\mathbf{h}(t),\boldsymbol{\lambda}(t));\mathbf{h}(t))\right] - \mathbf{x}(t)\right] + \boldsymbol{\lambda}_2^T(t)\mathbf{f}_2(\mathbf{x}(t)). \quad (27)$$

Let $\boldsymbol{\lambda}(t)$ be given and consider the conditional expected value of the stochastic subgradient $\mathbb{E}[\hat{\mathbf{s}}(t)\,|\,\boldsymbol{\lambda}(t)]$. Focus first on the second component $\hat{\mathbf{s}}_2(t)$. With $\boldsymbol{\lambda}(t)$ given, primal variables $\mathbf{x}(t)$ are unequivocally determined [cf. (6)] thus implying [cf. (8)]

$$\mathbb{E}[\hat{\mathbf{s}}_2(t)\,|\,\boldsymbol{\lambda}(t)] = \mathbf{f}_2(\mathbf{x}(t)). \quad (28)$$

The first component $\hat{\mathbf{s}}_1(t)$, however, depends on $\mathbf{p}(t)$ and $\mathbf{x}(t)$ [cf. (8)] with $\mathbf{p}(t)$ further dependent on $\boldsymbol{\lambda}(t)$ and $\mathbf{h}(t)$ [cf. (7)]. Thus

$$\begin{aligned}\mathbb{E}[\hat{\mathbf{s}}_1(t)\,|\,\boldsymbol{\lambda}(t)] &= \mathbb{E}\left[\mathbf{f}_1(\mathbf{p}(\mathbf{h}(t),\boldsymbol{\lambda}(t));\mathbf{h}(t))\,|\,\boldsymbol{\lambda}(t)\right] - \mathbb{E}[\mathbf{x}(t)\,|\,\boldsymbol{\lambda}(t)] \\ &= \mathbb{E}_{\mathbf{h}(t)}\left[\mathbf{f}_1(\mathbf{p}(\mathbf{h}(t),\boldsymbol{\lambda}(t));\mathbf{h}(t))\right] - \mathbf{x}(t) \end{aligned} \quad (29)$$

where the second equality follows from $\mathbb{E}[\mathbf{x}(t)\,|\,\boldsymbol{\lambda}(t)] = \mathbf{x}(t)$ and the fact that with $\boldsymbol{\lambda}(t)$ given, expectation with respect to the processes' probability distribution is equivalent to expectation with respect to $\mathbf{h}(t)$.

Substitute now the expressions in (28) and (29) into (27) to rewrite the dual function value $g(t)$ as

$$\begin{aligned} g(t) &= f_0(\mathbf{x}(t)) + \mathbb{E}\left[\hat{\mathbf{s}}_1^T(t)\,|\,\boldsymbol{\lambda}(t)\right]\boldsymbol{\lambda}_1(t) + \mathbb{E}\left[\hat{\mathbf{s}}_2^T(t)\,|\,\boldsymbol{\lambda}(t)\right]\boldsymbol{\lambda}_2(t) \\ &= f_0(\mathbf{x}(t)) + \mathbb{E}\left[\hat{\mathbf{s}}^T(t)\,|\,\boldsymbol{\lambda}(t)\right]\boldsymbol{\lambda}(t). \end{aligned} \quad (30)$$

To finalize the proof consider the dual function value $g(\boldsymbol{\lambda})$ for arbitrary $\boldsymbol{\lambda}$ that we can bound as

$$\begin{aligned} g(\boldsymbol{\lambda}) &\geq f_0(\mathbf{x}(t)) \\ &\quad + \boldsymbol{\lambda}_1^T\left[\mathbb{E}_{\mathbf{h}(t)}\left(\mathbf{f}_1(\mathbf{p}(\mathbf{h}(t),\boldsymbol{\lambda}(t));\mathbf{h}(t))\right) - \mathbf{x}(t)\right] + \boldsymbol{\lambda}_2^T\mathbf{f}_2(\mathbf{x}(t)) \\ &= f_0(\mathbf{x}(t)) + \mathbb{E}\left[\hat{\mathbf{s}}^T(t)\,|\,\boldsymbol{\lambda}(t)\right]\boldsymbol{\lambda}. \end{aligned} \quad (31)$$

The inequality in (31) is true because evaluating the RHS of (26) for $\mathbf{p}(\mathbf{h}(t)) = \mathbf{p}(\mathbf{h}(t),\boldsymbol{\lambda}(t))$ and $\mathbf{x} = \mathbf{x}(t)$ lower bounds the maxima that yields the dual function's value $g(\boldsymbol{\lambda})$. Subtracting (31) from (30) yields (23). ■

Proposition 1 establishes that the expected value of the stochastic subgradient is a descent direction of the dual function. To see this substitute $\boldsymbol{\lambda} = \boldsymbol{\lambda}^*$ into (23). For $\boldsymbol{\lambda} = \boldsymbol{\lambda}^*$ we have $g(\boldsymbol{\lambda}^*) = D$ and since D is the minimum of the dual function it follows that

$$\mathbb{E}\left[\hat{\mathbf{s}}^T(t)\,|\,\boldsymbol{\lambda}(t)\right](\boldsymbol{\lambda}(t) - \boldsymbol{\lambda}^*) \geq g(t) - D \geq 0. \quad (32)$$

If the inner product between vectors $\boldsymbol{\lambda}^* - \boldsymbol{\lambda}(t)$ and $-\hat{\mathbf{s}}(t)$ is positive on average, the angle between them tends to be smaller than $\pi/2$. Thus, standing on $\boldsymbol{\lambda}(t)$, $-\hat{\mathbf{s}}(t)$ points, on average, towards $\boldsymbol{\lambda}^*$. Because the dual function is always convex we expect that a descent algorithm constructed by replacing $\mathbb{E}[\hat{\mathbf{s}}(t)\,|\,\boldsymbol{\lambda}(t)]$ for $\hat{\mathbf{s}}(t)$ in (9), would eventually approach an optimal dual variable $\boldsymbol{\lambda}^*$ as is indeed not difficult to prove ([17], Ch. 2). Since the stochastic subgradient $\hat{\mathbf{s}}(t)$ varies randomly around its mean $\mathbb{E}[\hat{\mathbf{s}}(t)\,|\,\boldsymbol{\lambda}(t)]$ it is reasonable to expect iterates $\boldsymbol{\lambda}(t)$ of (9) to also come close to $\boldsymbol{\lambda}^*$. This argument can be formalized to establish the following result.

Theorem 2: Consider the ESO algorithm as defined by (6)–(9) with the same hypotheses and definitions of Theorem 1. Assume the dual variable $\boldsymbol{\lambda}(T_0)$ at given time T_0 is given and define the best value of the dual function at time t as $g_{\text{best}}[t\,|\,\boldsymbol{\lambda}(T_0)] := \min_{u\in[T_0,t]} g(u)$. Such best value almost surely converges to within $(\epsilon\hat{S}^2)/2$ of D, i.e.

$$\lim_{t\to\infty} g_{\text{best}}\left[t\,|\,\boldsymbol{\lambda}(T_0)\right] - D \leq \frac{\epsilon\hat{S}^2}{2} \quad \text{a.s.} \quad (33)$$

Proof: See the Appendix. ■

According to Theorem 2 it holds that for almost all processes and arbitrary $\delta > 0$, $g(t) - D$ falls below $\epsilon\hat{S}^2/2 + \delta$ at least once as t grows. Furthermore, this happens infinite many times because at any time T_0 there exists a future time t at which $g(t) - D \leq \epsilon\hat{S}^2/2 + \delta$. Note that (33) is true if and only if the limit infimum $\liminf_{t\to\infty} g(t) - D \leq \epsilon\hat{S}^2/2$ for almost all sequences. The explicit formulation in (33) is more convenient to establish forthcoming proofs in Section III-A.

The result in (33) is a weak characterization of the dual iterates. All that is claimed by Theorem 2 is that the dual iterates $\boldsymbol{\lambda}(t)$ eventually visit a neighborhood of the optimal set $\boldsymbol{\lambda}^*$. In between these visits no claim is made about $\boldsymbol{\lambda}(t)$. While a stronger characterization is possible, for the purpose of studying convergence properties of primal sequences $\mathbf{p}(\mathbb{N})$ and $\mathbf{x}(\mathbb{N})$, Theorem 2 is sufficient. We pursue these proofs after the following remark.

Remark 3: As pointed out in the introduction the ESO algorithm (S1)–(S3) is related to subgradient descent on the dual function. The primal iteration of subgradient descent consists of finding arguments that maximize the Lagrangian $\mathcal{L}[\boldsymbol{\lambda}(t), \mathbf{x}, m(\mathbf{p}(\mathbf{h}))]$ [cf. step (S1)]. The constraints are then evaluated at these maximizing arguments to compute a dual function subgradient [cf. step (S2)] and the dual variables descend opposite the subgradient direction [cf. step (S3)]. To appreciate similarities and differences let us develop dual subgradient descent equations for the problem in (1). To compute a subgradient for the dual function we find primal arguments $\mathbf{x}(t)$ and $\{m(\mathbf{p}(\mathbf{h}), t)\}_{\mathbf{h}\in\mathcal{H}}$ that maximize the Lagrangian, i.e.,

$$\begin{aligned} &\mathbf{x}(t), \{m(\mathbf{p}(\mathbf{h}), t)\}_{\mathbf{h}\in\mathcal{H}} := \\ &\qquad \arg\max \quad \mathcal{L}[\boldsymbol{\lambda}(t), \mathbf{x}, m(\mathbf{p}(\mathbf{h}))] \\ &\qquad \text{s.t.} \qquad \mathbf{x}\in\mathcal{X},\ \{m(\mathbf{p}(\mathbf{h})) : \mathbf{p}(\mathbf{h}) \in \mathcal{P}(\mathbf{h})\}_{\mathbf{h}\in\mathcal{H}}. \end{aligned} \tag{34}$$

As noted in (24) in the proof of Proposition 1 the Lagrangian $\mathcal{L}[\boldsymbol{\lambda}(t), \mathbf{x}, m(\mathbf{p}(\mathbf{h}))]$ exhibits a separable structure. Variables $\mathbf{x}$ and $\mathbf{p}(\mathbf{h})$ appear in different summands and the maximization of $\mathbb{E}_{m(\mathbf{p}(\mathbf{h}))}(\boldsymbol{\lambda}_1^T \mathbf{f}_1(\mathbf{p}(\mathbf{h}); \mathbf{h}))$ can be reduced to separate maximizations with respect to each individual distribution $m(\mathbf{p}(\mathbf{h}))$. Therefore, the maximizers in (34) can be computed separately as

$$\mathbf{x}(t) = \arg\max_{\mathbf{x}\in\mathcal{X}} f_0(\mathbf{x}) - \boldsymbol{\lambda}_1^T(t)\mathbf{x} + \boldsymbol{\lambda}_2^T(t)\mathbf{f}_2(\mathbf{x}), \tag{35}$$

$$m(\mathbf{p}(\mathbf{h}), t) = \arg\max_{m(\mathbf{p}(\mathbf{h})):\mathbf{p}(\mathbf{h})\in\mathcal{P}(\mathbf{h})} \mathbb{E}_{m(\mathbf{p}(\mathbf{h}))}\left(\boldsymbol{\lambda}_1^T(t)\mathbf{f}_1(\mathbf{p}(\mathbf{h}); \mathbf{h})\right). \tag{36}$$

A subgradient $\mathbf{s}(t) = \mathbf{s}(\boldsymbol{\lambda}(t)) = [\mathbf{s}_1^T(t), \mathbf{s}_2^T(t)]^T$ of the dual function can now obtained by evaluating the constraints at the maximizing arguments $\mathbf{x}(t)$ and $\{m(\mathbf{p}(\mathbf{h}), t)\}_{\mathbf{h}\in\mathcal{H}}$, ([17, Ch. 2]). Components $\mathbf{s}_1^T(t)$ and $\mathbf{s}_2^T(t)]^T$ are therefore given by

$$\begin{aligned} \mathbf{s}_1(t) &:= \mathbb{E}_{\mathbf{h}}\left[\mathbb{E}_{m(\mathbf{p}(\mathbf{h}),t)}\left(\boldsymbol{\lambda}_1^T \mathbf{f}_1(\mathbf{p}(\mathbf{h}); \mathbf{h})\right)\right] - \mathbf{x}(t) \\ \mathbf{s}_2(t) &:= \mathbf{f}_2(\mathbf{x}(t)). \end{aligned} \tag{37}$$

The dual update has the same functional form of (9) with $\mathbf{s}(t)$ replacing $\hat{\mathbf{s}}(t)$. The subgradient descent algorithm for (1) consists of iterative application of the primal iteration (35)–(37) and a dual iteration having the same form as (9) with $\mathbf{s}(t)$ replacing $\hat{\mathbf{s}}(t)$.

