Learning in Wireless Control Systems Over Nonstationary Channels

Mark Eisen ⓘ, Konstantinos Gatsis ⓘ, *Member, IEEE*, George J. Pappas ⓘ, *Fellow, IEEE*, and Alejandro Ribeiro ⓘ

Abstract—**This paper considers a set of multiple independent control systems that are each connected over a nonstationary wireless channel. The goal is to maximize control performance over all the systems through the allocation of transmitting power within a fixed budget. This can be formulated as a constrained optimization problem examined using Lagrangian duality. By taking samples of the unknown wireless channel at every time instance, the resulting problem takes on the form of empirical risk minimization, a well-studied problem in machine learning. Due to the nonstationarity of wireless channels, optimal allocations must be continuously learned and updated as the channel evolves. The quadratic convergence property of Newton's method motivates its use in learning approximately optimal power allocation policies over the sampled dual function as the channel evolves over time. Conditions are established under which Newton's method learns approximate solutions with a single update, and the subsequent suboptimality of the control problem is further characterized. Numerical simulations illustrate the near-optimal performance of the method and resulting stability on a wireless control problem.**

Index Terms—**Wireless control systems, learning, Newton's method, nonstationary channel.**

I. INTRODUCTION

THE recent developments in autonomy in industrial control environments, teams of robotic vehicles, and the Internet-of-Things have motivated intelligent design of wireless systems. Even though wireless communication facilitates connectivity, it also introduces uncertainty that may affect stability and performance. To guarantee performance and safety of the control application it is common to employ model-based approaches. However wireless communication is naturally uncertain and time-varying due to effects that are not always amenable to modeling, such as mobility in the environment. In this paper we propose an alternative learning-based approach, where au-

tonomy relies on collected channel samples to optimize performance in a non-stationary environment. The connection between the two approaches is based on the observation that a sampled version of the model-based design approach can be cast as an empirical risk minimization (ERM) problem, a typical machine learning problem. Even so, standard techniques developed for solving ERM problems in machine learning do not address the additional challenges present in wireless autonomous systems, namely the non-stationarity of sample distributions.

The traditional model-based approach is motivated by the desire to build wireless control systems with stability and optimal performance. To counteract channel uncertainties it is natural to include a model of the wireless communication, for example an i.i.d. or Markov link quality, alongside the model of the physical system to be controlled. These models have been valuable to help analysis and control/communication design. For example, one can characterize that it is impossible to estimate and/or stabilize an unstable plant if its growth rate is larger than the rate at which the link drops packets [3]–[6], or below a certain channel capacity [7], [8]. Additionally models facilitate the design of controllers [9]–[11], as well as the allocation of communication resources to optimize control performance, for example power allocation over fading channels with known distributions [12], [13], or event-triggered control [14]–[18].

In practice wireless autonomous systems operate under unpredictable channel conditions following unknown time-varying distributions. While one approach would be to estimate the distributions using channel samples and then follow the above model-based design approach, in this paper we propose an alternative learning-based approach which bypasses the channel-modeling phase. We exploit channel samples taken from the time-varying channel distributions with the goal to learn directly the solution to communication design problems. To apply this approach we exploit a connection between the model-based and the learning-based design problems. Existing works [19]–[21] study related problems in multiple-access wireless control systems and resource allocation problems in wireless systems but under a stationary channel distribution. These works generally employ first-order stochastic methods, which have slow convergence rates and hence not suitable for the present framework. A significant challenge remains in how to continuously learn optimal policies over a wireless channel that is time-varying. This shortcoming of existing sample-based approaches used in [19]–[21] and more general machine learning scenarios motivates the higher-order learning approach proposed in this paper. Some existing machine learning methods account for nonstationarity

by optimizing an averaged objective over all time [22]–[24]. Our approach differs in that we seek and track optimality locally with respect to the current channel distribution at every time epoch.

In this paper we consider a wireless autonomous system where the design goal is to maximize a level of control performance for multiple systems while meeting a desired transmit power budget over the wireless channel (Section II). The wireless channel is modeled as a fading channel with a time-varying and unknown distribution, and only available through samples taken over time. We derive in Section II-A a wireless control problem that finds optimal power allocation policies for an individual time epoch where the wireless channel distribution does not change, and then proceed to derive the Lagrange dual (Section II-B). We show in Section III that the dual of the power allocation problem can be rewritten using channel samples as an empirical risk minimization problem, a common machine learning problem in which an expected loss function over an unknown distribution is approximated by optimized over a set of samples. Here the risk is loosely related to how far the current solution is from the desired optimal power allocation.

Because the wireless channel is varying over time, we develop a new approach to solving a sequence of ERM problems. We collect and store a window of channel samples taken from consecutive distributions to reduce sampling complexity and employ Newton's method to learn new policies quickly (Section IV). More specifically, the quadratic convergence rate of Newton's method is shown to be sufficient to find approximate solutions to slowly varying objectives with a single update. Using Newton's method, we propose an algorithm that uses channel samples to approximate the solution of a power allocation wireless control problem over a non-stationary channel. We prove that, under specific conditions, the algorithm reaches an approximately optimal point in a single iteration of Newton's method (Section V). This result establishes both a suboptimality bound with respect to the sampled problem (Section V-A) as well as with respect to control performance metric in the wireless control problem (Section V-B). We additionally show a stability result for a particular problem description common in wireless control systems (Section V-C) and provide considerations for practical implementation of the method (Section VI). These results are further demonstrated in a numerical demonstration of learning power allocation policies across multiple control systems over a time-varying channel (Section VII).

II. WIRELESS CONTROL PROBLEM

We consider a wireless control problem (WCP) with m independent control systems labeled $i = 1, \ldots, m$, as shown in Fig. 1. Each control system/agent i communicates at time t its state x_t^i over a wireless channel in order to close a loop and maximize a level of control performance. In particular, system i tries to close the control loop over the wireless channel by transmitting with power level $p^i \in [0, p_0]$. Due to propagation effects the channel fading conditions that each system i experiences, denoted by $h^i \in \mathbb{R}_+$, change unpredictably over time [25, Ch. 3]. Together, the channel fading h^i and transmit power p^i deter-
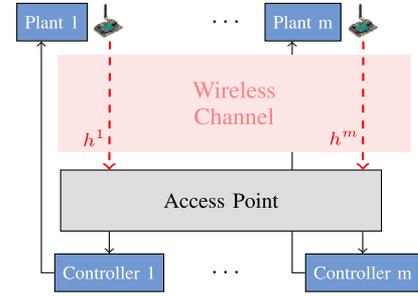

Fig. 1. Wireless control system. Plants communicate state information to access point/controllers over wireless medium.

mine the signal-to-noise ratio (SNR) at the receiver for system i, which in turn affects the probability of successful decoding of the transmitted packet at the receiver. We consider a function $q(h^i, p^i)$ that, given a current channel state and transmit power, determines the probability of successful transmission and decoding of the transmitted packet – see, e.g., [12], [13] for more details on this model. Transmission are assumed on different frequencies/bands and are not subject to contention – see [19], [20] for alternative formulations.

Because these fading conditions vary quickly and unpredictably, they can be modeled as independent random variables drawn from distribution $\mathcal{H}$ that itself is non-stationary, or time-varying. Channel fading is assumed constant during each transmission slot and it is independently distributed over time slots (block fading). Furthermore, the channel distribution $\mathcal{H}$ may vary across *time epochs*, but will in general be stationary within a single time epoch. In particular, consider an epoch index $k = 0, 1, \ldots$ that specifies a particular channel distribution $\mathcal{H}_k$ with realization h_k^i for system i. In Fig. 2, we display a time axis rendering of this model. The state variables change at each transmission slot t, while the channel changes at scale k, which will in general contain multiple time steps. This is to say that we assume that the channel distribution $\mathcal{H}_k$ changes at a rate slower than the system evolution, and that within a single time epoch the channel is effectively stationary.

We proceed to derive a formal description of the wireless control problem of interest within a single time epoch, where the channel is assumed stationary. In Section IV we extend this formulation to the non-stationary setting.

A. WCP in Single Epoch

Within a particular time epoch k with channel distribution $\mathcal{H}_k$, we can derive a formulation that characterizes the optimal power allocations between the m control systems so as to maximize the aggregate control performance across all systems, where p_0 reflects a maximum transmission power of the system. Given a random channel state $h_k^i \in \mathbb{R}_+$ drawn from the distribution $\mathcal{H}_k$. We wish to determine the amount of transmit power $p_k^i(h_k^i) : \mathbb{R}_+ \rightarrow [0, p_0]$ to be used when attempting to close its loop— see [12] for details. We note that we are looking for transmit power as a function of current channel conditions, as the power necessary to close the loop will indeed change with channel conditions. We assume the current channel gain h_k^i is available
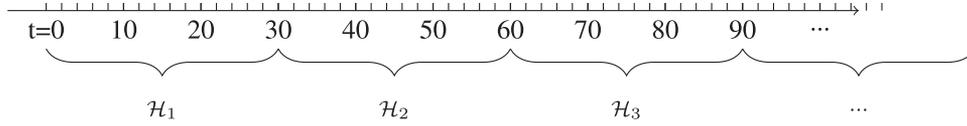
Fig. 2. Time axis showing evolution of time t and epochs k. Each channel distribution $\mathcal{H}_k$ is stationary for a set of time instances.

at the transmitter at each slot, as this can generally be obtained via short pilot signals—see [12]. Then the probability of closing the loop is given by the value

$$y_k^i := \mathbb{E}_{h_k^i} \left\{ q(h_k^i, p_k^i(h_k^i)) \right\}. \tag{1}$$

The variable $y_k^i \in [0, 1]$ is the expectation of successful transmission over the channel distribution $\mathcal{H}_k$.

Using the variable y_k^i we use a monotonically increasing concave function $J^i : [0, 1] \to \mathbb{R}$ that returns a measure of control system performance as a function of the probability of successful transmission. Such a function can take on many forms and, in general, can be derived in relation to the particular control task of interest. In the following example, we derive such a measure for a typical wireless control problem setting, namely the quadratic control performance of a switched linear dynamical system – see, e.g., [4], [6].

Example 1: Consider for example that a control system i is a scalar linear dynamical system of the form

$$x_{t+1}^i = A_o^i x_t^i + B^i u_t^i + w_t^i \tag{2}$$

where $x_t^i \in \mathbb{R}$ is the state of the system at transmission time t, A_o^i is the open loop (potentially unstable) dynamics of the system, $u_t^i \in \mathbb{R}$ is the control input applied to the system at time t, and w_t^i is some zero-mean i.i.d. disturbance process with variance W^i. Consider a given linear state feedback is applied to the system as the control input when a transmission is successful, i.e.,

$$u_t^i = \begin{cases} K^i x_t^i & \text{if loop closes} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

As a result, the system switched between an open loop mode A_o^i and a closed loop stable mode $A_c^i = A_o^i + B^i K^i$, as in

$$x_{t+1}^i = \begin{cases} A_c^i x_t^i + w_t^i & \text{if loop closes} \\ A_o^i x_t^i + w_t^i & \text{otherwise} \end{cases} \tag{4}$$

The goal is to regulate the system state close to zero, i.e., the system attempts to close the loop at a high rate in order to minimize an expected quadratic control cost objective of the form

$$\lim_{N \to \infty} \frac{1}{N} \sum_{t=0}^{N-1} \mathbb{E}(x_t^i)^2 \tag{5}$$

Assuming the control loop in (4) is closed with the success probability y_k^i in (1) *at all time steps*, it is possible to express the above cost explicitly as a function of y_k^i. Using the system dynamics (4), the variance of the system state satisfies the recursive formula

$$\mathbb{E}(x_{t+1}^i)^2 = y_k^i (A_c^i)^2 \mathbb{E}(x_t^i)^2 + (1 - y_k^i)(A_o^i)^2 \mathbb{E}(x_t^i)^2 + W^i \tag{6}$$

that is, with probability y_k^i the variance grows according to the open loop dynamics, and with probability $1 - y_k^i$ the variance shrinks according to the closed loop stable dynamics.