Stochastic subgradients $\hat{\mathbf{s}}(t)$ are easier to compute. To determine $\mathbf{s}_1(t)$ it is necessary to solve the maximization in (36) for a large number of states $\mathbf{h}$ in order to obtain a good approximation of the expected value in (37). To compute $\hat{\mathbf{s}}_1(t)$ only one such maximization, for $\mathbf{h} = \mathbf{h}(t)$ is required. Further note that the maximization in (36) is with respect to probability distributions $m(\mathbf{p}(\mathbf{h}))$ while the maximization in (7) is with respect to values $\mathbf{p}(\mathbf{h}(t))$. Also, while to implement subgradient descent the state probability distribution $m(\mathbf{h})$ is needed to compute the expected value in (29), the stochastic version requires access only to current state realizations $\mathbf{h}(t)$.

A. Feasibility

To prove that sequences $\mathbf{p}(\mathbb{N})$ and $\mathbf{x}(\mathbb{N})$ are almost surely feasible in the sense of (10) and (11), the idea is to show that if ESO sequences $\mathbf{p}(\mathbb{N})$ and $\mathbf{x}(\mathbb{N})$ generated by iterative application of (6)–(9) are *not* feasible in that sense, the corresponding dual iterates $\boldsymbol{\lambda}(t)$ grow unbounded. Assuming that the set of optimal multipliers is bounded, i.e., $\|\boldsymbol{\lambda}^*\| < \infty$ for all optimal $\boldsymbol{\lambda}^*$, this would imply that the distance between $\boldsymbol{\lambda}(t)$ and any optimal $\boldsymbol{\lambda}^*$ becomes arbitrarily large. Because the dual function is convex the latter would in turn imply that $g(t) = g[\boldsymbol{\lambda}(t)]$ grows unbounded. But if this happens, then (33) of Theorem 2 is not true and since the set of sequences for which (33) is true has probability 1, realizations $\mathbf{h}(\mathbb{N})$ with infeasible $\mathbf{p}(\mathbb{N})$ and $\mathbf{x}(\mathbb{N})$ have zero probability.

A technical detail in this chain of argument is the need to ensure boundedness of the optimal set of dual variables, i.e., $\|\boldsymbol{\lambda}^*\| < \infty$. A sufficient condition for the latter is the existence of a strictly feasible set of primal variables $\mathbf{x}_0$ and $m_0(\mathbf{p}(\mathbf{h}))$ such that for some strictly positive constant $C > 0$, $\mathbb{E}_{\mathbf{h}}[\mathbb{E}_{m_0(\mathbf{p}(\mathbf{h}))}(\mathbf{f}_1(\mathbf{p}(\mathbf{h}); \mathbf{h}))] - \mathbf{x}_0 > C$ and $\mathbf{f}_2(\mathbf{x}_0) > C$. Recall that this is a hypothesis of Theorem 1. If such strictly feasible variables exist they can be used to bound the norm of $\boldsymbol{\lambda}$ by the value of the dual function $g(\boldsymbol{\lambda})$. To do this, note that being $g(\boldsymbol{\lambda}) := \max_{\mathbf{x}, m(\mathbf{p}(\mathbf{h}))} \mathcal{L}[\boldsymbol{\lambda}, \mathbf{x}, m(\mathbf{p}(\mathbf{h}))]$ the maximum value of the Lagrangian over primal variables, it has to be $g(\boldsymbol{\lambda}) \geq \mathcal{L}[\boldsymbol{\lambda}, \mathbf{x}_0, m_0(\mathbf{p}(\mathbf{h}))]$. Writing $\mathcal{L}[\boldsymbol{\lambda}, \mathbf{x}_0, m_0(\mathbf{p}(\mathbf{h}))]$ in explicit form and using the bounds $\mathbb{E}_{\mathbf{h}}[\mathbb{E}_{m_0(\mathbf{p}(\mathbf{h}))}(\mathbf{f}_1(\mathbf{p}(\mathbf{h}); \mathbf{h}))] - \mathbf{x}_0 > C$ and $\mathbf{f}_2(\mathbf{x}_0) > C$ yields

$$\begin{aligned} &g(\boldsymbol{\lambda}) \\ &\geq f_0(\mathbf{x}_0) + \boldsymbol{\lambda}_1^T\left[\mathbb{E}_{\mathbf{h}}\left[\mathbb{E}_{m_0(\mathbf{p}(\mathbf{h}))}\left(\mathbf{f}_1(\mathbf{p}(\mathbf{h}); \mathbf{h})\right)\right] - \mathbf{x}_0\right] + \boldsymbol{\lambda}_2^T \mathbf{f}_2(\mathbf{x}_0) \\ &\geq f_0(\mathbf{x}_0) + C\boldsymbol{\lambda}^T\mathbf{1}. \end{aligned} \tag{38}$$

Because $\boldsymbol{\lambda} \geq \mathbf{0}$ it follows that each component of $\boldsymbol{\lambda}$ is smaller than the sum $\boldsymbol{\lambda}^T\mathbf{1}$. Using this fact and reordering terms in (38) yields

$$\boldsymbol{\lambda} \leq \frac{1}{C}\Big[g(\boldsymbol{\lambda}) - f_0(\mathbf{x}_0)\Big], \tag{39}$$

where the inequality denotes a componentwise bound on the vector $\boldsymbol{\lambda}$ by the scalar value $(1/C)[g(\boldsymbol{\lambda}) - f_0(\mathbf{x}_0)]$. The inequality in (39) tells us that if strictly feasible $\mathbf{x}_0$ and $m_0(\mathbf{p}(\mathbf{h}))$ exist, as required by hypothesis, then a finite value of the dual function $g(\boldsymbol{\lambda})$ implies a finite argument $\boldsymbol{\lambda}$. In particular, optimal arguments $\boldsymbol{\lambda}^*$ are bounded because D is. Relying on this observation we can formalize the former argument to establish almost sure feasibility of $\mathbf{p}(\mathbb{N})$ and $\mathbf{x}(\mathbb{N})$ in the sense of (10) and (11).

Proof (Result (i) of Theorem 1): Before proceeding with the proof, recall that Theorem 2 assures the almost sure existence of a time $t \geq T_0$, for arbitrarily selected T_0 at which $g(t) - D \leq \epsilon\hat{S}^2/2$. Combining this result with the dual value bound on dual variables shown in (39), we can equivalently guarantee that for almost all sequences and arbitrarily large time T_0 there exists a time t at which

$$\boldsymbol{\lambda}(t) \leq \frac{1}{C}\left[\frac{\epsilon\hat{S}^2}{2} + D - f_0(\mathbf{x}_0)\right] \quad \text{for some } t \geq T_0. \tag{40}$$

Start the proof proper by noting that (10) would be true if

$$\lim_{t\to\infty} \frac{1}{t}\sum_{u=1}^{t}[\mathbf{f}_1(\mathbf{p}(\mathbf{h}(u)); \mathbf{h}(u)) - \mathbf{x}(u)] \geq \mathbf{0} \quad \text{a.s.} \tag{41}$$

For (11) notice that it holds

$$\mathbf{f}_2\left[\lim_{t\to\infty}\frac{1}{t}\sum_{u=1}^{t}\mathbf{x}(u)\right]=\lim_{t\to\infty}\mathbf{f}_2\left[\frac{1}{t}\sum_{u=1}^{t}\mathbf{x}(u)\right] \geq \lim_{t\to\infty}\tfrac{1}{t}\sum_{u=1}^{t}\mathbf{f}_2[\mathbf{x}(u)] \tag{42}$$

where the equality follows from continuity of $\mathbf{f}_2(\mathbf{x})$—continuity being implied by concavity—and the inequality from its concavity. From the latter inequality in (42) we conclude that (11) would follow if

$$\lim_{t\to\infty}\frac{1}{t}\sum_{u=1}^{t}\mathbf{f}_2[\mathbf{x}(u)]\geq\mathbf{0}\quad\text{a.s.} \tag{43}$$

Comparing the inequalities (41) and (43) with the definition of the stochastic subgradient $\hat{\mathbf{s}}(t)$ in (8) we see that the summands in (41) and (43) coincide with the correspondent components $\hat{\mathbf{s}}_1(u)$ and $\hat{\mathbf{s}}_2(u)$. Therefore, it is just a matter of definition that (41) and (43) can be more compactly written as

$$\lim_{t\to\infty}\frac{1}{t}\sum_{u=1}^{t}\hat{\mathbf{s}}(u)\geq\mathbf{0}\quad\text{a.s.} \tag{44}$$

Almost sure feasibility follows if we establish that (44) is true. To do this, assume that it is not. Hypothesize then the existence of a nonzero probability set of sequences for which (44) is not true. Existence of this set implies existence of a sufficiently small positive constant $\delta>0$ such that with nonzero probability we have

$$\lim_{t\to\infty}\frac{1}{t}\sum_{u=1}^{t}\hat{\mathbf{s}}(u)\leq-2\delta<0. \tag{45}$$

For each of the sequences for which (45) is true there exists a time T_1, possibly different for each sequence, such that for all $t\geq T_1$

$$\frac{1}{t}\sum_{u=1}^{t}\hat{\mathbf{s}}(u)\leq-\delta<0. \tag{46}$$

Consider now the dual iteration in (9) for a generic time index t. Since projection to the positive orthant increases the value of $\boldsymbol{\lambda}(t+1)$, it holds

$$\boldsymbol{\lambda}(t+1)\geq\boldsymbol{\lambda}(t)-\epsilon\hat{\mathbf{s}}(t)\geq\boldsymbol{\lambda}(1)-\epsilon\sum_{u=1}^{t}\hat{\mathbf{s}}(u) \tag{47}$$

where the second inequality follows from a recursive argument. Combining (46) and (47), we conclude that for any sequence for which (45) holds, we have a time T_1, such that for all $t\geq T_1$

$$\boldsymbol{\lambda}(t+1)\geq\boldsymbol{\lambda}(1)+(\epsilon\delta)t\geq(\epsilon\delta)t \tag{48}$$

where the second inequality is true because we dropped the nonnegative $\boldsymbol{\lambda}(1)$. Select now a time index $T_2\geq T_1$, that also satisfies

$$T_2>\frac{1}{C(\epsilon\delta)}\left[\frac{\epsilon\hat{S}^2}{2}+D-f_0(\mathbf{x}_0)\right]. \tag{49}$$

For all times $t\geq T_2$, we then have [cf. (48) and (49)]

$$\boldsymbol{\lambda}(t+1)>\frac{1}{C}\left[\frac{\epsilon\hat{S}^2}{2}+D-f_0(\mathbf{x}_0)\right]\quad\text{for all}\,t\geq T_2. \tag{50}$$

The hypothesized existence of a nonzero probability set of sequences with the property in (45) implies that a nonzero probability subset satisfies (50). Elements of this set also satisfy (40) almost surely because of Theorem 2. Selecting $T_0=T_2$, in (40)—which we can do because T_0 may be arbitrarily selected—it follows that sequences in this set comply with the inequality in (50) for *all* $t\geq T_2$ and the inequality in (40) for *some* $t\geq T_2=T_0$. This is absurd. Therefore the hypothesized existence of a nonzero probability set of sequences with the property in (45) is also absurd. The opposite, i.e., that almost all sequences satisfy (44), is true. The proof follows because (44) implies (10) and (11) as already shown. ■

The just proved feasibility claim stated in Result (i) of Theorem 1 assures that sequences $\mathbf{p}(\mathbb{N})$ and $\mathbf{x}(\mathbb{N})$ satisfy problem constraints with probability 1 [cf. (10) and (11)]. This is a stronger claim when compared with the optimality results for primal and dual variables that establish a typically small but not null performance gap [cf. (12) of Theorem 1 and (33) of Theorem 2]. It is also worth remarking that this is true independently of the step size ϵ. While we think of ϵ as a small number, and it is indeed desirable to select small ϵ, this is not necessary to ensure feasibility of $\mathbf{p}(\mathbb{N})$ and $\mathbf{x}(\mathbb{N})$. The strength of Result (i) of Theorem 1 is important from a practical standpoint. A small optimality gap is acceptable in general, but a small violation of problem constraints results in a set of variables incompatible with the physical constraints of the system.