Operating recursively and using the geometric series sum, we can rewrite the variance at time t as

$$\mathbb{E}(x_t^i)^2 = [y_k^i (A_c^i)^2 + (1 - y_k^i)(A_o^i)^2]^t \mathbb{E}(x_0^i)^2 \tag{7}$$

$$+ W^i \frac{1 - [y_k^i (A_c^i)^2 + (1 - y_k^i)(A_o^i)^2]^t}{1 - [y_k^i (A_c^i)^2 + (1 - y_k^i)(A_o^i)^2]}. \tag{8}$$

As follows from the above expression, the system is stable, i.e., the variance is bounded, if the packet success rate satisfies $[y_k^i (A_c^i)^2 + (1 - y_k^i)(A_o^i)^2] < 1$ so that the sum above is bounded – see also [4], [6]. In that case, the state variance as well as the average (5) converge to the same limit value, which we can define as our control performance function

$$J^i(y_k^i) = -\frac{W^i}{1 - [y_k^i (A_c^i)^2 + (1 - y_k^i)(A_o^i)^2]} \tag{9}$$

This control performance function satisfies the assumption of concavity, and it is also monotonically increasing because we have added the negative sign in front of the expression. It is also possible to extend this analysis to include a cost on the control input, as is common in the Linear Quadratic Control problem, i.e., replace the cost in (5) with $\mathbb{E}(x_t^i)^2 + (u_t^i)^2$.

Remark 1: In Example 1, observe that the control system performance in (5) is a long term objective asymptotically for $t \to \infty$. As the channel fading distribution $\mathcal{H}_k$ will change unpredictably in the future it is hard to define an accurate value of this control performance. As a surrogate, in the above example we write a control system performance in (9) with respect to the current channel distribution $\mathcal{H}_k$, i.e., as if this channel distribution is stationary and will not change in the future. Later, in Section V-C we argue that this approximation and the power allocation algorithm we develop can indeed guarantee system stability.

To derive the full formulation of the wireless control problem for current channel distribution $\mathcal{H}_k$, we first define using boldface vectors the set of m channel states $\mathbf{h}_k := [h_k^1; h_k^2; \ldots; h_k^m] \sim \mathcal{H}_k^m$ observed by the control systems and the set of power allocation policies $\mathbf{p}_k(\mathbf{h}_k) := [p_k^1(h_k^1); p_k^2(h_k^2); \ldots; p_k^m(h_k^m)] \in \mathcal{P} := [0, p_0]^m$. We further define the vector of transmission probabilities at specific channel states $\mathbf{q}(\mathbf{h}_k, \mathbf{p}_k(\mathbf{h}_k)) := [q(h_k^1, p_k^1(h_k^1)); \ldots; q(h_k^m, p_k^m(h_k^m))]$ and expected transmission probabilities $\mathbf{y}_k := [y_k^1; y_k^2; \ldots; y_k^m]$ from (1). The goal is to select $\mathbf{p}_k(\mathbf{h}_k)$ whose expected aggregate value is within a maximum power budget $p_{\max}$ while maximizing the total system performance $\sum J^i$ over m agents. Because J^i is monotonically increasing, we can relax the equality in (1) to an inequality

constraint and write the following optimization problem.

$$\{\mathbf{p}_k^*(h), \mathbf{y}_k^*\} := \operatorname*{argmax}_{\mathbf{p}_k \in \mathcal{P}, \mathbf{y}_k \in \mathbb{R}^m} \ J(\mathbf{y}_k) := \sum_{i=1}^{m} J^i(y_k^i) \quad (\text{WCP}_k)$$

$$\text{s.t.} \quad \mathbf{y}_k \leq \mathbb{E}_{\mathbf{h}_k} \left\{ \mathbf{q}(\mathbf{h}_k, \mathbf{p}_k(\mathbf{h}_k)) \right\},$$

$$\sum_{i=1}^{m} \mathbb{E}_{h_k^i} \left(p_k^i(h_k^i) \right) \leq p_{\max}$$

The problem in (WCP_k) states the optimal power allocation policy $\mathbf{p}_k^*(\mathbf{h}_k)$ is the one that maximizes the expected aggregate control performance over channel states while guaranteeing that the expected total transmitting power is below an available budget $p_{\max}$. We stress that this only provides the optimal policy with respect to a particular channel distribution $\mathcal{H}_k$. In the non-stationary wireless setting we are interested in solving (WCP_k) for all k.

B. Dual Formulation of (WCP_k)

Solving this optimization problem directly has a number of significant challenges. The first is that the problem is non-convex, in particular due to the first constraint in (WCP_k). The second challenge is that the problem is optimized over an infinite-dimensional variable $\mathbf{p}_k(\mathbf{h}_k)$. It is very difficult to solve such a problem if there is no assumed parameterization of $\mathbf{p}_k^*(\mathbf{h}_k)$. We can show, however, from a result in [26] that a naturally occurring parameterization of $\mathbf{p}_k^*(\mathbf{h}_k)$ indeed can be derived from Lagrangian duality theory.

We proceed then to derive the dual problem from the constrained problem in (WCP_k). To simplify the presentation, we first introduce a set of augmented variables, denoted with tildes. Define the augmented vectors $\check{\mathbf{q}}(\mathbf{h}_k, \mathbf{p}_k(\mathbf{h}_k)) \in \mathbb{R}^{m+1}$ and $\check{\mathbf{y}}_k \in \mathbb{R}^{m+1}$ as

$$\check{\mathbf{q}}(\mathbf{h}_k, \mathbf{p}_k(\mathbf{h}_k)) := \begin{bmatrix} q(h_k^1, p_k^1(h_k^1)) \\ \vdots \\ q(h_k^m, p_k^m(h_k^m)) \\ -\sum_{i=1}^{m} p_k^i(h_k^i)] \end{bmatrix} \quad \check{\mathbf{y}}_k := \begin{bmatrix} y_k^1 \\ \vdots \\ y_k^m \\ -p_{\max} \end{bmatrix}. \quad (10)$$

The augmented $\check{\mathbf{q}}(\mathbf{h}_k, \mathbf{p}_k(\mathbf{h}_k))$ includes transmission probabilities augmented with the total power allocation while $\check{\mathbf{y}}_k$ includes auxiliary variables augmented with total power budget. Using this new notation, the Lagrangian function is formed as

$$\mathcal{L}_k(\mathbf{p}_k(\mathbf{h}_k), \mathbf{y}_k, \boldsymbol{\mu}_k) := \sum_{i=1}^{m} J^i(y_k^i)$$

$$+ \boldsymbol{\mu}_k^T \left(\mathbb{E}_{\mathbf{h}_k} \check{\mathbf{q}}(\mathbf{h}_k, \mathbf{p}_k(\mathbf{h}_k)) - \check{\mathbf{y}}_k \right), \quad (11)$$

where $\boldsymbol{\mu}_k := [\mu_k^1; \ldots; \mu_k^m; \tilde{\mu}] \in \mathbb{R}_+^{m+1}$ contains the dual variables associated with each of the $m+1$ constraints in (WCP_k). From the Lagrangian function in (11), the Lagrangian dual loss function is defined as $L_k(\boldsymbol{\mu}_k) := \max_{\mathbf{p}_k, \mathbf{y}_k} \mathcal{L}_k(\mathbf{p}_k(\mathbf{h}_k), \mathbf{y}_k, \boldsymbol{\mu}_k)$—see, e.g., [27]—and the corresponding dual

problem as

$$\tilde{\boldsymbol{\mu}}_k^* := \operatorname*{argmin}_{\boldsymbol{\mu}_k \geq 0} L_k(\boldsymbol{\mu}_k)$$

$$L_k(\boldsymbol{\mu}_k) := \mathbb{E}_{\mathbf{h}_k} \left\{ \max_{\mathbf{p}_k, \mathbf{y}_k} \sum_{i=1}^{m} J^i(y_k^i) + \boldsymbol{\mu}_k^T (\check{\mathbf{q}}(\mathbf{h}_k, \mathbf{p}_k(\mathbf{h}_k)) - \check{\mathbf{y}}_k) \right\}. \quad (12)$$

Note in (12) that the expectation operator and maximization were exchanged without loss of generality—see, e.g., [19, Proposition 2]. It is important to stress here the connection between the dual problem in (12) with the original problem in (WCP_k). While (WCP_k) is indeed not convex, problems of this form can be shown to exhibit zero duality gap under the technical assumption that the primal problem is strictly feasible and that the channel probability distribution is non-atomic [26]. This implies that the optimal primal variable $\mathbf{p}_k^*(\mathbf{h}_k)$ in (WCP_k) can be recovered from the optimal dual variable $\tilde{\boldsymbol{\mu}}_k^*$ in (12). Thus, the power allocation policy for each agent i is found indirectly by solving (12) and recovering as

$$p_k^i(h_k^i, \boldsymbol{\mu}_k) = \operatorname*{argmax}_{p_k^i \in [0, p_0]} \mu_k^i q(h_k^i, p_k^i(h_k^i)) - \tilde{\mu} p_k^i(h_k^i), \quad (13)$$

$$y_k^i(\boldsymbol{\mu}_k) = \operatorname*{argmax}_{y_k^i} J^i(y_k^i) - \mu_k^i y_k^i. \quad (14)$$

The optimal policy is subsequently recovered using the optimal dual variable as $\mathbf{p}_k^*(\mathbf{h}_k) := [p_k^1(h_k^1, \tilde{\boldsymbol{\mu}}_k^*); \ldots; p_k^m(h_k^m, \tilde{\boldsymbol{\mu}}_k^*)]$. Observe that the problem in (12) is a simply constrained stochastic problem that is known to always be convex from duality theory, and can be solved efficiently with a variety of projected stochastic descent methods [19], [20], [28]–[30]. Thus, the non-convex, infinite-dimensional optimization problem in (WCP_k) can be solved indirectly but exactly with the convex, finite-dimensional problem in (12).

Remark 2: The problem formulation given in (WCP_k) that we use in this paper assumes there is a fixed power budget and the metric to be optimized is a measure of control performance. An alternative formulation of resource allocation that may be more relevant in some settings would instead fix a bound on the required control performance, typically derived from a stability margin for the control system. Here the objective would instead be to minimize total power usage, subject to the constraint on control performance. Indeed, these two problems are very similar when reformulated in the dual domain, and can thus be studied almost identically as such. We specifically focus on the problem in (WCP_k) in this paper but stress that all the results will apply to this alternative problem as well.

III. ERM FORMULATION OF (WCP_k)

The stochastic program in (12) features an objective that is the expectation taken over a random variable, and can thus be considered as a particular case of the empirical risk minimization (ERM) problem. Empirical risk minimization is a common problem studied in machine learning due to its ubiquity in training classifiers, and the same structure appears naturally in the dual formulation of the WCP. A generic ERM

problem considers a convex loss function $f(\boldsymbol{\mu}_k, \mathbf{h}_k)$ of a decision variable $\boldsymbol{\mu}_k \in \mathbb{R}^{m+1}$ and random variable $\mathbf{h}_k$ drawn from distribution $\mathcal{H}_k$ and seeks to minimize the expected loss $L_k(\boldsymbol{\mu}_k) := \mathbb{E}_{\mathbf{h}_k}[f(\boldsymbol{\mu}_k, \mathbf{h}_k)]$. For the WCP in (WCP_k), we rewrite the loss function L and associated ERM problem in terms of a function $f(\boldsymbol{\mu}_k, \mathbf{h}_k)$ using its dual as

$$\tilde{\boldsymbol{\mu}}_k^* := \underset{\boldsymbol{\mu}_k \geq 0}{\operatorname{argmin}} \, L_k(\boldsymbol{\mu}_k) := \underset{\boldsymbol{\mu}_k \geq 0}{\operatorname{argmin}} \, \mathbb{E}_{\mathbf{h}_k} f(\boldsymbol{\mu}_k, \mathbf{h}_k),$$

$$f(\boldsymbol{\mu}_k, \mathbf{h}_k) := J(\mathbf{y}_k(\boldsymbol{\mu}_k)) + \boldsymbol{\mu}_k^T(\check{\mathbf{q}}(\mathbf{h}_k, \mathbf{p}_k(\mathbf{h}_k, \boldsymbol{\mu}_k)) - \check{\mathbf{y}}_k(\boldsymbol{\mu}_k)). \tag{15}$$

Typically the distribution $\mathcal{H}_k$ is not known by the user, so the expected loss cannot be evaluated directly, but is instead replaced by an *empirical risk* by taking n samples labeled $\mathbf{h}_k^1, \mathbf{h}_k^2, \ldots, \mathbf{h}_k^n \in \mathcal{H}_k^m$, (where $\mathbf{h}_k^l := [h_k^{1,l}; \ldots; h_k^{m,l}]$). In practice, such samples can be obtained through the use of short pilot signals sent from the users to measure channel conditions—see [12]. We then consider the empirical average loss function

$$\hat{L}_k(\boldsymbol{\mu}_k) := \frac{1}{n} \sum_{l=1}^{n} f(\boldsymbol{\mu}_k, \mathbf{h}_k^l) := \frac{1}{n} \sum_{l=1}^{n} f_k^l(\boldsymbol{\mu}_k). \tag{16}$$

To characterize the closeness of the empirical risk $\hat{L}_k(\boldsymbol{\mu}_k)$ with n samples with respect to the expected loss $L_k(\boldsymbol{\mu}_k)$, we define a constant V_n called the *statistical accuracy* of $\hat{L}_k$. The statistical accuracy V_n provides a bound of the difference in the empirical and expected loss for all $\boldsymbol{\mu}_k$ with high probability (i.e., at least $1 - \gamma$ for some small γ). In other words, we define V_n to be the constant that satisfies

$$\sup_{\boldsymbol{\mu}_k} |\hat{L}_k(\boldsymbol{\mu}_k) - L_k(\boldsymbol{\mu}_k)| \leq V_n \quad \text{w.h.p.} \tag{17}$$

The upper bounds on V_n are well studied in the learning literature and in general may involve a number of parameters of the loss function f as well as, perhaps most importantly, the number of samples n. For $\hat{L}_k(\boldsymbol{\mu}_k)$ defined in (16), a bound for the statistical accuracy V_n can be obtained in the order of $\mathcal{O}(1/\sqrt{n})$ or, in some cases, $\mathcal{O}(1/n)$ [31], [32]. This further implies a suboptimality of $\hat{L}_k^* := \min \hat{L}_k(\boldsymbol{\mu}_k)$ of the same accuracy, i.e., $|L_k^* - \hat{L}_k^*| \leq 2V_n$ [31].

As is often the case in machine learning problems, the statistical accuracy informs the proper use of regularization terms in the empirical loss function. We can add regularizations to prescribe desirable properties on the empirical risk $\hat{L}_k(\boldsymbol{\mu}_k)$, such as strong convexity, without adding additional bias beyond that already accrued by the empirical approximation. In other words, as $\hat{L}_k^*$ will be of order V_n from the optimal expected value L^*, any additional bias of order V_n or less is permissible. With that in mind, we add the regularization term $\alpha V_n/2 \|\boldsymbol{\mu}_k\|^2$ where $\alpha > 0$ to the empirical risk in (16) to impose strong convexity. We can further remove the non-negativity constraint on the dual variables in (15) through the use of a logarithmic barrier. To preserve smoothness for small $\boldsymbol{\mu}_k$, we use an ϵ-thresholded log function, defined as

$$\log_\epsilon(\boldsymbol{\mu}_k) := \begin{cases} \log(\boldsymbol{\mu}_k) & \boldsymbol{\mu}_k \geq \epsilon \\ \ell_{2,\epsilon}(\boldsymbol{\mu}_k - \epsilon) & \boldsymbol{\mu}_k < \epsilon, \end{cases} \tag{18}$$

where $\ell_{2,\epsilon}(\boldsymbol{\mu}_k)$ is a second order Taylor series expansion of $\log(\boldsymbol{\mu}_k)$ centered at ϵ for some small $0 < \epsilon < 1$. We then use $-\beta V_n \mathbf{1}^T \log_\epsilon \boldsymbol{\mu}_k$ where $\beta > 0$ as a second regularization term, and obtain a regularized empirical risk function

$$R_k(\boldsymbol{\mu}_k) := \frac{1}{n} \sum_{l=1}^{n} f_k^l(\boldsymbol{\mu}_k) + \frac{\alpha V_n}{2} \|\boldsymbol{\mu}_k\|^2 - \beta V_n \mathbf{1}^T \log_\epsilon \boldsymbol{\mu}_k. \tag{19}$$

From here, we can seek a minimizer of the strongly convex regularized risk $R_k(\boldsymbol{\mu}_k)$ without explicitly enforcing a non-negativity constraint on $\boldsymbol{\mu}_k$ and find a solution with suboptimality of order $\mathcal{O}(V_n)$ with respect to (15). Such a deterministic and strongly convex loss function as in (19) can be minimized using a wide array of optimization methods [30], [33]–[35]. However, all such methods only solve the problem for a particular epoch k, or otherwise assume a stationary channel distribution $\mathcal{H}_k$ as is typical in machine learning settings.

IV. ERM OVER NONSTATIONARY CHANNEL

The ERM problem we are interested in solving in wireless autonomous systems is further complicated by the non-stationarity of $\mathcal{H}$, making existing solution methods insufficient. This is due to the fact that finding the minimizer to $R_k(\boldsymbol{\mu})$ will only provide an optimal power allocation for the respective channel distribution $\mathcal{H}_k$. In wireless systems, we instead must *continuously learn* optimal policies as the channel varies, or in other words, find optimal points for $R_k(\boldsymbol{\mu})$ for $k = 0, 1, \ldots$. To formulate the non-stationarity, however, we first define an epoch-indexed empirical risk function. While we may use a simple empirical risk as we did in (16), we instead define a more general statistical loss function for a non-stationary channel using samples from the previous M epochs. We define a windowed empirical loss function $\tilde{L}_k(\boldsymbol{\mu})$ at epoch k as

$$\tilde{L}_k(\boldsymbol{\mu}) := \frac{1}{M} \sum_{j=k-M+1}^{k} \hat{L}_j(\boldsymbol{\mu}) \tag{20}$$

By keeping a window of samples, we may retain $N = Mn$ total samples while drawing only n new samples at each epoch. If the successive channel distributions $\mathcal{H}_{k-M+1}, \ldots, \mathcal{H}_k$ are not very different, we may expect the old channel samples to still be of interest. We define the associated statistical accuracy $\tilde{V}_N$ as the constant that satisfies

$$\sup_{\boldsymbol{\mu}} |\tilde{L}_k(\boldsymbol{\mu}) - L_k(\boldsymbol{\mu})| \leq \tilde{V}_N \quad \text{w.h.p.} \tag{21}$$

Here we stress that the bounds on this constant $\tilde{V}_N$ are not as easily obtainable or well-studied as in the stationary setting. Such a bound over non-i.i.d. samples may be dependent upon many parameters such as the sample batch size n, window size M, and correlation between successive distributions $\mathcal{H}_j$ and $\mathcal{H}_{j+1}$. Therefore, finding precise bounds on $\tilde{V}_N$ would require a sophisticated statistical analysis and is outside the scope of this work. We instead define a user-selected accuracy $\hat{V}$ that may estimate the statistical accuracy $\tilde{V}_N$. We assume that $\hat{V} \geq \tilde{V}_N$, with equality holding in cases where $\tilde{V}_N$ is known. Using the same regularizations introduced previously, we obtain the

regularized windowed empirical loss function

$$\tilde{R}_k(\boldsymbol{\mu}) := \frac{1}{M} \sum_{j=k-M+1}^{k} \hat{L}_j(\boldsymbol{\mu}) + \frac{\alpha\hat{V}}{2}\|\boldsymbol{\mu}\|^2 - \beta\hat{V}\mathbf{1}^T \log_\epsilon \boldsymbol{\mu}. \tag{22}$$

We subsequently define $\tilde{\boldsymbol{\mu}}_k^* := \operatorname{argmin}_{\boldsymbol{\mu}} \tilde{R}_k(\boldsymbol{\mu})$. The definition of the loss function in (22) includes the batches of n samples taken from the previous M channel distributions $\mathcal{H}_{k-M+1}$, $\ldots, \mathcal{H}_k$. This definition is, in a sense, a generalization of the simpler empirical risk $R_k(\boldsymbol{\mu})$ in (19). Observe that, by using a window size of $M = 1$, we use only samples from the current channel and recover $R_k(\boldsymbol{\mu})$. In the following proposition we establish the accuracy of an optimal point of our regularized empirical risk function $\tilde{R}_k(\boldsymbol{\mu})$ relative to the optimal point of the original dual loss function $L_k(\boldsymbol{\mu})$.

Proposition 1: Consider $L_k^* = L_k(\boldsymbol{\mu}_k^*)$ and $\tilde{L}_k^* = \min_{\boldsymbol{\mu} \geq 0} \tilde{L}_k(\boldsymbol{\mu})$, and define $\tilde{R}_k^* := \min_{\boldsymbol{\mu}} \tilde{L}_k(\boldsymbol{\mu}) + \alpha\hat{V}/2\|\boldsymbol{\mu}\|^2 - \beta\hat{V}\mathbf{1}^T \log \boldsymbol{\mu}$ as the optimal value of the regularized empirical risk. Define $\tilde{V}_N$ by (21). Assuming $\tilde{V}_N \leq \hat{V}$, the difference $|L_k^* - \tilde{R}_k^*|$ is upper bounded on the order of statistical accuracy $\hat{V}$, i.e., for some $\rho > 0$

$$|L_k^* - \tilde{R}_k^*| \leq 2\tilde{V}_N + \rho\hat{V} \leq (2+\rho)\hat{V}, \quad \text{w.h.p.} \tag{23}$$

Proof: To obtain the result in (23), consider expanding and upper bounding $|L_k^* - \tilde{R}^*| = |L_k^* - \tilde{L}_k^* + \tilde{L}_k^* + \tilde{R}_k^*| \leq |L_k^* - \tilde{L}_k^*| + |\tilde{L}_k^* + \tilde{R}_k^*|$. The first term is bounded by $2\tilde{V}_N$ as previously discussed. The second term, can be decomposed into the bias introduced by the logarithmic barrier $-\beta\hat{V}\mathbf{1}^T \log \boldsymbol{\mu}$ and the bias introduced by the quadratic regularizer $c\hat{V}/2\|\boldsymbol{\mu}\|^2$. The former of these is known to produce an optimality bias of $(m+1)\beta\hat{V}$ [27, Section 11.2.2], while the latter is known to introduce a bias on the order of $\mathcal{O}(\hat{V})$ [36]. Combining these, we get a total suboptimality between the regularized risk function optimal and the true optimal of $2\tilde{V}_N + \rho\hat{V}$ for some constant $\rho > 0$. As we assume that $\tilde{V}_N \leq \hat{V}$, the rightmost bound in (23) follows. ∎

A key observation to be made here is that any exact solution to (22) only minimizes the expected loss L_k to within accuracy $\hat{V}$ (assuming $\hat{V} \geq \tilde{V}_N$). There is therefore no need to minimize (22) exactly but is in fact sufficient to find a $\hat{V}$-accurate solution, as this incurs no additional error relative to the statistical approximation itself. While many optimization methods can be used to find a minimizer to (22), we demonstrate in the next section that fast second order methods can be used to learn *approximate* minimizers—and by Proposition 1 approximately solve (15)—at each epoch k with just single updates as the channel distribution $\mathcal{H}_k$ changes, thus tracking near-optimal points at every epoch. This is done by exploiting an important property of second order optimization methods, namely *local quadratic convergence*.

Remark 3: Observe in the text of Proposition 1 that we define $\tilde{R}_k^*$ to be the optimal point of the loss function $\tilde{L}_k(\boldsymbol{\mu}_k)$ regularized with a standard log barrier $-\log(\boldsymbol{\mu}_k)$, rather than the thresholded barrier $-\log_\epsilon(\boldsymbol{\mu}_k)$ used in the definition in (22). Indeed, using the thresholded barrier does not explicitly enforce

nonnegativity for values smaller than ϵ. However, this thresholding is necessary to preserve smoothness of the barrier, which will be necessary for the proof of Lemma 1 in Section V. The threshold ϵ can be made as small as necessary to enforce non-negativity, although this comes at the cost of a worse smoothness constant. In practice, however, we observe this thresholding to not be explicitly needed and is just included here for ease of analysis. We also stress that the smoothness constant itself does not play a pivotal role in the proceeding analysis.

A. Learning via Newton's Method

In this paper, we use Newton's method to approximately minimize (22) efficiently as the channel $\mathcal{H}_k$ changes over epochs. Motivated by the recent use of Newton's method in solving large scale ERM problems through adaptive sampling policies [34], [35], we use the N samples drawn from recent distributions to find an iterate $\boldsymbol{\mu}_k$ that approximately solves for $\tilde{\boldsymbol{\mu}}_k^*$. At the next epoch, the iterate $\boldsymbol{\mu}_k$ provides a "soft" start towards finding a point $\boldsymbol{\mu}_{k+1}$ that approximately minimizes $\tilde{R}_{k+1}(\boldsymbol{\mu})$. In this way, with single iterations we may find near-optimal solutions for each regularized empirical loss function, and thereby efficiently learn the optimal power allocation of the wireless channel as the channel distribution evolves over time epochs.

We proceed by presenting the details of Newton's method. At epoch k, we compute a new iterate $\boldsymbol{\mu}_{k+1}$ by subtracting from the current iterate $\boldsymbol{\mu}_k$ the product of the Hessian inverse and the gradient of the function $\tilde{R}_{k+1}(\boldsymbol{\mu}_k)$. For the empirical dual loss function $\tilde{R}_k$ defined in (22), we define the gradient $\nabla\tilde{R}_k(\boldsymbol{\mu})$ and Hessian $\nabla^2\tilde{R}_k(\boldsymbol{\mu})$. The new approximate solution $\boldsymbol{\mu}_{k+1}$ is then found from current approximate solution $\boldsymbol{\mu}_k$ using the Newton update

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k - \mathbf{H}_{k+1}^{-1}\nabla\tilde{R}_{k+1}(\boldsymbol{\mu}_k), \tag{24}$$

where we use $\mathbf{H}_{k+1} := \nabla^2\tilde{R}_{k+1}(\boldsymbol{\mu}_k)$ as simplified notation.