IV. Optimality

We now turn attention to establish optimality of the ergodic limit $\mathbf{x}=\lim_{t\to\infty}(1/t)\sum_{u=1}^{t}\mathbf{x}(u)$. As per Result (ii) of Theorem 1 we want to show that with probability 1 $f_0(\mathbf{x})$ is within $\epsilon\hat{S}^2/2$ of optimal. Since $f_0(\mathbf{x})$ is a concave function it is possible to write

$$P-f_0\left[\frac{1}{t}\sum_{u=1}^{t}\mathbf{x}(u)\right]\leq P-\frac{1}{t}\sum_{u=1}^{t}f_0(\mathbf{x}(u)):=\bar{f}_0(t) \tag{51}$$

where we have defined the ergodic average of the optimality gaps $\bar{f}_0(t):=P-(1/t)\sum_{u=1}^{t}f_0(\mathbf{x}(u))$. The bound in (51) can be read as stating that the optimality gap of the ergodic average $P-f_0[(1/t)\sum_{u=1}^{t}\mathbf{x}(u)]$ is smaller than the ergodic average of the optimality gaps $\bar{f}_0(t)$.

Comparing (12) with (51) it follows that to prove near-optimality of the ergodic limit $\mathbf{x}$ as per (12) it suffices to bound $\bar{f}_0(t)$ by a small positive constant for all sufficiently large t.

The latter will be established through an ergodic argument, i.e., by translating properties of $\lim_{t\to\infty}\mathbb{E}[\bar{f}_0(t)]$ into properties of $\lim_{t\to\infty}\bar{f}_0(t)$. A prerequisite for that is to establish that the expected value $\mathbb{E}[\bar{f}_0(t)]$ of the ergodic average $\bar{f}_0(t)$ does become small as t grows. This is the subject of the following theorem.

Theorem 3 (Average Near-Optimality): Consider sequences $\mathbf{p}(\mathbb{N})$ and $\mathbf{x}(\mathbb{N})$ generated by the ESO algorithm defined by (6)–(9) and assume the same hypotheses and definitions of Theorem 1. Then,

$$\lim_{t\to\infty} \mathbb{E}[\bar{f}_0(t)\,|\,\boldsymbol{\lambda}(T_0)] = P - \lim_{t\to\infty} \mathbb{E}\left[\frac{1}{t}\sum_{u=1}^{t} f_0(\mathbf{x}(u)) \,\middle|\, \boldsymbol{\lambda}(T_0)\right] \leq \frac{\epsilon \hat{S}^2}{2}. \quad (52)$$

I.e., the expected value of the ergodic averages of $f_0(\mathbf{x}(t))$, conditioned on arbitrary dual variable $\boldsymbol{\lambda}(T_0)$ at arbitrary time T_0, converges to a point with optimality gap smaller than $\epsilon\hat{S}^2/2$.

Proof: In (6), $\mathbf{x}(t)$ is defined as a maximizing argument. Therefore, evaluating the maximand at $\mathbf{x} = \mathbf{x}(t)$ upper bounds the maximand values for any other $\mathbf{x} \in \mathcal{X}$ yielding

$$f_0(\mathbf{x}(t)) - \boldsymbol{\lambda}_1^T(t)\mathbf{x}(t) + \boldsymbol{\lambda}_2^T(t)\mathbf{f}_2(\mathbf{x}(t)) \geq f_0(\mathbf{x}) - \boldsymbol{\lambda}_1^T(t)\mathbf{x} + \boldsymbol{\lambda}_2^T(t)\mathbf{f}_2(\mathbf{x}). \quad (53)$$

Likewise, resource allocation $\mathbf{p}(t)$ is a maximizing argument for the expression in (7). It then follows that for arbitrary $\mathbf{p}(t) \in \mathcal{P}(\mathbf{h}(t))$ it holds

$$\boldsymbol{\lambda}_1^T(t)\mathbf{f}_1(\mathbf{p}(t);\mathbf{h}(t)) \geq \boldsymbol{\lambda}_1^T(t)\mathbf{f}_1(\mathbf{p}(\mathbf{h}(t));\mathbf{h}(t)). \quad (54)$$

The above being true for all $\mathbf{p}(\mathbf{h}(t)) \in \mathcal{P}(\mathbf{h}(t))$ implies that it is still true after integrating over arbitrary probability distribution $m(\mathbf{p}(\mathbf{h}))t$. Then, for any pdf $m(\mathbf{p}(\mathbf{h}))t$ we have

$$\boldsymbol{\lambda}_1^T(t)\mathbf{f}_1(\mathbf{p}(t);\mathbf{h}(t)) \geq \boldsymbol{\lambda}_1^T(t)\mathbb{E}_{m(\mathbf{p}(\mathbf{h}(t)))}(\mathbf{f}_1(\mathbf{p}(\mathbf{h}(t));\mathbf{h}(t))). \quad (55)$$

Summing up the inequalities in (53) and (55) we obtain after reordering terms

$$f_0(\mathbf{x}(t)) + \boldsymbol{\lambda}_1^T(t)[\mathbf{f}_1(\mathbf{p}(t);\mathbf{h}(t)) - \mathbf{x}(t)] + \boldsymbol{\lambda}_2^T(t)\mathbf{f}_2(\mathbf{x}(t)) \geq f_0(\mathbf{x}) + \boldsymbol{\lambda}_1^T(t)\Big[\mathbb{E}_{m(\mathbf{p}(\mathbf{h}(t)))}\Big(\mathbf{f}_1(\mathbf{p}(\mathbf{h}(t));\mathbf{h}(t))\Big) - \mathbf{x}\Big] + \boldsymbol{\lambda}_2^T(t)\mathbf{f}_2(\mathbf{x}). \quad (56)$$

The second and third summands in the left-hand side (LHS) of (56) coincide with the expression for the stochastic subgradient $\hat{\mathbf{s}}(t)$ in (8). This allows rewriting (56) in the more compact form

$$f_0(\mathbf{x}(t)) + \boldsymbol{\lambda}^T(t)\hat{\mathbf{s}}(t) \geq f_0(\mathbf{x}) + \boldsymbol{\lambda}_1^T(t)\Big[\mathbb{E}_{m(\mathbf{p}(\mathbf{h}(t)))}\Big(\mathbf{f}_1(\mathbf{p}(\mathbf{h}(t));\mathbf{h}(t))\Big) - \mathbf{x}\Big] + \boldsymbol{\lambda}_2^T(t)\mathbf{f}_2(\mathbf{x}). \quad (57)$$

Let $\boldsymbol{\lambda}(t)$ be given and consider the conditional expected value of both sides of (57). With $\boldsymbol{\lambda}(t)$ given, probability distributions of terms in (57) depend only on the pdf $m(\mathbf{h}(t))$. Thus, conditional expectations can be computed by integrating with respect to $m(\mathbf{h}(t))$ only; i.e., $\mathbb{E}[\cdot\,|\,\boldsymbol{\lambda}(t)] = \mathbb{E}_{\mathbf{h}(t)}[\cdot\,|\,\boldsymbol{\lambda}(t)]$. Further note that because $\mathbf{x}(t)$ is a deterministic function of $\boldsymbol{\lambda}(t)$ it holds $\mathbb{E}[f_0(\mathbf{x}(t))\,|\,\boldsymbol{\lambda}(t)] = f_0(\mathbf{x}(t))$ and that on the RHS of (57) $\mathbf{x}$ and $m(\mathbf{p}(\mathbf{h}))t$ are arbitrary and given. Using these observations we can write

$$f_0(\mathbf{x}(t)) + \boldsymbol{\lambda}^T(t)\mathbb{E}[\hat{\mathbf{s}}(t)\,|\,\boldsymbol{\lambda}(t)] \geq f_0(\mathbf{x}) + \boldsymbol{\lambda}_1^T(t)\Big[\mathbb{E}_{\mathbf{h}}\Big[\mathbb{E}_{m(\mathbf{p}(\mathbf{h}))}\Big(\mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h})\Big)\Big] - \mathbf{x}\Big] + \boldsymbol{\lambda}_2^T(t)\mathbf{f}_2(\mathbf{x}) \quad (58)$$

where we also used the fact that states $\mathbf{h}(t)$ are i.i.d. to drop the time index t. Restrict now attention to feasible primal variables, i.e., to those variables $\mathbf{x}$ and $\{m(\mathbf{p}(\mathbf{h}))\}_{\mathbf{h}\in\mathcal{H}}$ that besides constraints $\mathbf{x} \in \mathcal{X}$ and $\{m(\mathbf{p}(\mathbf{h})) : \mathbf{p}(\mathbf{h}) \in \mathcal{P}(\mathbf{h})\}_{\mathbf{h}\in\mathcal{H}}$, also satisfy $\mathbf{x} \leq \mathbb{E}_{\mathbf{h}}[\mathbb{E}_{m(\mathbf{p}(\mathbf{h}))}(\mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h}))]$ and $\mathbf{f}_2(\mathbf{x}) \geq \mathbf{0}$. Because dual variables $\boldsymbol{\lambda}_1(t) \geq \mathbf{0}$ and $\boldsymbol{\lambda}_2(t) \geq \mathbf{0}$ are nonnegative, it follows that for feasible primal variables the terms $\boldsymbol{\lambda}_1^T(t)[\mathbb{E}_{\mathbf{h}}[\mathbb{E}_{m(\mathbf{p}(\mathbf{h}))}(\mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h}))] - \mathbf{x}] \geq 0$ and $\boldsymbol{\lambda}_2^T(t)\mathbf{f}_2(\mathbf{x}) \geq 0$. Using these bounds and reordering terms in (58) it follows

$$f_0(\mathbf{x}(t)) \geq f_0(\mathbf{x}) - \boldsymbol{\lambda}^T(t)\mathbb{E}[\hat{\mathbf{s}}(t)\,|\,\boldsymbol{\lambda}(t)]. \quad (59)$$

As discussed here, (59) is true for any $\mathbf{x} \in \mathcal{X}$ satisfying $\mathbf{f}_2(\mathbf{x}) \geq \mathbf{0}$ and for which exist resource allocation pdfs $\{m(\mathbf{p}(\mathbf{h})) : \mathbf{p}(\mathbf{h}) \in \mathcal{P}(\mathbf{h})\}_{\mathbf{h}\in\mathcal{H}}$ such that $\mathbf{x} \leq \mathbb{E}_{\mathbf{h}}[\mathbb{E}_{m(\mathbf{p}(\mathbf{h}))}(\mathbf{f}_1(\mathbf{p}(\mathbf{h});\mathbf{h}))]$. In particular, for optimal ergodic limits $\mathbf{x}^*$ and resource allocation distributions $m^*(\mathbf{p}(\mathbf{h}))$, it holds $f_0(\mathbf{x}^*) = P$ and

$$f_0(\mathbf{x}(t)) \geq P - \boldsymbol{\lambda}^T(t)\mathbb{E}[\hat{\mathbf{s}}(t)\,|\,\boldsymbol{\lambda}(t)]. \quad (60)$$

The latter is true for all time indexes t. Fixing arbitrary T_0 and summing between $T_0 + 1$ and t yields

$$\frac{1}{t - T_0}\sum_{u=T_0+1}^{t} f_0(\mathbf{x}(u)) \geq P - \frac{1}{t - T_0}\sum_{u=T_0+1}^{t} \boldsymbol{\lambda}^T(u)\mathbb{E}[\hat{\mathbf{s}}(u)\,|\,\boldsymbol{\lambda}(u)]. \quad (61)$$

The expected value of the latter conditional on $\boldsymbol{\lambda}(T_0)$ with $t \to \infty$ yields after reordering terms

$$P - \lim_{t\to\infty} \mathbb{E}\left[\frac{1}{t - T_0}\sum_{u=T_0+1}^{t} f_0(\mathbf{x}(u))\,|\,\boldsymbol{\lambda}(T_0)\right] \leq \lim_{t\to\infty} \mathbb{E}\left[\frac{1}{t - T_0}\sum_{u=T_0+1}^{t} \boldsymbol{\lambda}^T(u)\hat{\mathbf{s}}(u)\,|\,\boldsymbol{\lambda}(T_0)\right]. \quad (62)$$

To finalize the proof we need to bound $\lim_{t\to\infty}\mathbb{E}[(1/(t - T_0))\sum_{u=T_0+1}^{t}\boldsymbol{\lambda}^T(u)\hat{\mathbf{s}}(u)\,|\,\boldsymbol{\lambda}(T_0)]$ in the RHS of (62). Establishing this bound requires a rather lengthy argument that we present in the next lemma.