To understand the full algorithm, consider that $\boldsymbol{\mu}_k$ is a $\hat{V}$-accurate solution of current loss function $\tilde{R}_k$, i.e., $\tilde{R}_k(\boldsymbol{\mu}_k) - \tilde{R}_k^* \leq \hat{V}$. Recall that the new loss function $\tilde{R}_{k+1}$ differs from $\tilde{R}_k$ only in the discarding of old samples $\hat{L}_{k-M+1}$ and inclusion of samples $\hat{L}_{k+1}$ drawn from $\mathcal{H}_{k+1}$. If we consider that the distributions are varying slowly across successive time epochs, i.e., $\mathcal{H}_{k+1}$ is close to $\mathcal{H}_k$, then the respective loss functions $\tilde{R}_{k+1}$ and $\tilde{R}_k$ and their optimal values $\tilde{R}_{k+1}^*$ and $\tilde{R}_k^*$ will also not differ greatly under some smoothness assumptions. Therefore, under such conditions a single step of Newton's method as performed in (24) can in fact be sufficient to reach a $\hat{V}$-accurate solution of the new loss function $\tilde{R}_{k+1}$. This is possible precisely because of the Newton method's property of *local quadratic convergence*, meaning that Newton's method will find a near-optimal solution very quickly when it is already in a local neighborhood of the optimal point. Given then a $\hat{V}$-accurate solution $\boldsymbol{\mu}_0$ of initial loss $\tilde{R}_0$, the proceeding and all subsequent iterates $\boldsymbol{\mu}_k$ will remain within the statistical accuracy of their respective losses $\tilde{R}_k$ as the channel distribution varies over time. The formal presentation of the exploitation of this property and other technical details of this result are discussed in Section V of this paper.

Algorithm 1: Learning via Newton's Method.

1: **Parameters:** Sample size increase constants $n_0 > 0$, $M_0 \geq 1$ backtracking params $0 < \gamma < 1 < \Gamma$, and accuracy $\hat{V}$.

2: **Input:** Initial sample size $n = m_0$ and argument $\boldsymbol{\mu}_n = \boldsymbol{\mu}_{m_0}$ with $\|\nabla \tilde{R}_n(\boldsymbol{\mu}_{k+1})\| < (\sqrt{2\alpha})\hat{V}$

3: **for** $k = 0, 1, 2, \ldots$ **do** {main loop}

4:     Reset factor $n = n_0$, $M = M_0$.

5:     **repeat** {sample size backtracking loop}

6:         Draw n samples from $\mathcal{H}_k$, discard from $\mathcal{H}_{k-M}$.

7:         Compute Gradient $\nabla \tilde{R}_k(\boldsymbol{\mu}_{k-1})$, Hessian $\mathbf{H}_k$.

8:         Newton Update [cf. (24)]:

$$\boldsymbol{\mu}_k = \boldsymbol{\mu}_{k-1} - \mathbf{H}_k^{-1} \nabla \tilde{R}_k(\boldsymbol{\mu}_{k-1})$$

9:         Determine power allocation, aux. variables [cf. (13), (14)]:

$$p_k^i(h_k^i, \boldsymbol{\mu}_k) = \underset{p_k^i \in [0, p_0]}{\operatorname{argmax}} \mu_k^i q(h_k^i, p_k^i(h_k^i)) - \tilde{\mu} p_k^i(h_k^i),$$

$$y_k^i(\boldsymbol{\mu}_k) = \underset{y_k^i}{\operatorname{argmax}} J^i(y_k^i) - \mu_k^i y_k^i.$$

10:      Backtrack sample draw $n = \Gamma n$, window size $M = \gamma M$.

11:     **until** $\|\nabla \tilde{R}_k(\boldsymbol{\mu}_k)\| < (\sqrt{2\alpha})\hat{V}$

12: **end for**

The learning algorithm is presented in Algorithm 1. After preliminaries and initializations in Steps 1–4, the backtracking loop starts in Step 5. Each iteration begins in Step 6 with the the drawing of n samples from the new channel distribution $\mathcal{H}_k$ and discarding of old samples from $\mathcal{H}_{k-M}$ to form $\tilde{R}_k$. Note that samples will be only be discarded for $k > M$. The gradient $\nabla \tilde{R}_k$ and Hessian $\mathbf{H}_k$ of the regularized dual loss function are computed in Step 7. The Newton step is taken with respect to $\tilde{R}_{k+1}$ in Step 8. In Step 9, the optimal primal variables are computed with respect to the updated dual variables. This includes both the auxiliary variables $\mathbf{y}(\boldsymbol{\mu}_k)$ and the power allocation policy $\mathbf{p}(\mathbf{h}, \boldsymbol{\mu}_k)$ itself. Because there are function and channel system parameters that are not known in practice, we include a backtracking step for the parameters n and M in Step 10 to ensure the new iterate $\boldsymbol{\mu}_k$ is within the intended accuracy $\hat{V}$ of $\boldsymbol{\mu}_k^*$. Further details on the specifics of the backtracking procedure are discussed in Section VI after the presentation of the theoretical results.

V. CONVERGENCE ANALYSIS

In this section we provide a theoretical analysis of the Newton learning update in (24). We do so by first analyzing the convergence properties of the ERM problem in (22). We subsequently return to the WCP in (WCP$_k$) and establish a control performance result.

A. Convergence of ERM Problem

We begin by analyzing the ERM formulation of the power allocation problem in (22) and establish a theoretical result that,

under certain conditions, guarantees each iterate $\boldsymbol{\mu}_k$ is within the statistically accuracy of the risk function at epoch k. Our primary theoretical result characterizes such conditions dependent on statistical accuracy and rate of non-stationarity. We begin by presenting a series of assumptions made in our analysis regarding the dual loss functions f.

Assumption 1: The expected loss function L_k and empirical loss functions $f(\boldsymbol{\mu}, \mathbf{h}_k)$ are convex with respect to $\boldsymbol{\mu}$ for all values of $\mathbf{h}_k$. Moreover, their gradients $\nabla L_k(\boldsymbol{\mu})$ and $\nabla f(\boldsymbol{\mu}, \mathbf{z})$ are Lipschitz continuous with constant Δ.

Assumption 2: The loss functions $f(\boldsymbol{\mu}, \mathbf{h})$ are self-concordant with respect to $\boldsymbol{\mu}$ for all $\mathbf{h}$, i.e., for all i,

$$|\partial^3/\partial \mu_i^3 f(\boldsymbol{\mu}, \mathbf{h})| \leq 2 \partial^2/\partial \mu_i^2 f(\boldsymbol{\mu}, \mathbf{h})^{3/2}.$$

Assumption 1 implies that the regularized empirical risk gradients $\nabla \tilde{R}_k$ are Lipschitz continuous with constant $\Delta + c\hat{V}$ where $c := \alpha + \beta/\epsilon^2$ and α, β, ϵ are the regularization constants in (22). The function $\tilde{R}_k$ is also strongly convex with constant $\alpha\hat{V}$. This implies an upper and lower bound of the eigenvalues of the Hessian of $\tilde{R}_k$, namely

$$\alpha \hat{V} \mathbf{I} \preceq \mathbf{H}_k \preceq (\Delta + c\hat{V})\mathbf{I}. \tag{25}$$

Assumption 2 states the loss functions are additionally self concordant, which is a common assumption made in the analysis of second-order methods—see, e.g., [27, Ch. 9], for such an analysis. It also follows that the functions $\tilde{R}_{k+1}$ are therefore self concordant because both the quadratic and thresholded log regularizers are self-concordant. We present a brief remark regarding the implications of these assumptions on the dual risk function on the wireless control problem.

Remark 4: We state the preceding assumptions in terms of the sampled dual functions f due to their direct use in the proceeding analysis. However, they indeed have implications on the primal domain problem in (WCP$_k$). While the dual function is always convex, the smoothness condition in Assumption 1 can be obtained from the strong concavity of the control performance $\sum_i J^i$ with strong concavity $1/\Delta$. The self-concordance property on the dual function in Assumption 2, however, is not easily derived from properties of $J^i(\cdot)$ or $q(\cdot)$. We point to work that establishes self concordance of the dual for various machine learning problems [37], [38].

The two preceding assumptions deal specifically with the properties of the empirical dual loss functions used in the ERM problem. To connect the solving of the sampled functions f^l with the expected loss function L, we additionally include two assumptions regarding the statistics of the expected and empirical losses.

Assumption 3: The difference between the gradients of the empirical risk $\hat{L}_k$ and the statistical average loss L_k is bounded by $V_N^{1/2}$ for all $\boldsymbol{\mu}$ and k with high probability,

$$\sup_{\boldsymbol{\mu}} \|\nabla L_k(\boldsymbol{\mu}) - \nabla \hat{L}_k(\boldsymbol{\mu})\| \leq V_N^{1/2}, \qquad \text{w.h.p.} \tag{26}$$

Assumption 4: The difference between two successive expected loss functions $L_k(\boldsymbol{\mu}) = \mathbb{E}_{h_k} f(\boldsymbol{\mu}, h_k)$ and $L_{k+1}(\boldsymbol{\mu}) = \mathbb{E}_{h_{k+1}} f(\boldsymbol{\mu}, h_{k+1})$ and the difference between their gradients are bounded respectively by a bounded sequence of constants

$\{D_k\}, \{\bar{D}_k\} \geq 0$ for all $\boldsymbol{\mu}$,

$$\sup_{\boldsymbol{\mu}} |L_k(\boldsymbol{\mu}) - L_{k+1}(\boldsymbol{\mu})| \leq D_k, \tag{27}$$

$$\sup_{\boldsymbol{\mu}} \|\nabla L_k(\boldsymbol{\mu}) - \nabla L_{k+1}(\boldsymbol{\mu})\| \leq \bar{D}_k. \tag{28}$$

Assumption 3 bounds the difference between gradients of the expected loss and the empirical risk with N samples by $V_N^{1/2}$, which can be readily obtained using the law of large numbers. Assumption 4 bounds the point-wise difference in the expected loss functions and their gradients at epochs k and $k+1$. This can be interpreted as the rate at which the channel evolves between epochs, and is used to establish that optimal dual variables for two consecutive empirical risk functions $\tilde{R}_k$ and $\tilde{R}_{k+1}$ are not very different. We discuss practical implications of this assumption in Section VI.

Remark 5: Observe that the bounds provided in Assumption (4) are with respect to the dual function rather than explicitly on the non-stationary statistics of the channel. They are provided as such because this is the manner in which the non-stationarity appears in the proceeding analysis. To see how the channel characteristics play a role in the provided bound, consider that, e.g., (27) can be expanded using the definition of the dual function $L_k(\boldsymbol{\mu})$ as

$$\sup_{\boldsymbol{\mu}} \Big| \mathbb{E}_{\mathbf{h}_k} \Big\{ \max_{\mathbf{p} \in \mathcal{P}} \boldsymbol{\mu}^T \check{\mathbf{q}}(\mathbf{h}_k, \mathbf{p}(\mathbf{h}_k)) \Big\}$$
$$- \mathbb{E}_{\mathbf{h}_{k+1}} \Big\{ \max_{\mathbf{p} \in \mathcal{P}} \boldsymbol{\mu}^T \check{\mathbf{q}}(\mathbf{h}_{k+1}, \mathbf{p}(\mathbf{h}_{k+1})) \Big\} \Big| \leq D_k. \tag{29}$$

The exact condition this imposes upon the channel distribution variation thus depends both on the form of the distributions $\mathcal{H}_k$, $\mathcal{H}_{k+1}$, and the function $q(\mathbf{h}, \mathbf{p})$. Thus, the exact manner in which the varying channel conditions effect this bound are indeed problem-specific, and a generic condition on non-stationarity of the channel is only present in the proceeding analysis indirectly through the condition in (29).

The proceeding analysis is organized in the following manner. Our goal is to establish conditions on the parameters of the statistical accuracy—$\hat{V}$—and the non-stationarity—D_k and $\bar{D}_k$—that guarantee that, starting from an approximate solution to $\tilde{R}_k$, a single step of Newton's method generates an approximately accurate solution to $\tilde{R}_{k+1}$. From there, we can recursively say that, assuming an initial point $\boldsymbol{\mu}_0$ that is within the intended accuracy of $\tilde{R}_0$, the method will continue to find a $\hat{V}$-accurate solution at each epoch as the channel distribution changes with k. We achieve this result in two steps. We first find a condition that guarantees that a $\hat{V}$-accurate solution of $\tilde{R}_k$ is also in the quadratic convergence region of $\tilde{R}_{k+1}$. Second, we find a condition that guarantees that such a point within the quadratic convergence region of $\tilde{R}_{k+1}$ will reach its intended accuracy with a single update as in (24).

We begin by establishing the condition in the first step, namely that a $\hat{V}$-accurate solution to $\tilde{R}_k$, labeled $\boldsymbol{\mu}_k$ is in in the quadratic convergence region of $\tilde{R}_{k+1}$ if certain conditions hold. The quadratic convergence region is a region local to the optimum in which Newton's method is known to converge at a fast quadratic rate. The analysis of Newton's method commonly characterizes quadratic convergence in terms of a quantity called the Newton decrement, explicitly defined as $\lambda_{k+1}(\boldsymbol{\mu}) := \|\nabla^2 \tilde{R}_{k+1}(\boldsymbol{\mu})^{-1/2} \nabla \tilde{R}_{k+1}(\boldsymbol{\mu})\|$. We say the dual iterate $\boldsymbol{\mu}$ is in the quadratic convergence region of $\tilde{R}_{k+1}$ when $\lambda_{k+1}(\boldsymbol{\mu}) < 1/4$—see [27, Chapter 9.6.4]. In the following proposition, we give conditions under which any iterate $\boldsymbol{\mu}_k$ that is within the accuracy $\hat{V}$ of the optimal point $\tilde{R}_k^* = \min_{\boldsymbol{\mu}} \tilde{R}_k(\boldsymbol{\mu})$ is also within the quadratic convergence region of the subsequent loss function $\tilde{R}_{k+1}$.

Lemma 1: Consider $\boldsymbol{\mu}_k$ as a $\hat{V}$-accurate optimal solution of the loss $\tilde{R}_k$, i.e., $\tilde{R}_k(\boldsymbol{\mu}_k) - \tilde{R}_k^* \leq \hat{V}$. In addition, define $\lambda_{k+1}(\boldsymbol{\mu}) := (\nabla \tilde{R}_{k+1}(\boldsymbol{\mu})^T \nabla^2 \tilde{R}_{k+1}(\boldsymbol{\mu})^{-1} \nabla \tilde{R}_{k+1}(\boldsymbol{\mu}))^{1/2}$ as the Newton decrement of variable $\boldsymbol{\mu}$ associated with the loss $\tilde{R}_{k+1}$. If Assumptions 1–4 hold, then Newton's method at point $\boldsymbol{\mu}_k$ is in the quadratic convergence phase for the objective function $\tilde{R}_{k+1}$, i.e., $\lambda_{k+1}(\boldsymbol{\mu}_k) < 1/4$, if we have

$$\left(\frac{2(\Delta + c\hat{V})\hat{V}}{\alpha \hat{V}} \right)^{1/2} + \frac{2\tilde{V}_N^{1/2} + \bar{D}_k}{(\alpha \hat{V})^{1/2}} < \frac{1}{4}. \quad \text{w.h.p.} \tag{30}$$

Proof: See Appendix. ∎

Lemma 1 provides the first necessary condition in our analysis by identifying the statistical parameters under which every iterate $\boldsymbol{\mu}_k$ is in the quadratic region of $\tilde{R}_{k+1}$. From here we can show the second step, in which such a point in the quadratic convergence region of $\tilde{R}_{k+1}$ can reach its statistical accuracy with a single Newton step as given in (24). To achieve this, we first present the following lemma that upper bounds the suboptimality of the point $\boldsymbol{\mu}_k$ with respect to the optimal solution of R_{k+1}^*.

Lemma 2: Consider a point $\boldsymbol{\mu}_k$ that minimizes the loss function $\tilde{R}_k$ to within accuracy $\hat{V}$, i.e., $\tilde{R}_k(\boldsymbol{\mu}_k) - \tilde{R}_k^* \leq \hat{V}$. Provided that Assumptions 1–4 hold, the sub-optimality $\tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \tilde{R}_{k+1}^*$ is upper bounded w.h.p. as

$$\tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \tilde{R}_{k+1}^* \leq 4\tilde{V}_N + \hat{V} + 2D_k \tag{31}$$

Proof: See Appendix. ∎

In Lemma 2 we establish a bound on the suboptimality of $\boldsymbol{\mu}_k$ with respect to $\tilde{R}_{k+1}$. The following lemma now bounds the suboptimality of $\boldsymbol{\mu}_{k+1}$ in terms of the suboptimality of $\boldsymbol{\mu}_k$ with a quadratic rate.

Lemma 3: Consider $\boldsymbol{\mu}_k$ to be in the quadratic neighborhood of the loss $\tilde{R}_{k+1}$, i.e., $\lambda_{k+1}(\boldsymbol{\mu}_k) \leq 1/4$. Recall the definition of the variable $\boldsymbol{\mu}_{k+1}$ in (24) as the updated variable using Newton's method. If Assumptions 1–3 hold, then the difference $\tilde{R}_{k+1}(\boldsymbol{\mu}_{k+1}) - \tilde{R}_{k+1}^*$ is upper bounded by

$$\tilde{R}_{k+1}(\boldsymbol{\mu}_{k+1}) - \tilde{R}_{k+1}^* \leq 144(\tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \tilde{R}_{k+1}^*)^2. \tag{32}$$

Proof: See Appendix. ∎

With Lemma 3 we establish the known quadratic rate of convergence of the suboptimality of the Newton update in (24). Now by substituting the upper bound on $\tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \tilde{R}_{k+1}^*$ from Lemma 2, a condition can easily be derived under which the suboptimality of the new iterate is within the accuracy $\hat{V}$ of $\tilde{R}_{k+1}$. Using the results of Lemmata 1–3, we present our main result in the following theorem.

Theorem 1: Consider Newton's method defined in (24) and the full learning method detailed in Algorithm 1. Define $\tilde{V}_N$ to be the statistical accuracy of $N = Mn$ samples by (21), with n samples taken from each of the M most recent channel distributions $\mathcal{H}_k$. Further consider the variable $\boldsymbol{\mu}_k$ as a $\hat{V}$-optimal solution of the loss $\tilde{R}_k$, and suppose Assumptions 1–4 hold. If the sample size n and window size M are chosen such that the following conditions

$$\left(\frac{2(\Delta + c\hat{V})\hat{V}}{\alpha\hat{V}} \right)^{1/2} + \frac{2\hat{V}^{1/2} + \bar{D}_k}{(\alpha\hat{V})^{1/2}} < \frac{1}{4} \qquad (33)$$

$$144(4\tilde{V}_N + \hat{V} + 2D_k)^2 \leq \hat{V} \qquad (34)$$

are satisfied, then the variable $\boldsymbol{\mu}_{k+1}$ computed from (24) has the suboptimality of $\hat{V}$ with high probability, i.e.,

$$\tilde{R}_{k+1}(\boldsymbol{\mu}_{k+1}) - \tilde{R}_{k+1}^* \leq \hat{V}, \qquad \text{w.h.p.} \qquad (35)$$

The inequalities (33)–(34) in Theorem 1 specify conditions under which $\boldsymbol{\mu}_{k+1}$ as generated by (24) is a $\hat{V}$-optimal solution of $\tilde{R}_{k+1}$. Note that these conditions come directly from the preceding lemmata. Thus, when these conditions are satisfied, single iterations of Newton's method at each epoch k—as detailed in Algorithm 1—successively generate approximately optimal dual parameters. A further discussion of the satisfaction of such conditions in regards to practical implementation is provided later in Section VI. We first extend the theoretical result of Theorem 1 to establish properties of the resulting WCP solution.

Remark 6: Observe in Theorem 1 that the provided conditions cannot be satisfied if the true statistical accuracy $\tilde{V}_N$ is greater than the selected $\hat{V}$. While we assume in our analysis this is not the case, (i.e., $\hat{V}$ is a conservative estimate of $\tilde{V}_N$), this may not be guaranteed if very little information is known about V_N. In the case $\hat{V} < V_N$, we point out that the results in Theorem 1 can simply be modified by replacing achieved accuracy $\hat{V}$ by V_N. In other words, the accuracy we can achieve is limited by the greater of these terms. We do not go through the details of this analysis for clarity of presentation, but such result can be obtained through the same steps of the preceding analysis.

B. Suboptimality in Wireless Control System

Because the proposed Newton method indeed solves (15) to within a statistical approximation $\hat{V}$, it is important to consider the effect of such an approximation on the original WCP in (WCP$_k$). In this section we provide a sequence of results that characterize the accuracy of the solutions generated by the Newton update in (24) in the original primal control problem in (WCP$_k$). Firstly, recall the constraints in (WCP$_k$) reflect both a power budget limited by $p_{\max}$ and that the auxiliary variable y^i should not exceed the expected packet success function $q(\cdot)$. In solving the dual problem approximately, we may then also violate these constraints by a small margin. We can specifically characterize such a constraint violation, as well as address the suboptimality in terms of the primal objective. Both these results together can then be combined to demonstrate the stability of

the switched system WCP introduced in Example 1. To do so, we first introduce an assumption regarding the feasibility and boundedness of the dual loss solutions L_k^* and the optimal dual point $\boldsymbol{\mu}_k^*$.

Assumption 5: For all epochs k, the problem in (WCP$_k$) under distribution $\mathcal{H}_k$ is strictly feasible. There also exists constants $\mathcal{K}$ and $\hat{\mathcal{K}}$ such that the optimal dual objective value L_k^* is bounded as $L_k^* \leq \mathcal{K}$ and optimal dual variable bounded as $\|\boldsymbol{\mu}_k^*\| \leq \hat{\mathcal{K}}$.

From strict feasibility of the primal problem in (WCP$_k$), we also obtain a finite upper bound on the value of the dual function. This can be used with the suboptimality result in Theorem 1 to bound the norm of the dual variables $\boldsymbol{\mu}_k$ generated from the Newton update in (24). This is presented in the following corollary.

Corollary 1: The norm of the dual variables $\boldsymbol{\mu}_k$ generated by the update in (24) is bounded as $\|\boldsymbol{\mu}_k\| \leq \sqrt{(2/\alpha)} + \hat{\mathcal{K}}$.

Proof: From strong convexity we have that $\|\boldsymbol{\mu}_k - \tilde{\boldsymbol{\mu}}_k^*\|^2 \leq (2/\alpha\hat{V})(\tilde{R}_k(\boldsymbol{\mu}_k) - \tilde{R}_k^*)$. Using the reverse triangle inequality with (35) and Assumption 5, we obtain the intended result. ∎

Observe that the boundedness of the solutions to the regularized dual function in Assumption 5 in effect states that, for all distributions $\mathcal{H}_k$, the empirical, or sampled, versions of the constrained problem in (WCP$_k$) will be strictly feasible. From here, we can establish through duality a bound on each constraint violation that may occur from solving the dual problem to its statistical accuracy. This result is stated in the following proposition.

Proposition 2: Consider $\boldsymbol{\mu}_k$ to be a $\hat{V}$-optimal minimizer of $\tilde{R}_k$, i.e., $\tilde{R}_k(\boldsymbol{\mu}_k) - \tilde{R}_k^* \leq \hat{V}$. Further consider $\mathbf{p}(\mathbf{h}, \boldsymbol{\mu}_k)$ and $\mathbf{y}(\boldsymbol{\mu}_k)$ to be the Lagrangian maximizers over dual parameter $\boldsymbol{\mu}_k$. If Assumptions 1 and 5 hold, then the norm of the constraint violations in (WCP$_k$) can each be upper bounded as

$$\left| \sum_{i=1}^{m} \mathbb{E}_{h_k^i} \left(p^i(h_k^i, \boldsymbol{\mu}) \right) - p_{\max} \right| \leq \sqrt{2\Delta(\tilde{V}_N + C\hat{V})},$$
$$(36)$$

$$\|\mathbf{y}(\boldsymbol{\mu}_k) - \mathbb{E}_{\mathbf{h}_k} \{ \mathbf{q}(\mathbf{h}_k, \mathbf{p}(\mathbf{h}_k, \boldsymbol{\mu}_k)) \}\| \leq \sqrt{2\Delta(\tilde{V}_N + C\hat{V})},$$
$$(37)$$

where $C := 1 + \rho + \beta\kappa$ and κ such that $\mathbf{1}^T \log_\epsilon(\boldsymbol{\mu}_k) \leq \kappa$.