Lemma 1 (Average Complementary Slackness): Consider sequences $\mathbf{p}(\mathbb{N})$ and $\mathbf{x}(\mathbb{N})$ generated by (6)–(9) and the same hypotheses and definitions of Theorem 1. For arbitrary dual variable $\boldsymbol{\lambda}(T_0)$ and time T_0 it holds

$$\lim_{t\to\infty} \mathbb{E}\left[\frac{1}{t - T_0}\sum_{u=T_0+1}^{t} \hat{\mathbf{s}}^T(u)\boldsymbol{\lambda}(u)\,|\,\boldsymbol{\lambda}(T_0)\right] \leq \frac{\epsilon\hat{S}^2}{2}. \quad (63)$$

Proof: Start considering the square of the dual iterates' norm $||\boldsymbol{\lambda}(t+1)||^2$ at time $t+1$. Using the expression for the dual iteration in (9) we can write

$$||\boldsymbol{\lambda}(t+1)||^2 = \left\|[\boldsymbol{\lambda}(t) - \epsilon\hat{\mathbf{s}}(t)]^+\right\|^2 \leq ||\boldsymbol{\lambda}(t)||^2 + \epsilon^2||\hat{\mathbf{s}}^2(t)||^2 - 2\epsilon\boldsymbol{\lambda}^T(t)\hat{\mathbf{s}}(t) \quad (64)$$

where the inequality follows by expanding the square after using the fact that projecting in the positive orthant reduces the norm of $\boldsymbol{\lambda}(t) - \epsilon\hat{\mathbf{s}}(t)$.

Compute the conditional expected value of both sides of (64) with $\boldsymbol{\lambda}(t)$ given. By hypothesis, the second term in the RHS is bounded as $\mathbb{E}[||\hat{\mathbf{s}}^2(t)||^2 \,|\, \boldsymbol{\lambda}(t)] \leq \hat{S}^2$. Therefore, we can bound $\mathbb{E}[||\boldsymbol{\lambda}(t+1)||^2 \,|\, \boldsymbol{\lambda}(t)]$ as

$$\mathbb{E}\Big[||\boldsymbol{\lambda}(t+1)||^2 \,|\, \boldsymbol{\lambda}(t)\Big] \leq ||\boldsymbol{\lambda}(t)||^2 + \epsilon^2\hat{S}^2 - 2\epsilon\boldsymbol{\lambda}^T(t)\mathbb{E}[\hat{\mathbf{s}}(t) \,|\, \boldsymbol{\lambda}(t)]. \quad (65)$$

Note that the latter bound in (65) is true for all time indexes t. Conditioning on $\boldsymbol{\lambda}(t-1)$ and integrating (65) over all possible $\boldsymbol{\lambda}(t)$ yields

$$\mathbb{E}\Big[||\boldsymbol{\lambda}(t+1)||^2 \,|\, \boldsymbol{\lambda}(t-1)\Big] \leq \mathbb{E}\Big[||\boldsymbol{\lambda}(t)||^2 \,|\, \boldsymbol{\lambda}(t-1)\Big] + \epsilon^2\hat{S}^2 - 2\epsilon\mathbb{E}\Big[\boldsymbol{\lambda}^T(t)\hat{\mathbf{s}}(t) \,|\, \boldsymbol{\lambda}(t-1)\Big]. \quad (66)$$

We can now bound the first expected value in the RHS of (66) using the expression in (65) to obtain

$$\mathbb{E}\Big[||\boldsymbol{\lambda}(t+1)||^2 \,|\, \boldsymbol{\lambda}(t-1)\Big] \leq ||\boldsymbol{\lambda}(t-1)||^2 + 2\epsilon^2\hat{S}^2 - 2\epsilon\sum_{u=t-1}^{t}\mathbb{E}\Big[\boldsymbol{\lambda}^T(u)\hat{\mathbf{s}}(u) \,|\, \boldsymbol{\lambda}(t-1)\Big]. \quad (67)$$

We can now repeat the arguments leading from (65) to (66), i.e., conditioning in the previous iterate, $\boldsymbol{\lambda}(t-2)$ in this case, and integrating over the iterate under consideration, $\boldsymbol{\lambda}(t-1)$ in this case. The resulting expression will contain a term of the form $\mathbb{E}[||\boldsymbol{\lambda}(t-1)||^2 \,|\, \boldsymbol{\lambda}(t-2)]$ that can be bounded using (65) as was done in obtaining (67) from (66). This process can be repeated iteratively $t - T_0$ times to yield

$$\mathbb{E}\Big[||\boldsymbol{\lambda}(t+1)||^2 \,|\, \boldsymbol{\lambda}(T_0)\Big] \leq ||\boldsymbol{\lambda}(T_0)||^2 + (t-T_0)\epsilon^2\hat{S}^2 - 2\epsilon\sum_{u=T_0+1}^{t}\mathbb{E}\Big[\boldsymbol{\lambda}^T(u)\hat{\mathbf{s}}(u) \,|\, \boldsymbol{\lambda}(T_0)\Big]. \quad (68)$$

Observe now that the LHS of (68) is nonnegative because it is the expected value of a squared norm. After reordering terms and dividing by $2\epsilon(t-T_0)$ this observation leads to

$$\frac{1}{t-T_0}\sum_{u=T_0+1}^{t}\mathbb{E}\Big[\boldsymbol{\lambda}^T(u)\hat{\mathbf{s}}(u) \,|\, \boldsymbol{\lambda}(T_0)\Big] \leq \frac{||\boldsymbol{\lambda}(T_0)||^2}{2\epsilon(t-T_0)} + \frac{\epsilon\hat{S}^2}{2}. \quad (69)$$

To obtain the result in (63) interchange sum and expected value in (69) and consider the limit as $t \to \infty$. ■

To prove Theorem 3 substitute the bound (63) of Lemma 1 into (62) and observe that for any given T_0, $\lim_{t\to\infty}(1/(t-T_0))\sum_{u=T_0+1}^{t} f_0(\mathbf{x}(u)) = \lim_{t\to\infty}(1/t)\sum_{u=1}^{t} f_0(\mathbf{x}(u))$ to obtain (52). ■

Characterizing the mean of a stochastic process across different realizations in the manner of Theorem 3 is of limited value. In practice, a single sequence is observed and the interest is therefore in claims about individual paths. This is the difference between the just proved Theorem 3 and Result (ii) of Theorem 1. The interest in Theorem 3 is as an intermediate step to establish almost sure optimality, as is pursued next.

A. Almost Sure Optimality

The ergodic average $\bar{f}_0(t)$ of the optimality gap at time t is a random variable whose probability distribution is determined by the probability distribution of the state process $\mathbf{H}(\mathbb{N})$. The ergodic limit $\lim_{t\to\infty}\bar{f}_0(t) = P - \lim_{t\to\infty}(1/t)\sum_{u=1}^{\infty} f_0(\mathbf{x}(t))$ exists almost surely because $f_0[\mathbf{x}(t)]$ is a deterministic function of dual variables $\boldsymbol{\lambda}(t)$ that form a Markov process with time invariant transition probabilities. Therefore, $\lim_{t\to\infty}\bar{f}_0(t)$ is also a properly defined random variable. Our goal is to show that with probability one this random variable is smaller than $\epsilon\hat{S}^2/2$. Having proved that the limit of the expected optimality gap $\lim_{t\to\infty}\mathbb{E}[\bar{f}_0(t)]$ converges to a small value in Theorem 3 the simplest path would be to use an ergodicity argument. A stochastic process is ergodic when the ergodic limits, e.g., $\lim_{t\to\infty}\bar{f}_0(t)$, are equal for all process realizations. In such case, $\lim_{t\to\infty}\bar{f}_0(t)$ would not be random and its equivalence with $\lim_{t\to\infty}\mathbb{E}[\bar{f}_0(t)]$ would follow. But there is no indication that ergodicity holds for $\mathbf{X}(\mathbb{N})$ motivating a more involved argument that we develop in this section. The idea is to prove that even though different ergodic limits $\lim_{t\to\infty}\bar{f}_0(t)$ are possible all of them are smaller than $\epsilon\hat{S}^2/2$. The proof of this is based in the following three arguments.

(A1) If $\lim_{t\to\infty}\bar{f}_0(t)$ exists for a given sequence, then $\bar{f}_0(t)$ gets arbitrarily close to its limit for t sufficiently large. Therefore, it is possible to find a time t such that with probability arbitrarily close to 1 all ergodic averages $\bar{f}_0(t)$ are close to their limit. This is proved in Lemma 2.

(A2) Let $\boldsymbol{\lambda}$ denote a given dual variable, t a large time index and consider two sequences $\boldsymbol{\lambda}_1(\mathbb{N})$ and $\boldsymbol{\lambda}_2(\mathbb{N})$ that go through $\boldsymbol{\lambda}$ at time t, i.e., $\boldsymbol{\lambda}_1(t) = \boldsymbol{\lambda}_2(t) = \boldsymbol{\lambda}$. Is it possible for these sequences to have ergodic limits $\lim_{t\to\infty}\bar{f}_{01}(t)$ and $\lim_{t\to\infty}\bar{f}_{02}(t)$ that are significantly far apart? The answer is with large probability no. Ergodic limits $\lim_{t\to\infty}\bar{f}_{0i}(t)$ for $i = 1, 2$ depend only on future values $\boldsymbol{\lambda}_i(t+1:\infty)$. From this observation and the Markov property it follows that since $\boldsymbol{\lambda}_1(t) = \boldsymbol{\lambda}_2(t)$ it is equally likely for a sequence with history $\boldsymbol{\lambda}_1(1:t)$ to end up having the ergodic limit $\lim_{t\to\infty}\bar{f}_{02}(t)$ as it is for a sequence with history $\boldsymbol{\lambda}_2(1:t)$. Thus, $\lim_{t\to\infty}\bar{f}_{01}(t)$ and $\lim_{t\to\infty}\bar{f}_{02}(t)$ are both possible limits for a sequence with history $\boldsymbol{\lambda}_1(1:t)$ for which the current average is $\bar{f}_{01}(t)$. According to (A1), for sufficiently large t the current value $\bar{f}_{01}(t)$ of the ergodic average is expected to be close to both of these possible limits $\lim_{t\to\infty}\bar{f}_{01}(t)$ and $\lim_{t\to\infty}\bar{f}_{02}(t)$. But if the limits $\lim_{t\to\infty}\bar{f}_{01}(t)$ and $\lim_{t\to\infty}\bar{f}_{02}(t)$ are not close to each other then $\bar{f}_{01}(t)$ is not close to at least one of them. This

contradicts (A1). Hence, for sufficiently large t ergodic limits of sequences $\boldsymbol{\lambda}_1(\mathbb{N})$ and $\boldsymbol{\lambda}_2(\mathbb{N})$ that go through a particular dual variable $\boldsymbol{\lambda}$ at time t are close to each other. This is proved in Lemma 3.

(A3) Assume now that a certain sequence $\boldsymbol{\lambda}_1(\mathbb{N})$ has ergodic limit $\lim_{t\to\infty}\bar{f}_{01}(t) > \epsilon\hat{S}^2/2$. For sufficiently large T_0 it must hold $\bar{f}_{01}(T_0) > \epsilon\hat{S}^2/2$. Consider then all sequences with $\boldsymbol{\lambda}(T_0) = \boldsymbol{\lambda}_1(T_0)$. All such sequences have ergodic limits close to $\lim_{t\to\infty}\bar{f}_{01}(t)$ and, therefore, $\lim_{t\to\infty}\bar{f}_0(t) > \epsilon\hat{S}^2/2$ for all sequences with $\boldsymbol{\lambda}(T_0) = \boldsymbol{\lambda}_1(T_0)$. This would imply that $\mathbb{E}[\bar{f}_0(t)\,|\,\boldsymbol{\lambda}(T_0)] > \epsilon\hat{S}^2/2$ contradicting the mean optimality result of Theorem 3. This argument is presented as the proof of Result (ii) of Theorem 1.

Let us then start with the proof of (A1) formally stated in the following Lemma.