Proof: See Appendix. ∎

In Proposition 2, we establish a bound that is proportional to $\hat{V}$ on the violation of the constraints in (WCP$_k$). There are two points to be stressed here. First, is that this constraint violation can indeed be made small by controlling the target accuracy $\hat{V}$. Additionally, we point out that the violation of the budget constraint can be controlled by adding a slack term to the maximum power as $\hat{p}_{\max} = p_{\max} - 2\Delta C\hat{V}$. In this way, any such violation will still be within the true intended budget $p_{\max}$.

We proceed by establishing suboptimality of the generated variables $\mathbf{y}(\boldsymbol{\mu}_k)$ in terms of control performance. Recall the final result in Theorem 1 that establishes at each epoch k, the current dual function value $\tilde{R}_k(\boldsymbol{\mu}_k)$ will be within accuracy $\hat{V}$ of the optimal value $\tilde{R}_k(\tilde{\boldsymbol{\mu}}_k^*)$ (after satisfying the necessary

conditions). To establish that the control systems induced by such dual parameters $\boldsymbol{\mu}_k$ remain stable, we first connect the accuracy of the dual function value to the accuracy of associated primal variables $\mathbf{p}(\mathbf{h}, \boldsymbol{\mu}_k)$ and $\mathbf{y}(\boldsymbol{\mu}_k)$ with respect to their optimal values $\mathbf{p}_k^*(\mathbf{h}) := \mathbf{p}(\mathbf{h}, \boldsymbol{\mu}_k^*)$ and $\mathbf{y}_k^* := \mathbf{y}(\tilde{\boldsymbol{\mu}}_k^*)$. This bound is established in the following theorem.

Theorem 2: Consider $\boldsymbol{\mu}_k$ to be a $\hat{V}$-optimal minimizer of $\tilde{R}_k$, i.e., $\tilde{R}_k(\boldsymbol{\mu}_k) - \tilde{R}_k^* \leq \hat{V}$. Further consider $\mathbf{p}(\mathbf{h}, \boldsymbol{\mu}_k)$ and $\mathbf{y}(\boldsymbol{\mu}_k)$ to be the Lagrangian maximizers over dual parameter $\boldsymbol{\mu}_k$. Under Assumptions 1–5 the primal objective function sub-optimality $J(\mathbf{y}(\boldsymbol{\mu}_k)) - J(\mathbf{y}_k^*)$ can be upper bounded as

$$J(\mathbf{y}(\boldsymbol{\mu}_k)) - J(\mathbf{y}_k^*) \leq (1 + C)\Delta \left(\frac{1}{\alpha} + 2\hat{V}(\sqrt{2/\alpha} + \hat{\mathcal{K}}) \right). \tag{38}$$

Proof: See Appendix. ∎

In Theorem 2, we derive a bound on the suboptimality of the primal objective function $J(\mathbf{y})$ that is proportional also to the statistical accuracy $\hat{V}$ plus a constant. Recall that this function is, in general, a measure of the control performance of the system. Thus, solving the dual problem approximately indeed can be translated into approximate accuracy in terms of our original utility metric with respect to the control system. In many problems, the performance $J(\mathbf{y})$ will also effectively establish a stability margin for control systems that have unstable regions of operation. To demonstrate the effect of using the proposed Newton's method over a non-stationary wireless channel, we return to the switched dynamical system in Example 1.

C. Stability of Switched Dynamical System (Example 1)

Consider the switched dynamical system given in (4) and the derived performance metric $J(\mathbf{y})$ in (9) that tracks the asymptotic behavior of the state x_t. In this system, if the open loop gain is unstable $|A_o| > 1$ it can indeed cause the system to grow in an unstable manner if the system is not closed sufficiently often. As mentioned in Example 1 the system reaches instability if $yA_c^2 + (1 - y)A_o^2$ becomes close to 1. A question of interest in this example is, using the power allocation policy found using Newton's method over a time-varying channel, whether or not the system remains stable over time. We can indeed demonstrate this to be true with the following argument.

From Theorem 2, we obtained that the primal suboptimality with respect to the control performance function $J(\mathbf{y})$ is bounded by a term proportional to $\hat{V}$. Assuming that $J(\mathbf{y}_k^*)$ is finite for all epochs k, it follows then that the generated performance $J(\mathbf{y}(\boldsymbol{\mu}_k))$ is also finite. Considering the expression for $J^i(y^i)$ given in (9), this is finite if and only if the denominator is positive, i.e., there exists a ω such that

$$1 - y^i(\boldsymbol{\mu}_k)((A_o^i)^2 - (A_c^i)^2) \leq \omega < 1 \tag{39}$$

at all epochs k.

Moreover from Proposition 2 we also have that the actual packet success rate during epoch k satisfies

$$\mathbb{E}_{h_k} \left\{ q(h_k^i, p(h_k^i, \mu_k)) \right\} \geq y^i(\boldsymbol{\mu}_k) - \sqrt{2\Delta(\tilde{V}_N + C\hat{V})}, \tag{40}$$

If the statistical accuracy at the right hand side of this expression is sufficiently small, then using (39) we also get that

$$1 - \mathbb{E}_{h_k} \left\{ q(h_k^i, p(h_k^i, \mu_k)) \right\} ((A_o^i)^2 - (A_c^i)^2) \leq \tilde{\omega} < 1 \tag{41}$$

In particular this holds if $\sqrt{2\Delta(\tilde{V}_N + C\hat{V})}((A_o^i)^2 - (A_c^i)^2) < 1 - \omega$.

Substituting (41) back into the recursive expression in (6), we get that the variance of the state at each time step satisfies

$$\mathbb{E}(x_{t+1}^i)^2 \leq \tilde{\omega}\mathbb{E}(x_t^i)^2 + W^i. \tag{42}$$

Operating recursively and using the geometric series as in Example 1, we can bound (42) as

$$\mathbb{E}(x_{t+1}^i)^2 \leq \tilde{\omega}^{t+1}\mathbb{E}(x_0^i)^2 + W^i \frac{1 - \tilde{\omega}^{t+1}}{1 - \tilde{\omega}}. \tag{43}$$

As both terms on the right hand side of (43) are finite, we can conclude that the state variables remain bounded in variance for all t in the non-stationary channel.

VI. DETAILS OF IMPLEMENTATION

In this section we provide a discussion of necessary considerations for practical implementation of the result in Theorem 1. Observe that the conditions in (33) and (34) are functions of four primary terms, $\hat{V}$, $\tilde{V}_N$, D_k, and $\bar{D}_k$. While $\hat{V}$ is user-selected, the latter three terms come directly from statistical properties of the control performance functions and the channel distribution. They can, however, be indirectly controlled for with some careful implementation techniques.

First, consider that the latter two terms D_k and $\bar{D}_k$ provide a bound on the difference of the neighboring expected loss functions L_k and L_{k+1} and their gradients, respectively. Thus, these terms collectively can be interpreted as a bound on the degree of non-stationarity of the channel distribution $\mathcal{H}$ between successive time iterations, or in other words the rate at which the channel changes over time epochs. In a practical sense, this rate is controllable by determining how much real time makes up a single discrete time epoch. That is, time epochs k and $k + 1$ that are closer together in a real time-sense will naturally have a lower bound for D_k, and $\bar{D}_k$, assuming the rate of change of the channel distribution is indeed smooth. In this sense, D_k and $\bar{D}_k$ can be lowered to satisfy the conditions in (33) and (34) by considering shorter time between discrete epochs. This is to say that, because the channel conditions are not in our control, if necessary we may change the rate at which we apply our algorithm in a real time sense. By using shorter epochs, we collect channel samples and run the proposed Newton step more often to adapt to quickly changing channel conditions.

The second term present in the conditions of Theorem 1— namely $\tilde{V}_N$—represents the statistical accuracy of the non-i.i.d. samples taken from the window of M most recent channel distributions with respect to the current channel distribution. A condition on $\tilde{V}_N$ in fact then indirectly provides conditions on the sample size n and window size M used to define $\tilde{R}_k$ necessary to learn a $\hat{V}$-optimal solution. We reiterate here that, in the simpler setting of $M = 1$, a well-studied bound on $\tilde{V}_N$ exists of the order $\mathcal{O}(1/\sqrt{n})$. For the case of windowed sampling the
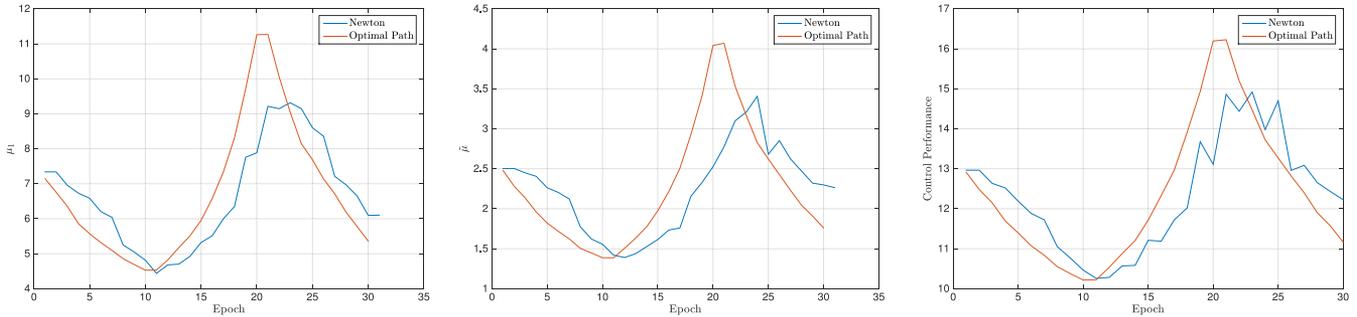
Fig. 3. Convergence paths of optimal values vs. values generated by the Newton learning method for time-varying $\mathcal{H}_k$ for dual variables (left) μ^1, (center) $\tilde{\mu}$, and (right) control performance $\sum J^i(y^i)$. Newton's method is able to find an approximately optimal value for the dual variables and respective control performance at each iteration.

bound on $\tilde{V}_N$ can nonetheless still be varied through various choices of window size M and sample draw size n. However, because the exact nature of both $\tilde{V}_N$ and D_k come from statistical properties not known in practice, precise selection of such parameters n and M can be chosen via a standard backtracking procedure.

The details of the backtracking procedure can be seen in Steps 10 and 11 in Algorithm 1. At each epoch k, the parameters n and M are initialized to n_0 and M_0 in Step 4. In the inner loop, in Step 10 these parameters are respectively increased and decreased by factors of Γ and γ after performing the Newton step. In Step 11, the accuracy of the new dual iterate $\boldsymbol{\mu}_{k+1}$ is checked to be within the intended accuracy $\hat{V}$. Note that, while the suboptimality cannot be checked directly without knowledge of $\tilde{R}_{k+1}^*$, it can be checked indirectly by checking the norm of the gradient $\|\nabla \tilde{R}_{k+1}(\boldsymbol{\mu}_{k+1})\| < (\sqrt{2\alpha})\hat{V}$ from the strong convexity property in (25). If the condition in Step 11 is satisfied, the parameters n and M require no further modification. Otherwise, they are further modified until μ_{k+1} is within the target accuracy which in turn may imply that the conditions in (33) and (34) are satisfied. Note that the backtracking rates γ, Γ are standard parameters used in the definition of a backtracking algorithm and effectively tradeoff the speed of the backtracking search vs. its thoroughness or accuracy. Generally speaking, values closer to 1 will result in a slower, more careful backtracking search while values of γ and Γ that are, respectively, smaller and larger will result in a faster, more aggressive search. Tuning of these parameters should thus reflect the desired tradeoff. With this practical considerations in mind, we proceed by simulating a wireless control learning problem using the proposed use of Newton's method on the ERM relaxation.

VII. SIMULATION RESULTS

We simulate the performance of our second order learning method on a simple WCP. Consider the 1-dimensional switched dynamical system in (1) governed by the transition constants A_o and A_c for $m = 4$ systems/states. The control performance for the ith agent $J^i(y^i)$ measures the mean square error performance and is now given by the expression in (9). The open and closed loop control gains for each agents are chosen between $[1.1, 1.5]$ and $[0, 0.8]$, respectively. The probability

of successful transmission for agent i is modeled as a negative exponential function of both the power and channel state, $q(h^i, p^i(h^i)) := 1 - e^{-h^i p^i(h^i)}$, while channel states at epoch k are drawn from an exponential distribution with mean u_k. The channel varies over time by the mean u_k changing for different times. We draw $n = 200$ samples and store a window of the previous $M = 5$ distributions for a total of $N = 1000$ samples at each epoch. As we assume the that channel statistics vary only vary across time *epochs*, but stay constant within a single epoch, we may consider it reasonable to collect 200 channel samples within an epoch.