Lemma 2: Let $\bar{f}_0(t)$ be the ergodic average of optimality gaps in (51). For arbitrary positive constants $\delta > 0$ and $\gamma > 0$ there exits a sufficiently large time index T_0 such that for all $t \geq T_0$

$$\Pr\Big[\Big|\bar{f}_0(t) - \lim_{u\to\infty}\bar{f}_0(u)\Big| > \gamma/2\Big] \leq \delta \quad \text{for all } t \geq T_0. \tag{70}$$

Proof: The claim is a straightforward restatement of the existence of the ergodic limits $\lim_{t\to\infty}\bar{f}_0(t)$ for almost all sequences. If the ergodic averages of almost all sequences converge, then the averages of almost all sequences eventually come within $\gamma/2$ of their limits. Equivalently, for sufficiently large time index t, only sequences with an arbitrarily small measure δ, have ergodic averages that are more than $\gamma/2$ away from their respective limits. To formalize this argument start fixing arbitrary γ and noting that for any sequence $\boldsymbol{\lambda}(\mathbb{N})$ with convergent ergodic averages there exist a time $t_\gamma[\boldsymbol{\lambda}(\mathbb{N})]$ such that

$$\Big|\bar{f}_0(t) - \lim_{u\to\infty}\bar{f}_0(u)\Big| \leq \gamma/2, \quad \forall t \geq t_\gamma[\boldsymbol{\lambda}(\mathbb{N})]. \tag{71}$$

While the set of sequences with convergent ergodic averages has probability 1, the set of times $t_\gamma[\boldsymbol{\lambda}(\mathbb{N})]$ is not necessarily bounded as the ergodic averages of some sequences may take arbitrarily long time to converge. Still, finite $t_\gamma[\boldsymbol{\lambda}(\mathbb{N})]$ exist for almost all sequences $\boldsymbol{\lambda}(\mathbb{N})$ implying that the probability of $t_\gamma[\boldsymbol{\lambda}(\mathbb{N})]$ exceeding t vanishes as the latter grows; i.e., $\lim_{t\to\infty}\Pr[t_\gamma[\boldsymbol{\lambda}(\mathbb{N})] > t] = 0$. Thus, given arbitrary $\delta > 0$ we can select a sufficiently large time T_0 such that the set of sequences with $t_\gamma[\boldsymbol{\lambda}(\mathbb{N})] > T_0$ has probability less than δ

$$\Pr[t_\gamma[\boldsymbol{\lambda}(\mathbb{N})] > T_0] \leq \delta. \tag{72}$$

For any sequence with $t_\gamma[\boldsymbol{\lambda}(\mathbb{N})] \leq T_0$, it holds $|\lim_{u\to\infty}\bar{f}_0(u) - \bar{f}_0(t)| \leq \gamma/2$ for all $t \geq T_0$. Negating both parts of the previous statement we obtain the equivalent fact that if a sequence $\boldsymbol{\lambda}(\mathbb{N})$ is such that $|\lim_{u\to\infty}\bar{f}_0(u) - \bar{f}_0(t)| > \gamma/2$ for some $t \geq T_0$, then $t_\gamma[\boldsymbol{\lambda}(\mathbb{N})] > T_0$. Therefore

$$\Pr\Big[\Big|\bar{f}_0(t) - \lim_{u\to\infty}\bar{f}_0(u)\Big| > \gamma/2\Big] \leq \Pr[t_\gamma[\boldsymbol{\lambda}(\mathbb{N})] > T_0]. \tag{73}$$

Combining (72) with (73) the claim in (70) follows. ■

We now focus on the claim in (A2) that all the ergodic limits $\lim_{t\to\infty}\bar{f}_0(t)$ for which $\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}$ at some large time index t span a small range γ. The idea here is that there might be some dual values $\boldsymbol{\lambda}$ such that if $\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}$ then it is possible to have ergodic limits separated more than γ. But these values form a transient set in that as t grows the probability of having $\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}$ for some λ in this set, vanishes.

The ergodic limit is a random variable. Restricting attention to sequences with $\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}$ is tantamount to considering the corresponding conditional random variable that we denote as $\lim_{u\to\infty}\bar{f}_0[u\,|\,\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}]$. For future reference, note that since $\boldsymbol{\Lambda}(\mathbb{N})$ is Markov the probability distribution of $\lim_{u\to\infty}\bar{f}_0[u\,|\,\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}]$, being conditioned on $\boldsymbol{\lambda}(t)$, is independent of $\bar{f}_0(t)$. Our interest is in the almost sure range of $\lim_{u\to\infty}\bar{f}_0[u\,|\,\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}]$, i.e., the smallest interval that contains a probability 1 set of possible ergodic limits $\lim_{u\to\infty}\bar{f}_0[u\,|\,\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}]$. If $\bar{f}_{0i}$ and $\bar{f}_{0j}$ are possible ergodic limits for sequences that go through $\boldsymbol{\lambda}$ we want to study the maximum possible difference $\max_{i,j}|\bar{f}_{0i} - \bar{f}_{0j}|$ between them for limits $\bar{f}_{0i}$ and $\bar{f}_{0j}$ that happen with meaningful probability. This is complicated by the need to left out isolated zero probability events prompting the following definition:

$$\bar{f}_{0\min}(\boldsymbol{\lambda}) := \sup_{\bar{f}_0}\bar{f}_0 : \Pr\Big[\lim_{u\to\infty}\bar{f}_0(u) \geq \bar{f}_0\,|\,\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}\Big] = 1. \tag{74}$$

To obtain $\bar{f}_{0\min}$ all possible ergodic limits $\lim_{u\to\infty}\bar{f}_0(u)$ attainable from $\boldsymbol{\lambda}$ are considered. We then look at all the possible $\bar{f}_0$ that lower bound these limits almost surely. The tightest bound is selected as $\bar{f}_{0\min}$. Likewise, define the tightest almost sure upper bound on the ergodic limits as

$$\bar{f}_{0\max}(\boldsymbol{\lambda}) := \inf_{\bar{f}_0}\bar{f}_0 : \Pr\Big[\lim_{u\to\infty}\bar{f}_0(u) \leq \bar{f}_0\,|\,\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}\Big] = 1. \tag{75}$$

The definitions of $\bar{f}_{0\min}(\boldsymbol{\lambda})$ and $\bar{f}_{0\max}(\boldsymbol{\lambda})$ implicitly assume that they do not depend on the conditioning time t at which $\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}$. This is correct since $\boldsymbol{\Lambda}(\mathbb{N})$ is Markov with time invariant transition probabilities.

Using the lower and upper bounds definitions in (74) and (75) we can write the almost sure range of the sequences that go through $\boldsymbol{\lambda}$ at some point in time as $\bar{f}_{0\max}(\boldsymbol{\lambda}) - \bar{f}_{0\min}(\boldsymbol{\lambda})$. Focus first on those values of $\boldsymbol{\lambda}$ for which the ergodic limit range is significant. Consider thus an arbitrary $\gamma > 0$ and define the set

$$\mathcal{T}_\gamma := \{\boldsymbol{\lambda} : \bar{f}_{0\max}(\boldsymbol{\lambda}) - \bar{f}_{0\min}(\boldsymbol{\lambda}) > \gamma\}. \tag{76}$$

We can think of variables $\boldsymbol{\lambda} \notin \mathcal{T}_\gamma$ as those for which all the possible ergodic limits attainable from $\boldsymbol{\lambda}$ are close to each other. Equivalently, variables $\boldsymbol{\lambda} \in \mathcal{T}_\gamma$ are those for which there is nonzero probability of attaining significantly different ergodic limits. At given time t, the ergodic average $\bar{f}_0(t)$ depends only in the past values of the sequence $\boldsymbol{\lambda}(1:t)$. The ergodic limit $\lim_{u\to\infty}\bar{f}_0(u)$, however, is completely determined by the future values of the sequence $\boldsymbol{\lambda}(t:\infty)$. Given that for $\boldsymbol{\lambda} \in \mathcal{T}_\gamma$ there exist possible limits separated by more than γ, some of this limits have to be separated from $\bar{f}_0(t)$ by more than $\gamma/2$. However, as per claim (70) of Lemma 2 the probability of the ergodic average $\bar{f}_0(t)$ coming arbitrarily close to the ergodic limit $\lim_{u\to\infty}\bar{f}_0(u)$ becomes arbitrarily close to 1. We formalize this contradiction in the following lemma to prove that $\Pr[\boldsymbol{\lambda}(t) \in \mathcal{T}_\gamma]$ vanishes as $t \to \infty$.

Lemma 3: For the ergodic average $\bar{f}_0(t)$ in (51) and arbitrary $\gamma > 0$ define the set $\mathcal{T}_\gamma$ of transient dual variables as in (76). As

time goes to infinity, the probability of $\boldsymbol{\lambda}(t)$ taking on an element of $\mathcal{T}_\gamma$ vanishes, i.e.

$$\lim_{t\to\infty} \Pr[\boldsymbol{\lambda}(t) \in \mathcal{T}_\gamma] = 0. \tag{77}$$

Proof: As per its definition in (76) the set T_γ contains dual variables $\boldsymbol{\lambda}$ from which ergodic limits $\lim_{u\to\infty} \bar{f}_0(u)$ separated by more than γ are attainable with nonzero probability. Formally stated, we have that for any $\boldsymbol{\lambda}_0 \in T_\gamma$ there exists $\delta_1(\boldsymbol{\lambda}_0)$ such that

$$\Pr\Big[\Big|\lim_{u\to\infty} \bar{f}_{01}(u) - \lim_{u\to\infty} \bar{f}_{02}(u)\Big| > \gamma \,|\, \boldsymbol{\lambda}_1(t) = \boldsymbol{\lambda}_2(t) = \boldsymbol{\lambda}_0\Big] > \delta_1(\boldsymbol{\lambda}_0). \tag{78}$$

The inequality in (78) is true for all time indices t because the ergodic limits $\lim_{u\to\infty} \bar{f}_{01}(u)$ and $\lim_{u\to\infty} \bar{f}_{02}(u)$ depend only on future iterates.

Consider now an arbitrary constant $\bar{f}_0$ and use the triangle inequality to write

$$\begin{aligned}\Big|\lim_{u\to\infty} \bar{f}_{01}(u) - \bar{f}_0\Big| + \Big|\lim_{u\to\infty} \bar{f}_{02}(u) - \bar{f}_0\Big| \\ \geq \Big|\lim_{u\to\infty} \bar{f}_{01}(u) - \lim_{u\to\infty} \bar{f}_{02}(u)\Big| > \gamma.\end{aligned} \tag{79}$$

The above implies that whenever the difference between limits $|\lim_{u\to\infty} \bar{f}_{01}(u) - \lim_{u\to\infty} \bar{f}_{02}(u)| > \gamma$ at least one of the differences in the LHS, $|\lim_{u\to\infty} \bar{f}_{01}(u) - \bar{f}_0|$ or $|\lim_{u\to\infty} \bar{f}_{02}(u)|$ is greater than $\gamma/2$. In particular, for every pair of sequences $\boldsymbol{\lambda}_1(\mathbb{N})$ and $\boldsymbol{\lambda}_2(\mathbb{N})$ that go through $\boldsymbol{\lambda}_0$ at time t and whose limits are separated by more than γ, at least one of them is separated from $\bar{f}_0$ by at least $\boldsymbol{\gamma}/2$. In terms of sets this observation can be written as

$$\begin{aligned}&\Big\{\Big|\lim_{u\to\infty} \bar{f}_{01}(u) - \lim_{u\to\infty} \bar{f}_{02}(u)\Big| > \gamma\Big\} \subset \\ &\quad \Big\{\Big|\lim_{u\to\infty} \bar{f}_{01}(u) - \bar{f}_0\Big| > \gamma/2\Big\} \cup \Big\{\Big|\lim_{u\to\infty} \bar{f}_{02}(u) - \bar{f}_0\Big| > \gamma/2\Big\}.\end{aligned} \tag{80}$$

Combining (78) and (80), it follows that for arbitrary $\bar{f}_0$ and all $\boldsymbol{\lambda}_0 \in T_\gamma$ there exists $\delta_1(\boldsymbol{\lambda}_0)$ such that

$$\Pr\Big[\Big|\lim_{u\to\infty} \bar{f}_0(u) - \bar{f}_0\Big| > \gamma/2 \,\Big|\, \boldsymbol{\lambda}(t) = \boldsymbol{\lambda}_0\Big] > \delta_1(\boldsymbol{\lambda}_0)/2. \tag{81}$$

The probability in (81) holds true for *any* $\bar{f}_0$. In particular if we consider given $\bar{f}_0(t)$, substitute $\bar{f}_0 = \bar{f}_0(t)$ and recall that the probability distribution of $\lim_{u\to\infty} \bar{f}_0[u \,|\, \boldsymbol{\lambda}(t) = \boldsymbol{\lambda}_0]$ is independent of $\bar{f}_0(t)$, we conclude

$$\Pr\Big[\Big|\lim_{u\to\infty} \bar{f}_0(u) - \bar{f}_0(t)\Big| > \gamma/2 \,|\, \boldsymbol{\lambda}(t) = \boldsymbol{\lambda}_0, \bar{f}_0(t) = \bar{f}_0\Big] > \delta_1(\boldsymbol{\lambda}_0)/2. \tag{82}$$