To demonstrate the ability of Newton's method to instantaneously learn an approximately optimal power allocation as the channel distribution varies over time, we perform Algorithm 1 over the ERM problem in (15) with the defined control performance $J(\cdot)$, transmission probabilities $q(\cdot)$ and channel distributions $\mathcal{H}_k$. In Fig. 3 we show the path of Newton's method at each time k for the dual variables μ_k^1, $\tilde{\mu}_k$, and the control performance $\sum_{i=1}^m J^i(y_k^i)$. The red line of each figure plots the optimal values for the current distribution parameter u_k as it changes with k. These values are obtained by solving the optimization problem at each epoch offline a priori. The blue line, alternatively, plots the values generated by Newton's method for each epoch k in an online manner. The channel evolves at each iteration by a fixed rate $u_{k+1} = u_k \pm r$ for some rate r. Observe that within some small error Newton's method is indeed able to quickly and approximately find each new solution as the channel varies over time.

To compare the effect of selecting different choices of accuracy $\hat{V}$ numerically, we present in Fig. 4 the simulation performance of two representative cases with respective accuracies of $\hat{V} = 0.01$ (left) and $\hat{V} = 0.03$ (right). In the top figures, we show the suboptimality relative to the optimal control performance and show on the bottom figures the resulting constraint violation (where a positive value reflect violation) over a set of time epochs where the channel varies. Here, we see an interesting case that highlights the need of proper selection or estimation of $\hat{V}$. Although the left hand figures strive for a better accuracy, the performance is better on the right hand figures. This is due to the fact that single iterations of Newton's method cannot reach accuracies of 0.01, resulting in a more suboptimal trajectory of resource allocation policies. On the other hand,
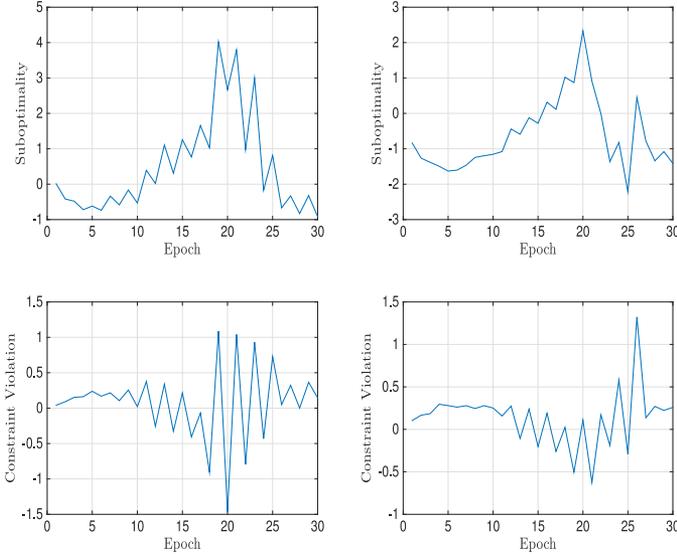
Fig. 4. Comparison of suboptimality (top) and constraint violation (bottom) for the case of $\hat{V} = 0.01$ (left) and $\hat{V} = 0.03$ (right). Although the right-hand figures strive for less accuracy, they perform better because Newton's method can adapt to the intended accuracy more easily with single iterations.
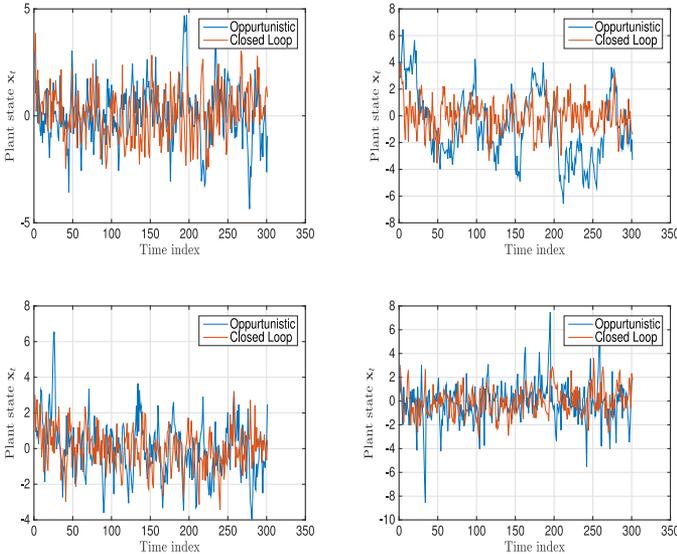

Fig. 5. Dynamic evolution of each of the 4 state variables over the time-varying channel. The blue curve shows the opportunistic power allocation policy found with Newton's method while the red curve shows the evolution assuming the loop can always be closed.

the more moderate goal of 0.03 allows for the learning method to reach intended accurate goals with each step of Newton's method as the channel varies.

Using the dual parameters found by Newton's method, we simulate the resulting dynamical system. The dual parameters are used to determine the power allocation policy, which is used to determine transmission probabilities given current channel conditions. In Fig. 5 we show the resulting state evolution of x_t^i for each of the 4 state variables. The blue curve shows the process using the opportunistic transmission policy from Newton's

method, while the red curve shows the process when the loop is always closed, i.e., no packet drops. Here, we observe that while there are some instances when the state variable grows large when the system is in open loop, overall the system remains stable over time.

VIII. CONCLUSION

In this paper we considered the wireless control system over a non-stationary wireless channel. The problem of maximizing a control utility subject to resource constraints can be formulated as a stochastic optimization problem in the dual domain. Because the wireless channel is random and time-varying, channel samples must be taken, resulting in a relaxed empirical risk minimization (ERM) problem. Standard ERM techniques do not suffice in the wireless setting because the channel is constantly changing. We propose the use of Newton's method, whose local quadratic convergence property allows us to continuously learn and adapt our optimal power allocation policies to changes in the channel distribution. We derive specific conditions on achieving instantaneous convergence to an approximate solution and characterize the suboptimality and stability in the wireless control problem (WCP). We additionally provide numerical simulations that demonstrate the use of Newton's method to learn and track optimal power allocations over a time varying channel. While this paper considers only resource allocation on contention-free links, consider the scheduling problem on a shared channel with non-stationary distributions remains an area of future work.

APPENDIX
PROOF OF LEMMA 1

We start with the definition of the Newton decrement at time $k + 1$. We can add and subtract $\nabla \tilde{R}_k(\boldsymbol{\mu}_k)$ and upper bound using the triangle inequality as

$$\lambda_{k+1}(\boldsymbol{\mu}_k) = \|\mathbf{H}_{k+1}^{-1/2} \nabla \tilde{R}_{k+1}(\boldsymbol{\mu})\| = \|\nabla \tilde{R}_{k+1}(\boldsymbol{\mu}_k)\|_{\mathbf{H}_{k+1}^{-1}}$$
$$\leq \|\nabla \tilde{R}_k(\boldsymbol{\mu}_k)\|_{\mathbf{H}_{k+1}^{-1}} + \|\nabla \tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \nabla \tilde{R}_k(\boldsymbol{\mu}_k)\|_{\mathbf{H}_{k+1}^{-1}}.$$
$$(44)$$

First, we will upper bound the second term in (44). By adding and subtracting the expected losses $\nabla L_k(\boldsymbol{\mu}_k)$ and $\nabla L_{k+1}(\boldsymbol{\mu}_k)$ and using the triangle inequality to obtain

$$\|\nabla \tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \nabla \tilde{R}_k(\boldsymbol{\mu}_k)\| \leq \|\nabla \hat{L}_{k+1}(\boldsymbol{\mu}_k) - \nabla L_{k+1}(\boldsymbol{\mu}_k)\|$$
$$+ \|\nabla L_k(\boldsymbol{\mu}_k) - \nabla \hat{L}_k(\boldsymbol{\mu}_k)\| + \|\nabla L_{k+1}(\boldsymbol{\mu}_k) - \nabla L_k(\boldsymbol{\mu}_k)\|.$$

The first two terms in the above sum are bounded by $\tilde{V}_N^{1/2}$ per (26), while the third term is the difference of two consecutive loss functions and is therefore bounded by $\bar{D}_k$ from (28). The norm weight $\mathbf{H}_{k+1}^{-1}$ additionally provides a bound of $\alpha \hat{V}$ as the strong convexity constant of $\tilde{R}_{k+1}$ providing an upper bound on the norm of Hessian inverse as in (25). Combining these, we obtain

$$\|\nabla \tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \nabla \tilde{R}_k(\boldsymbol{\mu}_k)\|_{\mathbf{H}_{k+1}^{-1}} \leq \frac{2\tilde{V}_N^{1/2} + \bar{D}_k}{(\alpha \hat{V})^{1/2}}. \quad (45)$$

We now can bound the first term in (44) using the Lipschitz continuity of the gradient $\Delta + c\hat{V}$, i.e.

$$\|\nabla \tilde{R}_k(\boldsymbol{\mu}_k)\|_{\mathbf{H}_{k+1}^{-1}} \leq \left(\frac{2(\Delta + c\hat{V})\|\boldsymbol{\mu}_k - \tilde{\boldsymbol{\mu}}_k^*\|}{\alpha \hat{V}} \right)^{1/2} \quad (46)$$

Recall that $\boldsymbol{\mu}_k$ is given to be a $\hat{V}$-accurate minimizer of $\tilde{R}_k$. The difference $\|\boldsymbol{\mu}_k - \tilde{\boldsymbol{\mu}}_k^*\|$ can subsequently be bounded with $\hat{V}$, resulting in the final bound for the first term

$$\|\nabla \tilde{R}_k(\boldsymbol{\mu}_k)\|_{\mathbf{H}_{k+1}^{-1}} \leq \left(\frac{2(\Delta + c\hat{V})\hat{V}}{\alpha \hat{V}} \right)^{1/2} \quad (47)$$

To be in the quadratic convergence region, i.e., $\lambda_{k+1}(\boldsymbol{\mu}_k) < 1/4$, follows by summing (45) and (47) as in (30).

APPENDIX
PROOF OF LEMMA 2

To prove this result, we start by expanding the term $\tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \tilde{R}_{k+1}^*$. By adding and subtracting $\tilde{R}_k(\boldsymbol{\mu}_k)$, $\tilde{R}_k^*$, and $\tilde{R}_k(\boldsymbol{\mu}_{k+1}^*)$, we obtain

$$\tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \tilde{R}_{k+1}^* = \tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \tilde{R}_k(\boldsymbol{\mu}_k)$$
$$+ \tilde{R}_k(\boldsymbol{\mu}_k) - \tilde{R}_k^*$$
$$+ \tilde{R}_k^* - \tilde{R}_k(\boldsymbol{\mu}_{k+1}^*)$$
$$+ \tilde{R}_k(\boldsymbol{\mu}_{k+1}^*) - \tilde{R}_{k+1}^*. \quad (48)$$

We now individually bound each of the four differences in (48). Firstly, the difference $\tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \tilde{R}_k(\boldsymbol{\mu}_k)$ becomes

$$\tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \tilde{R}_k(\boldsymbol{\mu}_k) = \hat{L}_{k+1}(\boldsymbol{\mu}_k) - \hat{L}_k(\boldsymbol{\mu}_k), \quad (49)$$

Using the same reasoning as in (45) with the functional statistical accuracy bound in place of the bound for gradients in (26) and using (27) in place of (28), we obtain the equivalent bound

$$\tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \tilde{R}_k(\boldsymbol{\mu}_k) \leq 2\tilde{V}_N + D_k. \quad (50)$$

For the second term in (48), we again use the fact that $\boldsymbol{\mu}_k$ as an $\hat{V}$-optimal solution for the sub-optimality $\tilde{R}_k(\boldsymbol{\mu}_k) - \tilde{R}_k^*$ to bound with the statistical accuracy as

$$\tilde{R}_k(\boldsymbol{\mu}_k) - \tilde{R}_k^* \leq \hat{V}. \quad (51)$$

We proceed with bounding the third term in (48). Based on the definition of $\boldsymbol{\mu}_k^*$ as the optimal solution of the loss $\tilde{R}_k$, the the difference $\tilde{R}_k^* - \tilde{R}_k(\boldsymbol{\mu}_{k+1}^*)$ is always negative, i.e.,

$$\tilde{R}_k^* - \tilde{R}_k(\boldsymbol{\mu}_{k+1}^*) \leq 0. \quad (52)$$

For the fourth term in (48), we use the triangle inequality to bound the difference $\tilde{R}_k(\boldsymbol{\mu}_{k+1}^*) - \tilde{R}_{k+1}^*$ in (48) as

$$\tilde{R}_k(\boldsymbol{\mu}_{k+1}^*) - \tilde{R}_{k+1}^* = \hat{L}_k(\boldsymbol{\mu}_{k+1}^*) - \hat{L}_{k+1}(\boldsymbol{\mu}_{k+1}^*)$$
$$\leq 2\tilde{V}_N + D_k. \quad (53)$$