Since the above inequality holds for all $\bar{f}_0(t)$ we can integrate over the marginal distribution of $\bar{f}_0(t)$ to write

$$\Pr\Big[\Big|\lim_{u\to\infty} \bar{f}_0(u) - \bar{f}_0(t)\Big| > \gamma/2 \,|\, \boldsymbol{\lambda}(t) = \boldsymbol{\lambda}_0\Big] > \delta_1(\boldsymbol{\lambda}_0)/2. \tag{83}$$

Consider now the unconditional probability of the difference between the limit $\lim_{u\to\infty} \bar{f}_0(u)$ and the current time average $\bar{f}_0(t)$ exceeding $\boldsymbol{\gamma}$. This can be written as a marginal integration over possible values of $\boldsymbol{\lambda}(t)$

$$\begin{aligned}&\Pr\Big[\Big|\lim_{u\to\infty} \bar{f}_0(u) - \bar{f}_0(t)\Big| > \gamma/2\Big] \\ &= \int \Pr\Big[\Big|\lim_{u\to\infty} \bar{f}_0(u) - \bar{f}_0(t)\Big| > \gamma/2 \,|\, \boldsymbol{\lambda}(t)\Big] d\Pr[\boldsymbol{\lambda}(t)] \\ &\geq \int_{\boldsymbol{\lambda}(t)\in T_\gamma} \delta_1(\boldsymbol{\lambda}(t))/2 \; d\Pr[\boldsymbol{\lambda}(t)].\end{aligned} \tag{84}$$

If the thesis in (77) is not true, then the probability $\Pr[\boldsymbol{\lambda}(t) \in \mathcal{T}_\gamma]$ is bounded away from 0 for at least some arbitrarily large time indexes. Thus there must exist a $\delta_2 > 0$ such that for any arbitrarily large T_0, there exists a time t for which $\Pr[\boldsymbol{\lambda}(t) \in \mathcal{T}_\gamma] > \delta_2$. Considering the inequality in (84) for such t we see that the RHS involves the integration of strictly positive numbers $\delta_1(\boldsymbol{\lambda}_0)/2 > 0$ over a set of positive measure $\Pr[\boldsymbol{\lambda}(t) \in \mathcal{T}_\gamma] > \delta_2$. It therefore amounts to a strictly positive number. Call it δ so as to write

$$\Pr\Big[\Big|\lim_{u\to\infty} \bar{f}_0(u) - \bar{f}_0(t)\Big| > \gamma/2\Big] \geq \delta > 0 \quad \text{for some } t \geq T_0. \tag{85}$$

The latter contradicts (70) of Lemma 2 because T_0 in (85) is arbitrary. The assumption $\Pr[\boldsymbol{\lambda}(T_0) \in \mathcal{T}_\gamma] > \delta_2$ for arbitrarily large T_0 is then incorrect. The opposite, which implies (77), is true. ■

Lemma 3 justifies the denomination of $\mathcal{T}_\gamma$ as a set of transient variables because as time increases the probability of having iterates $\boldsymbol{\lambda}(t) \in \mathcal{T}_\gamma$ vanishes. An alternative interpretation of Lemma 3 is that for large t, knowing $\boldsymbol{\lambda}(t)$ is sufficient to predict the ergodic limit $\lim_{u\to\infty} \bar{f}_0(u \,|\, \boldsymbol{\lambda}(t))$ with precision γ.

We are now able to prove Result (ii) of Theorem 1 using the results of Theorem 3 and Lemma 3. As per argument (A3) the idea is to consider the set of sequences with ergodic limits not satisfying (12), i.e., $\lim_{u\to\infty} \bar{f}_0(u) > \epsilon\hat{S}^2/2$. We then fix a certain time t and separate analysis into those sequences with $\boldsymbol{\lambda}(t) \in \mathcal{T}_\gamma$ and those with $\boldsymbol{\lambda}(t) \notin \mathcal{T}_\gamma$. The first set, i.e., $\boldsymbol{\lambda}(t) \in \mathcal{T}_\gamma$ has vanishing probability as per Lemma 3 reducing consideration to the second set with $\boldsymbol{\lambda}(t) \notin \mathcal{T}_\gamma$. For these sequences we know that with probability 1 the ergodic limits $\lim_{u\to\infty} \bar{f}_0(u \,|\, \boldsymbol{\lambda}(t))$ are within γ of each other. Thus, if for some sequence the ergodic limit is $\lim_{u\to\infty} \bar{f}_0(u \,|\, \boldsymbol{\lambda}(t)) > \epsilon\hat{S}^2/2 + \gamma$, then for all sequences that go through $\boldsymbol{\lambda}(t)$ it holds $\lim_{u\to\infty} \bar{f}_0(u \,|\, \boldsymbol{\lambda}(t)) > \epsilon\hat{S}^2/2$. From here we conclude that $\lim_{u\to\infty} \mathbb{E}[\bar{f}_0(u) \,|\, \boldsymbol{\lambda}(t)] > \epsilon\hat{S}^2/2$. This contradicts Theorem 3, consequently, it is not possible to have sequences with $\lim_{u\to\infty} \bar{f}_0(u \,|\, \boldsymbol{\lambda}(t)) > \epsilon\hat{S}^2/2 + \gamma$. But since γ is arbitrarily selected, the claim in (12) follows. This argument is formalized in the following proof.

Proof (Result (ii) of Theorem 1): We want to prove that the nonzero probability range of $\lim_{u\to\infty} \bar{f}_0(u)$ is bounded above by $\epsilon\hat{S}^2/2$. Equivalently define the set of sequences $\mathcal{A}_\gamma$ whose ergodic averages fall "above" $\epsilon\hat{S}^2/2$ by at least γ

$$\mathcal{A}_\gamma = \Big\{\boldsymbol{\lambda}(\mathbb{N}) : \lim_{u\to\infty} \bar{f}_0(u) > \epsilon\hat{S}^2/2 + \gamma\Big\}. \tag{86}$$

The result follows if we prove that $\Pr[\mathcal{A}_\gamma] = 0$ for all $\gamma > 0$. This is in turn equivalent to proving that for all $\gamma > 0$ and arbitrary $\delta > 0$ it holds $\Pr[\mathcal{A}_\gamma] < \delta$. We will prove this latter statement.

For that purpose fix time index t and consider the conditional random variable obtained by restricting attention to sequences with $\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}$. Denote this variable as $\lim_{u\to\infty} \bar{f}_0(u \,|\, \boldsymbol{\lambda}(t) = \boldsymbol{\lambda})$. The probability of the set $\mathcal{A}_\gamma$ can be then computed by conditioning on $\boldsymbol{\lambda}(t)$ leading to the expression

$$\Pr[\mathcal{A}_\gamma] = \int \Pr[\mathcal{A}_\gamma \,|\, \boldsymbol{\lambda}(t)]\, d\Pr[\boldsymbol{\lambda}(t)]. \tag{87}$$

The latter integral can be separated between the multiplier values $\boldsymbol{\lambda}(t)$ that belong to the transient set T_γ and those that do not, yielding

$$\begin{aligned}\Pr[\mathcal{A}_\gamma] &= \int_{\boldsymbol{\lambda}(t)\notin T_\gamma} \Pr[\mathcal{A}_\gamma \,|\, \boldsymbol{\lambda}(t)] d\Pr[\boldsymbol{\lambda}(t)] \\ &\quad + \int_{\boldsymbol{\lambda}(t)\in T_\gamma} \Pr[\mathcal{A}_\gamma \,|\, \boldsymbol{\lambda}(t)] d\Pr[\boldsymbol{\lambda}(t)] \\ &\le \int_{\boldsymbol{\lambda}(t)\notin T_\gamma} \Pr[\mathcal{A}_\gamma \,|\, \boldsymbol{\lambda}(t)] d\Pr[\boldsymbol{\lambda}(t)] + \Pr[\boldsymbol{\lambda}(t) \in \mathcal{T}_\gamma].\end{aligned} \tag{88}$$

The inequality in (88) follows from the fact that the second term in the RHS of (88) integrates a probability over the set $\boldsymbol{\lambda}(t) \in T_\gamma$. Its value, then, cannot be larger than the measure of the set $\Pr[\boldsymbol{\lambda}(t) \in \mathcal{T}_\gamma]$.

Note now that according to Lemma 3 the probability of the transient set vanishes. Therefore, we can select sufficiently large T_0 to bound that term by δ therefore yielding

$$\Pr[\mathcal{A}_\gamma] \le \int_{\boldsymbol{\lambda}(T_0)\notin T_\gamma} \Pr[\mathcal{A}_\gamma \,|\, \boldsymbol{\lambda}(T_0)]\, d\Pr[\boldsymbol{\lambda}(T_0)] + \delta. \tag{89}$$

We will now prove that the probability $\Pr \mathcal{A}_\gamma \,|\, \boldsymbol{\lambda}(T_0)$ for any $\boldsymbol{\lambda}(T_0) \notin T_\gamma$ is null. Assume the probability is not null. If this were true there would be a set of sequences going through $\boldsymbol{\lambda}(T_0)$ with limits

$$\Pr\Big[\lim_{u\to\infty} \bar{f}_0[u \,|\, \boldsymbol{\lambda}(T_0)] > \epsilon\hat{S}^2/2 + \gamma\Big] > \delta_1. \tag{90}$$

But according to the definition of T_γ, if $\boldsymbol{\lambda}(T_0) \notin T_\gamma$, then all the limits of sequences that go through $\boldsymbol{\lambda}(T_0)$ are almost surely within γ of each other [cf. (74)–(76)]. In particular, the latter implies that

$$\Pr\Big[\lim_{u\to\infty} \bar{f}_0[u \,|\, \boldsymbol{\lambda}(T_0)] > \epsilon\hat{S}^2/2\Big] = 1. \tag{91}$$

But if all the limits are larger than $\epsilon\hat{S}^2/2$

$$\mathbb{E}\Big[\lim_{u\to\infty} \bar{f}_0[u \,|\, \boldsymbol{\lambda}(T_0)]\Big] > \epsilon\hat{S}^2/2. \tag{92}$$

The latter contradicts Theorem 3. Therefore, we must have $\Pr[\mathcal{A}_\gamma \,|\, \boldsymbol{\lambda}(T_0)] = 0$ for all $\boldsymbol{\lambda}(T_0) \notin T_\gamma$. Substituting this into (89) yields $\Pr \mathcal{A}_\gamma \le \delta$. As argued before the latter implies that the ergodic average $\bar{f}_0(t)$ almost surely converges to a value below $\epsilon\hat{S}^2/2$ as indicated in (12). ■

V. Conclusion

We have developed ESO algorithms to realize optimal resource allocations in problems with long time horizons allowing optimality criteria defined in terms of ergodic limits. In the problems considered, a certain resource $\mathbf{p}(t)$ is allocated at time t, in response to a random state realization $\mathbf{h}(t)$. State and resource allocation constrain the production of a certain good $\mathbf{x}(t)$ that we seek to optimize. We proposed an algorithm using stochastic subgradient descent in the dual function and showed that with probability 1 problem constraints are satisfied and close to optimal production achieved. ESO algorithms do not require access to the state probability distribution and while they assume convexity of objective functions and constraints with respect to $\mathbf{x}(t)$, they do not require convexity with respect to the resource allocation $\mathbf{p}(t)$.

The type of optimal resource allocation problem solved by the ESO algorithm arise frequently in signal processing. An application to find the optimal operating point of an orthogonal frequency division multiplexing broadcast channel was considered as an example of a large scale nonconvex optimization problem that can be solved with reasonable computational cost. We have also considered applications to general wireless networking problems in [2] and to random access channels in [27]. Applications to different problems, e.g., cognitive radio, beamforming, and multiple input multiple access channels are left for further research.

Appendix
Convergence of Dual Iterates

As shown by (32) the stochastic subgradient in (8) is, on average, a descent direction for $g(\boldsymbol{\lambda}(t))$. A preliminary step in the proof of Theorem 2 is to quantify the descent associated with the dual iteration in (9). This is introduced in the following theorem.

Theorem 4: Conisder the ESO algorithm as defined by (6)–(9) with the same hypothesis and definitions of Theorem 1. Denote as $\boldsymbol{\lambda}^*$ any optimal argument of the convex function $g(\boldsymbol{\lambda})$. Then, the distances $\|\boldsymbol{\lambda}(t) - \boldsymbol{\lambda}^*\|^2$ between iterates $\boldsymbol{\lambda}(t)$ and the optimal argument $\boldsymbol{\lambda}^*$ satisfy the recursive inequality

$$\mathbb{E}\Big[\|\boldsymbol{\lambda}(t+1) - \boldsymbol{\lambda}^*\|^2 \,|\, \boldsymbol{\lambda}(t)\Big] \le \|\boldsymbol{\lambda}(t) - \boldsymbol{\lambda}^*\|^2 + \epsilon^2\hat{S}^2 - 2\epsilon[g(t) - D]. \tag{93}$$

Proof: Steps are similar to those used to prove a similar result for deterministic subgradients [17, Ch. 2]. Consider the 2-norm distance $\|\boldsymbol{\lambda}(t+1) - \boldsymbol{\lambda}^*\|$ at time $t+1$. Using (9) it is possible to relate $\|\boldsymbol{\lambda}(t+1) - \boldsymbol{\lambda}^*\|$ with the distance $\|\boldsymbol{\lambda}(t) - \boldsymbol{\lambda}^*\|$ in the previous time slot

$$\|\boldsymbol{\lambda}(t+1) - \boldsymbol{\lambda}^*\|^2 = \Big\|\Big[\boldsymbol{\lambda}(t) - \epsilon\hat{\mathbf{s}}(t)\Big]^+ - \boldsymbol{\lambda}^*\Big\|^2 \tag{94}$$

$$\le \|\boldsymbol{\lambda}(t) - \epsilon\hat{\mathbf{s}}(t) - \boldsymbol{\lambda}^*\|^2 \tag{95}$$

$$= \|\boldsymbol{\lambda}(t) - \boldsymbol{\lambda}^*\|^2 + \epsilon^2\|\hat{\mathbf{s}}(t)\|^2 - 2\epsilon\hat{\mathbf{s}}^T(t)(\boldsymbol{\lambda}(t) - \boldsymbol{\lambda}^*). \tag{96}$$

The inequality in (95) follows because setting negative components of $\boldsymbol{\lambda}(t) - \epsilon\hat{\mathbf{s}}(t)$ to zero reduces the distance to (positive) $\boldsymbol{\lambda}^*$. Expanding the square in (95) yields (96).

For given $\boldsymbol{\lambda}(t)$ the relations in (94)–(96) hold for all realizations of $\hat{\mathbf{s}}(t)$. Therefore, the expectation of (96) conditional on $\boldsymbol{\lambda}(t)$ can be written as

$$\mathbb{E}\Big[\|\boldsymbol{\lambda}(t+1)-\boldsymbol{\lambda}^*\|^2 \,|\, \boldsymbol{\lambda}(t)\Big] \leq \|\boldsymbol{\lambda}(t)-\boldsymbol{\lambda}^*\|^2 + \epsilon^2\mathbb{E}\Big[\|\hat{\mathbf{s}}(t)\|^2 \,|\, \boldsymbol{\lambda}(t)\Big] - 2\epsilon\mathbb{E}\Big[\hat{\mathbf{s}}^T(t) \,|\, \boldsymbol{\lambda}(t)\Big](\boldsymbol{\lambda}(t)-\boldsymbol{\lambda}^*). \quad (97)$$

The second moment of the norm of $\hat{\mathbf{s}}(t)$ is assumed bounded by hypothesis. The second term on the RHS of (97) can then be substituted for $\mathbb{E}[\|\hat{\mathbf{s}}(t)\|^2 \,|\, \boldsymbol{\lambda}(t)] \leq \hat{S}^2$. For the third term $\mathbb{E}[\hat{\mathbf{s}}^T(t) \,|\, \boldsymbol{\lambda}(t)](\boldsymbol{\lambda}(t)-\boldsymbol{\lambda}^*)$ use (23) of Proposition 1 with $\boldsymbol{\lambda}=\boldsymbol{\lambda}^*$ and $g(\boldsymbol{\lambda}) = g(\boldsymbol{\lambda}^*) = D$. Plugging both bounds in (97) yields (93). ■

For iterates $\boldsymbol{\lambda}(t)$ with large optimality gap $g(t)-D$, or equivalently for small enough stepsize ϵ, the term $2\epsilon[g(t)-D]$ dominates $\epsilon^2\hat{S}^2$. In such case we expect a reduction of the distance $\|\boldsymbol{\lambda}(t)-\boldsymbol{\lambda}^*\|$ between iterates $\boldsymbol{\lambda}(t)$ and any optimal dual argument $\boldsymbol{\lambda}^*$. Reasoning along this line it seems plausible to establish some form of convergence of $\boldsymbol{\lambda}(t)$ to $\boldsymbol{\lambda}^*$. A first complication is that as the optimality gap $g(t)-D$ is reduced, the decreasing term $2\epsilon[g(t)-D]$ eventually becomes smaller than the fixed term $\epsilon^2\hat{S}^2$ suggesting that $\boldsymbol{\lambda}(t)$ does not converge to $\boldsymbol{\lambda}^*$ but to a neighborhood of it. Because this is in line with the claim in (33) this is not a significant issue. A more challenging complication is the fact that (93) holds on average while our interest is in establishing almost sure convergence results. This mismatch can be addressed by resorting to an argument based on supermatingales as discussed in the following section.

A. Supermartingales and Proof of Theorem 2

A stochastic process $A(\mathbb{N})$ with realizations $\alpha(\mathbb{N})$ and a sequence of nested σ-algebras $\mathcal{A}(t)$ measuring at least the first t elements $\alpha(0:t)$ of $A(\mathbb{N})$ are said to form a supermartingale if $\mathbb{E}[\alpha(t+1) \,|\, \mathcal{A}(t)] \leq \alpha(t)$. The inequality $\mathbb{E}[\alpha(t+1) \,|\, \mathcal{A}(t)] \leq \alpha(t)$ implies that given the past history of the process as measured by $\mathcal{A}(t)$, the next value $\alpha(t+1)$ is, on average, not larger than $\alpha(t)$. If the process elements are deterministic functions they have to be monotone nonincreasing. In that sense, supermartingales can be interpreted as generalizing the concept of monotone nonincreasing function to stochastic processes. If we further lower bound the process, e.g., $\alpha(t) \geq 0$, then it is intuitively expected that $\lim_{t\to\infty}\alpha(t)$ exists. This is, in fact, a fundamental result in stochastic processes that we repeat here for convenience; see, e.g., [28, Th. E7.4].

Theorem 5 (Supermartingale Convergence Theorem): Consider nonnegative stochastic processes $A(\mathbb{N})$ and $B(\mathbb{N})$ with realizations $\alpha(\mathbb{N})$ and $\beta(\mathbb{N})$ having values $\alpha(t) \geq 0$ and $\beta(t) \geq 0$ and a sequence of nested σ-algebras $\mathcal{A}(0:t)$ measuring at least $\alpha(0:t)$ and $\beta(0:t)$. If

$$\mathbb{E}[\alpha(t+1) \,|\, \mathcal{A}(0:t)] \leq \alpha(t) - \beta(t) \quad (98)$$

the sequence $\alpha(t)$ converges almost surely and $\beta(t)$ is almost surely summable, i.e., $\sum_{u=1}^{\infty}\beta(u) < \infty$ a.s.

To prove convergence of stochastic descent algorithms supermartingale arguments are commonly used, e.g., ([28, Appendix E]). For stochastic subgradients using vanishing step sizes a simple proof of convergence using the supermartingale convergence theorem can be found in [17, Ch. 3]. For constant step sizes as considered here the proof is complicated by the fact that the relation in (93) is somewhat different from the definition of a supermartingale. The sequence $\alpha(t)$ in (98) is akin to the squared distance $\|\boldsymbol{\lambda}(t)-\boldsymbol{\lambda}^*\|^2$ in (93). Likewise, $\beta(t)$ is akin to the term $-\epsilon^2\hat{S}^2 + 2\epsilon[g(t)-D]$. However, the statistical relations in (93) and (98) are not identical because $-\epsilon^2\hat{S}^2 + 2\epsilon[g(t)-D]$ may become negative for small duality gap $g(t)-D < \epsilon\hat{S}^2/2$ while the sequence $\beta(t)$ in Theorem 5 is required to be nonnegative.

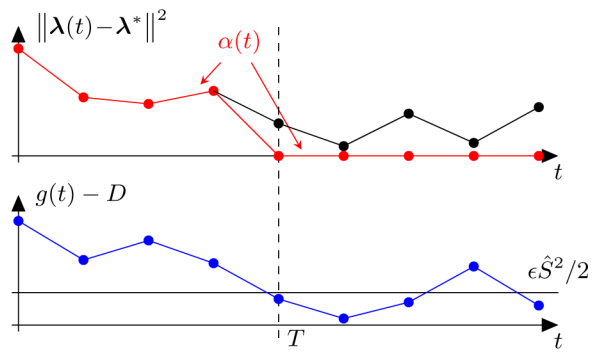

Fig. 5. Definition of the sequence $\alpha(t)$ in the proof of Theorem 2 [cf. (99)]. At time T, $g(t)-D$ falls below $\epsilon\hat{S}^2/2$ for the first time. For times smaller than T, $\alpha(t)$ coincides with $\|\boldsymbol{\lambda}(t)-\boldsymbol{\lambda}^*\|^2$. For times larger than T, $\alpha(t)=0$. A stochastic process with realizations $\alpha(\mathbb{N})$ is proved to be a supermartingale.

Notwithstanding, if at some time $t \geq T_0$ we have $g_{\text{best}}(t \,|\, (\boldsymbol{\lambda}(T_0)) - D < \epsilon\hat{S}^2/2$, then the inequality in (33) follows for the given sequence. If the opposite is true, i.e., $g_{\text{best}}(t \,|\, (\boldsymbol{\lambda}(T_0)) - D \geq \epsilon\hat{S}^2/2$ then it must be $g(u)-D \geq \epsilon\hat{S}^2/2$ for all time indexes $T_0 \leq u \leq t$. In such case, the squared distances $\|\boldsymbol{\lambda}(t)-\boldsymbol{\lambda}^*\|^2$ satisfy the supermartingale hypothesis in (98) for $T_0 \leq u \leq t$. It appears, then, that Theorem 5 can be used to study the path of the sequence $\boldsymbol{\lambda}(\mathbb{N})$ during those times for which $g(t)-D \geq \epsilon\hat{S}^2/2$. This is done in the following proof that uses Theorem 5 as the basis to establish Theorem 2.

Proof (Theorem 2): For simplicity of exposition let $T_0 = 0$. Start defining the sequence $\alpha(t)$ to track $\|\boldsymbol{\lambda}(t)-\boldsymbol{\lambda}^*\|^2$ until the optimality gap $g(t)-D$ falls below $\epsilon\hat{S}^2/2$ for the first time, i.e.

$$\alpha(t) := \|\boldsymbol{\lambda}(t)-\boldsymbol{\lambda}^*\|^2 \mathbb{I}\Big\{g_{\text{best}}(t) - D > (\epsilon\hat{S}^2)/2\Big\}. \quad (99)$$

To explain the definition of $\alpha(t)$ in (99) define $T := \min\{t : g(t)-D \leq \epsilon\hat{S}^2/2\}$ as the first time for which the optimality gap $g(t)-D$ drops below $\leq \epsilon\hat{S}^2/2$; see Fig. 5. For times $t < T$ it clearly holds $g_{\text{best}}(t)-D \geq (\epsilon\hat{S}^2)/2$, the indicator variable in (99) is 1 and the sequence $\alpha(t)$ coincides with $\|\boldsymbol{\lambda}(t)-\boldsymbol{\lambda}^*\|^2$. For $t \geq T$ the best dual iterate is $g_{\text{best}}(t)-D \leq g(T)-D \leq \epsilon\hat{S}^2/2$, the indicator variable is 0 and $\alpha(t)=0$. Consequently, $\alpha(t)$ coincides with $\|\boldsymbol{\lambda}(t)-\boldsymbol{\lambda}^*\|^2$ until the optimality gap $g(t)-D$ falls below $\epsilon\hat{S}^2/2$ for the first time. From this time on we set $\alpha(t)=0$.

Similarly, define the sequence

$$\beta(t) := \Big[2\epsilon[g(t)-D]-\epsilon^2\hat{S}^2\Big]\mathbb{I}\Big\{g_{\text{best}}(t)-D > \epsilon\hat{S}^2/2\Big\} \quad (100)$$

that follows $2\epsilon[g(t) - D] - \epsilon^2\hat{S}^2$ until $g(t) - D$ falls below $\epsilon\hat{S}^2/2$ for the first time. Let also, $\mathcal{A}(0 : t)$ be a sequence of σ-algebras measuring $\alpha(0 : t)$, $\beta(0 : t)$ and $\boldsymbol{\lambda}(0 : t)$. It will be shown that $\alpha(t)$ and $\beta(t)$ comply with the hypotheses of Theorem 5 with respect to the sequence of σ-algebras $\mathcal{A}(0 : t)$ The values of $\alpha(t)$ and $\beta(t)$ are completely determined by $\boldsymbol{\lambda}(t)$. Therefore, conditioning on $\mathcal{A}(0 : t)$ is equivalent to conditioning on $\boldsymbol{\lambda}(0 : t)$. Further recalling the Markov assumption on the process $\boldsymbol{\Lambda}(\mathbb{N})$ it follows that $\mathbb{E}[\alpha(t) \,|\, \mathcal{A}(0 : t)] = \mathbb{E}[\alpha(t) \,|\, \boldsymbol{\lambda}(0 : t)] = \mathbb{E}[\alpha(t) \,|\, \boldsymbol{\lambda}(t)]$. The latter expectation is further separated in the cases when $\alpha(t) = 0$ and $\alpha(t) \neq 0$ to write

$$\mathbb{E}[\alpha(t) \,|\, \mathcal{A}(0 : t)] = \mathbb{E}[\alpha(t) \,|\, \boldsymbol{\lambda}(t), \alpha(t) = 0]\Pr[\alpha(t) = 0] + \mathbb{E}[\alpha(t) \,|\, \boldsymbol{\lambda}(t), \alpha(t) \neq 0]\Pr[\alpha(t) \neq 0]. \quad (101)$$

Start considering the case when $\alpha(t) = 0$. The definitions in (99) and (100) dictate that if $\alpha(t) = 0$, then it must be $\beta(t) = 0$ and $\alpha(t+1) = 0$. To see this formally note that if $\alpha(t) = 0$ then, either $\|\boldsymbol{\lambda}(t) - \boldsymbol{\lambda}^*\|^2 = 0$, or $g_{\text{best}}(t) - D \leq (\epsilon\hat{S}^2)/2$ so that the indicator function is 0, or both are true. If $\|\boldsymbol{\lambda}(t) - \boldsymbol{\lambda}^*\|^2 = 0$ then it must be $\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}^*$, in which case $g(t) = g[(\boldsymbol{\lambda}^*)] = D$ in turn implying that $g(t) - D = 0 \leq (\epsilon\hat{S}^2)/2$. In any event, $\alpha(t) = 0$ implies that $g_{\text{best}}(t) - D \leq \epsilon\hat{S}^2/2$. From here it follows that $\beta(t) = 0$ when $\alpha(t) = 0$. Further noting that $g_{\text{best}}(t+1) \leq g_{\text{best}}(t)$ it follows that $\alpha(t) = 0$ implies $\alpha(t+1) = 0$. The following equality is therefore evident because all terms are zero

$$\mathbb{E}[\alpha(t+1) \,|\, \boldsymbol{\lambda}(t), \alpha(t) = 0] = \alpha(t) - \beta(t). \quad (102)$$

When $\alpha(t) \neq 0$ the conditional expected value of $\alpha(t+1)$ is

$$\begin{aligned}
&\mathbb{E}[\alpha(t+1) \,|\, \boldsymbol{\lambda}(t), \alpha(t) \neq 0] \\
&\quad = \mathbb{E}\Big[\|\boldsymbol{\lambda}(t+1) - \boldsymbol{\lambda}^*\|^2 \mathbb{I}\Big\{g(t+1) - D \geq \epsilon\hat{S}^2/2\Big\} \,|\, \boldsymbol{\lambda}(t)\Big] && (103)\\
&\quad \leq \mathbb{E}\Big[\|\boldsymbol{\lambda}(t+1) - \boldsymbol{\lambda}^*\|^2 \,|\, \boldsymbol{\lambda}(t)\Big] && (104)\\
&\quad \leq \|\boldsymbol{\lambda}(t) - \boldsymbol{\lambda}^*\|^2 + \epsilon^2\hat{S}^2 - 2\epsilon[g(t) - D] && (105)\\
&\quad = \alpha(t) - \beta(t). && (106)
\end{aligned}$$

The equality in (103) follows from the definition of $\alpha(t+1)$ in (99) and noting that because $\alpha(t) \neq 0$, $g_{\text{best}}(t+1) - D \leq \epsilon\hat{S}^2/2$ if and only if $g(t+1) - D \leq \epsilon\hat{S}^2/2$. The first inequality in (104) is true because the indicator term is not larger than 1. The second inequality in (105) follows from the expected distance reduction expression in (93) of Theorem 4. The last equality (106) follows from the definitions of $\alpha(t)$ and $\beta(t)$ in (99) and (100) after noting that $\alpha(t) \neq 0$ implies $\mathbb{I}\{g_{\text{best}}(t) - D > (\epsilon\hat{S}^2)/2\} = 1$.

Substituting (102) and (106) into (101), it finally follows

$$\begin{aligned}
&\mathbb{E}[\alpha(t+1) \,|\, \mathcal{A}(0 : t)] \\
&\quad \leq \Big[\alpha(t) - \beta(t)\Big]\Big[\Pr[\alpha(t) = 0] + \Pr[\alpha(t) \neq 0]\Big] \\
&\quad = \alpha(t) - \beta(t). && (107)
\end{aligned}$$

Given (107) and the fact that $\alpha(t)$ and $\beta(t)$ are nonnegative by definition if follows from Theorem 5 that: (i) $\alpha(t)$ converges almost surely; and (ii) the sum $\sum_{t=1}^{\infty} \beta(t) < \infty$ is almost surely finite. Writing the latter consequence in terms of the explicit value of $\beta(t)$ in (100) yields

$$\sum_{t=1}^{\infty} \Big[-\epsilon^2\hat{S}^2 + 2\epsilon[g(t) - D]\Big] \mathbb{I}\Big\{g_{\text{best}}(t) - D > \epsilon\hat{S}^2/2\Big\} \leq \infty \text{ a.s.} \quad (108)$$

In particular, the almost sure convergence of the sum in (108) implies that

$$\liminf_{t\to\infty} \Big[-\epsilon^2\hat{S}^2 + 2\epsilon[g(t) - D]\Big] \mathbb{I}\Big\{g_{\text{best}}(t) - D > \epsilon\hat{S}^2/2\Big\} = 0 \text{ a.s.} \quad (109)$$

The latter is true if either $g_{\text{best}}(t) - D \leq \epsilon\hat{S}^2/2$ for sufficiently large t—so that the indicator function is null—or if $\liminf_{t\to\infty}[-\epsilon^2\hat{S}^2 + 2\epsilon[g(t) - D]] = 0$. From any of these two events, (33) follows. ■

REFERENCES

[1] A. Ribeiro, "Ergodic stochastic optimization algorithms for wireless communication and networking," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, Dallas, TX, Mar. 14–19, 2010 [Online]. Available: http://www.seas.upenn.edu/~aribeiro/preprints/eso_icassp.pdf

[2] A. Ribeiro, "Stochastic learning algorithms for optimal design of wireless fading networks," in *Proc. IEEE Workshop on Signal Process. Adv. Wireless Commun.*, Marrakech, Morocco, Jun. 20–23, 2010 [Online]. Available: http://www.seas.upenn.edu/~aribeiro/preprints/eso_spawc.pdf, (to appear)

[3] X. Wang and G. B. Giannakis, "Resource allocation for wireless multiuser OFDM networks—Ergodic capacity and rate-guaranteed scheduling," *IEEE Trans. Inf. Theory*, Dec. 2008, submitted for publication.

[4] N. D. Sidiropoulos, T. N. Davidson, and Z. Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.

[5] V. Ntranos, N. D. Sidiropoulos, and L. Tassiulas, "On multicast beamforming for minimum outage," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3172–3181, June 2009.

[6] J. A. Bazerque and G. B. Giannakis, "Distributed scheduling and resource allocation for cognitive ofdma radios," *Mobile Nets, Apps.*, vol. 13, no. 5, pp. 452–462, Oct. 2008.

[7] Z. Quan, S. Cui, and A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 28–40, Feb. 2008.

[8] X. Lin, N. B. Shroff, and R. Srikant, "A tutorial on cross-layer optimization in wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1452–1463, Aug. 2006.

[9] A. Eryilmaz and R. Srikant, "Joint congestion control, routing, and MAC for stability and fairness in wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1514–1524, Aug. 2006.

[10] X. Wang and K. Kar, "Cross-layer rate optimization for proportional fairness in multihop wireless networks with random access," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1548–1559, Aug. 2006.

[11] L. Chen, S. H. Low, M. Chiang, and J. C. Doyle, "Cross-layer congestion control, routing and scheduling design in ad hoc wireless networks," in *Proc. IEEE INFOCOM*, Barcelona, Spain, Apr. 23–29, 2005, pp. 1–13.

[12] Y. Yi and S. Shakkottai, "Hop-by-hop congestion control over a wireless multi-hop network," *IEEE/ACM Trans. Netw.*, vol. 15, no. 133–144, pp. 1548–1559, Feb. 2007.

[13] M. Chiang, S. H. Low, R. A. Calderbank, and J. C. Doyle, "Layering as optimization decomposition," *Proc. IEEE*, vol. 95, no. 1, pp. 255–312, Jan. 2007.

[14] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic power allocation and routing for time-varying wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 1, pp. 89–103, Jan. 2005.

[15] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and crosslayer control in wireless networks," *Foundations and Trends in Networking*, vol. 1, no. 1, pp. 1–144, 2006.

[16] J.-W. Lee, R. R. Mazumdar, and N. B. Shroff, "Opportunistic power scheduling for dynamic multi-server wireless systems," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1506–1515, Jun. 2006.

[17] N. Z. Shor, *Minimization Methods for Non-Differentiable Functions*. Berlin, Heilderberg, Germany: Springer-Verlag, 1985.

[18] B. T. Polyak, "New stochastic approximation type procedures," *Autom. Remote Control*, vol. 51, pp. 937–946, 1990.

[19] B. T. Polyak and A. B. Juditsky, "Acceleration of stochastic approximation by averaging," *SIAM J. Control Optim.*, vol. 30, no. 4, pp. 838–855, Jul. 1992.

[20] A. Juditsky, G. Lan, A. Nemirovski, and A. Shapiro, "Stochastic approximation approach to stochastic programming," *SIAM J. Optim.*, vol. 19, no. 4, pp. 1574–1609, Feb. 2009.

[21] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Stat*, vol. 22, no. 3, pp. 400–407, Sep. 1951.

[22] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 2002.

[23] A. Ribeiro and G. B. Giannakis, "Separation principles in wireless networking," *IEEE Trans. Inf. Theory* vol. 56, no. 9, pp. 4488–4505, Sep. 2010.

[24] T. Larsson, M. Patriksson, and A. Str omberg, "Ergodic primal convergence in dual subgradient schemes for convex programming," *Math. Program.*, vol. 86, no. 2, pp. 283–312, Nov. 1999.

[25] A. Nedic and A. Ozdaglar, "Approximate primal solutions and rate analysis for dual subgradient methods," *SIAM J. Optim.*, vol. 19, no. 4, pp. 1757–1780, Feb. 2009.

[26] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Trans. Autom. Control*, vol. 37, no. 12, pp. 1936–1948, Dec. 1992.

[27] Y. Hu and A. Ribeiro, "Adaptive distributed algorithms for optimal random access channels," *IEEE Trans. Wireless Commun.* Feb. 2010 [Online]. Available: http://www.seas.upenn.edu/~yichuan/paper/ra_twc10.pdf, submitted for publication

[28] V. Solo and X. Kong, *Adaptive Signal Processing Algorithms: Stability and Performance*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

Alejandro Ribeiro (M'07) received the B.Sc. degree in electrical engineering from the Universidad de la Republica Oriental del Uruguay, Montevideo, in 1998. He received the M.Sc. and Ph.D. degrees in electrical engineering from the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, in 2005 and 2007, respectively.

He is an Assistant Professor with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, where he has been since 2008. From 1998 to 2003, he was a member of the Technical Staff at Bellsouth Montevideo. His research interests lie in the areas of communication, signal processing, and networking. His current research focuses on wireless communications and networking, distributed signal processing, and wireless sensor networks.

Dr. Ribeiro is a Fulbright scholar, received Best Student Paper awards at ICASSP 2005 and ICASSP 2006, and a National Science Foundation CAREER award in 2010.