Observe that (53) uses the same reasoning as (50). Replacing the differences in (48) by the upper bounds in (50)–(53),

$$\tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \tilde{R}_{k+1}^* \leq 4\tilde{V}_N + \hat{V} + 2D_k \quad \text{w.h.p.} \quad (54)$$

APPENDIX
PROOF OF LEMMA 3

The proof for this result follows from [34, Proposition 4], which we repeat here for completeness. We proceed by bounding the difference $\tilde{R}_{k+1}(\boldsymbol{\mu}) - \tilde{R}_{k+1}^*$ in terms of the Newton decrement parameter $\lambda_{k+1}(\boldsymbol{\mu})$. We first use the result in [39, Theorem 4.1.11], showing that

$$\lambda_{k+1}(\boldsymbol{\mu}) - \ln\left(1 + \lambda_{k+1}(\boldsymbol{\mu})\right)$$
$$\leq \tilde{R}_{k+1}(\boldsymbol{\mu}) - \tilde{R}_{k+1}^*$$
$$\leq -\lambda_{k+1}(\boldsymbol{\mu}) - \ln\left(1 - \lambda_{k+1}(\boldsymbol{\mu})\right). \quad (55)$$

We can use the Taylor's expansion of $\ln(1 + a)$ for $a = \lambda_{k+1}(\boldsymbol{\mu})$ to show that $\lambda_{k+1}(\boldsymbol{\mu}) - \ln(1 + \lambda_{k+1}(\boldsymbol{\mu}))$ is bounded below by $(1/2)\lambda_{k+1}(\boldsymbol{\mu})^2 - (1/3)\lambda_{k+1}(\boldsymbol{\mu})^3$ for $0 < \lambda_{k+1}(\boldsymbol{\mu}) < 1/4$. Likewise, we have that $(1/6)\lambda_{k+1}(\boldsymbol{\mu})^2 \leq (1/2)\lambda_{k+1}(\boldsymbol{\mu})^2 - (1/3)\lambda_{k+1}(\boldsymbol{\mu})^3$ and subsequently $\lambda_{k+1}(\boldsymbol{\mu}) - \ln(1 + \lambda_{k+1}(\boldsymbol{\mu}))$ is bounded below by $(1/6)\lambda^2$. We again use Taylor's expansion of $\ln(1 - a)$ for $a = \lambda_{k+1}(\boldsymbol{\mu})$ to show that $-\lambda_{k+1}(\boldsymbol{\mu}) - \ln\left(1 - \lambda_{k+1}(\boldsymbol{\mu})\right)$ is bounded above by $\lambda_{k+1}(\boldsymbol{\mu})^2$ for $\lambda_{k+1}(\boldsymbol{\mu}) < 1/4$; see e.g., [27, Ch. 9]. Considering these bounds and the inequalities in (55) we obtain that

$$\frac{1}{6}\lambda_{k+1}(\boldsymbol{\mu})^2 \leq \tilde{R}_{k+1}(\boldsymbol{\mu}) - \tilde{R}_{k+1}^* \leq \lambda_{k+1}(\boldsymbol{\mu})^2. \quad (56)$$

Because we assume that $\lambda_{k+1}(\boldsymbol{\mu}_k) \leq 1/4$, the quadratic convergence rate of Newton's method for self-concordant functions [27] implies that the Newton decrement has a quadratic convergence and we can write

$$\lambda_{k+1}(\boldsymbol{\mu}_{k+1}) \leq 2\lambda_{k+1}(\boldsymbol{\mu}_k)^2. \quad (57)$$

We combine the results in (56) and (57) to show that the optimality error $\tilde{R}_{k+1}(\boldsymbol{\mu}_{k+1}) - \tilde{R}_{k+1}^*$ has an upper bound which is proportional to $(\tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \tilde{R}_{k+1}^*)^2$. In particular, we can write $\tilde{R}_{k+1}(\boldsymbol{\mu}_{k+1}) - \tilde{R}_{k+1}^* \leq \lambda_{k+1}(\boldsymbol{\mu}_{k+1})^2$ based on the second inequality in (56). This observation in conjunction with the result in (57) implies that

$$\tilde{R}_{k+1}(\boldsymbol{\mu}_{k+1}) - \tilde{R}_{k+1}^* \leq 4\lambda_{k+1}(\boldsymbol{\mu}_k)^4. \quad (58)$$

The first inequality in (56) implies that $\lambda_{k+1}(\boldsymbol{\mu}_k)^4 \leq 36(\tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \tilde{R}_{k+1}^*)^2$. Thus, we can substitute $\lambda_{k+1}(\boldsymbol{\mu}_k)^4$ in (58) by $36(\tilde{R}_{k+1}(\boldsymbol{\mu}_k) - \tilde{R}_{k+1}^*)^2$ to obtain the result in (32).

APPENDIX
PROOF OF PROPOSITION 2

We begin by bounding the gradient of the expected dual loss $L(\boldsymbol{\mu}_k)$ at the kth dual iterate $\boldsymbol{\mu}_k$ by using Lipschitz continuity, i.e.

$$\|\nabla L_k(\boldsymbol{\mu}_k)\|^2 \leq 2\Delta(L_k(\boldsymbol{\mu}_k) - L_k^*). \quad (59)$$

We expand the sub-optimality $L(\boldsymbol{\mu}_k) - L^*$ by adding and subtracting terms as follows

$$\frac{1}{2\Delta}\|\nabla L_k(\boldsymbol{\mu}_k)\|^2 \leq L_k(\boldsymbol{\mu}_k) - \tilde{L}_k(\boldsymbol{\mu}_k) + \tilde{L}_k(\boldsymbol{\mu}_k)$$
$$- \tilde{R}_k(\boldsymbol{\mu}_k) + \tilde{R}_k(\boldsymbol{\mu}_k) - \tilde{R}_k^* + \tilde{R}_k^* - L_k^*, \quad (60)$$

where we recall the notation $\tilde{R}_k^* := \tilde{R}_k(\tilde{\boldsymbol{\mu}}_k^*)$. We now proceed by bounding each successive pair of terms in (60). The first difference $L_k(\boldsymbol{\mu}_k) - \tilde{L}_k(\boldsymbol{\mu}_k)$ comes from the sampling and is thus bounded by the statistical accuracy $\tilde{V}_N$. The second difference $\tilde{L}_k(\boldsymbol{\mu}_k) - \tilde{R}_k(\boldsymbol{\mu}_k)$ can be bounded by the regularizers as

$$\tilde{L}_k(\boldsymbol{\mu}_k) - \tilde{R}_k(\boldsymbol{\mu}_k) \le \beta\hat{V}\mathbf{1}^T \log_\epsilon(\boldsymbol{\mu}_k) - \frac{\alpha\hat{V}}{2}\|\boldsymbol{\mu}_k\|^2. \quad (61)$$

The second term on the right hand side of (61) is negative and can be ignored. Because the dual variable $\|\boldsymbol{\mu}_k\|$ was upper bounded in Corollary 1, we can place a finite bound on $\mathbf{1}^T \log_\epsilon(\boldsymbol{\mu}_k) \le \kappa$ and then bound the term $\beta\hat{V}\mathbf{1}^T \log_\epsilon(\boldsymbol{\mu}_k) \le \beta\hat{V}\kappa$. The third difference $\tilde{R}_k(\boldsymbol{\mu}_k) - \tilde{R}_k^*$ is bounded by the suboptimality $\hat{V}$ from the main result in (35) and the fourth difference $\tilde{R}_k^* - L_k^*$ can be bounded by $\rho\hat{V}$ from (23). We can therefore bound the gradient of the dual loss as

$$\|\nabla L_k(\boldsymbol{\mu}_k)\|^2 \le 2\Delta(\tilde{V}_N + C\hat{V}), \quad (62)$$

where $C := 1 + \rho + \beta m \log \kappa$. From here, consider that the norm of the dual gradient $\|\nabla L_k(\boldsymbol{\mu}_k)\|^2$ is the sum of squares of each constraint violation in (WCP_k), i.e.,

$$\left(\sum_{i=1}^m \mathbb{E}_{h_k^i}\left(p^i(h)\right) - p_{\max}\right)^2 + \sum_{i=1}^m \left(y^i - \mathbb{E}_{h_k^i}\left\{q(h, p^i(h))\right\}\right)^2$$
$$\le 2\Delta(\tilde{V}_N + C\hat{V}). \quad (63)$$

The results in (36) and (37) then follow from here.

APPENDIX
PROOF OF THEOREM 2

Consider that, using the definitions of the primal maximizers $\mathbf{p}(\mathbf{h}, \boldsymbol{\mu}_k)$ and $\mathbf{y}(\boldsymbol{\mu}_k)$ at a dual point $\boldsymbol{\mu}_k$, we can write the dual function as

$$L(\boldsymbol{\mu}_k) = J(\mathbf{y}(\boldsymbol{\mu}_k)) + \boldsymbol{\mu}_k^T \left(\mathbb{E}_\mathbf{h}\check{\mathbf{q}}(\mathbf{p}(\mathbf{h}, \boldsymbol{\mu}_k)) - \check{\mathbf{y}}(\boldsymbol{\mu}_k)\right). \quad (64)$$

Likewise, we know from strong duality that the optimal dual values L_k^* is equivalent to the optimal primal objective value $J(\mathbf{y}_k^*)$. Therefore, we can write the suboptimality of dual functions as

$$L(\boldsymbol{\mu}_k) - L_k^* = J(\mathbf{y}(\boldsymbol{\mu}_k)) - J(\mathbf{y}_k^*)$$
$$+ \boldsymbol{\mu}_k^T \left(\mathbb{E}_\mathbf{h}\check{\mathbf{q}}(\mathbf{p}(\mathbf{h}, \boldsymbol{\mu}_k)) - \check{\mathbf{y}}(\boldsymbol{\mu}_k)\right). \quad (65)$$

Using the bound on dual suboptimality that comes from combining strong convexity and the gradient bound in (62), we can upper bound (65) as

$$(1 + C)\Delta/\alpha \ge J(\mathbf{y}(\boldsymbol{\mu}_k)) - J(\mathbf{y}_k^*)$$
$$+ \boldsymbol{\mu}_k^T \left(\mathbb{E}_\mathbf{h}\check{\mathbf{q}}(\mathbf{p}(\mathbf{h}, \boldsymbol{\mu}_k)) - \check{\mathbf{y}}(\boldsymbol{\mu}_k)\right). \quad (66)$$

We can lower bound the right hand side of (66) by taking the negative of the absolute value of the final term. Rearranging terms we obtain

$$(1 + C)\Delta/\alpha + |\boldsymbol{\mu}_k^T \left(\mathbb{E}_\mathbf{h}\check{\mathbf{q}}(\mathbf{p}(\mathbf{h}, \boldsymbol{\mu}_k)) - \check{\mathbf{y}}(\boldsymbol{\mu}_k)\right)|$$
$$\ge J(\mathbf{y}(\boldsymbol{\mu}_k)) - J(\mathbf{y}_k^*). \quad (67)$$

From here, we can upper bound the second term on the left hand side using the Cauchy-Schwartz inequality. The norm $\|\boldsymbol{\mu}_k\|$ is bounded by $\sqrt{2/\alpha} + \hat{\mathcal{K}}$ from Corollary 1 while the norm $\|\mathbb{E}_\mathbf{h}\check{\mathbf{q}}(\mathbf{p}(\mathbf{h}, \boldsymbol{\mu}_k)) - \check{\mathbf{y}}(\boldsymbol{\mu}_k)\|$ is bounded by $2\Delta(1 + C)\hat{V}$ from (37). This provides us the final result as

$$(1 + C)\Delta\left(\frac{1}{\alpha} + 2\hat{V}(\sqrt{2/\alpha} + \hat{\mathcal{K}})\right) \ge J(\mathbf{y}(\boldsymbol{\mu}_k)) - J(\mathbf{y}_k^*). \quad (68)$$

REFERENCES

[1] M. Eisen, K. Gatsis, G. J. Pappas, and A. Ribeiro, "Learning statistically accurate resource allocations in non-stationary wireless systems," in *Proc. Int. Conf. Acoust. Speech Signal Process.*, 2018. [Online]. Available: http://www.fling.seas.upenn.edu/~maeisen/wiki/icassp2.pdf

[2] M. Eisen, K. Gatsis, G. J. Pappas, and A. Ribeiro, "Learning in non-stationary wireless control systems via Newton's method," in *Proc. Amer. Control Conf.*, 2018, [Online]. Available: http://www.fling.seas.upenn.edu/~maeisen/wiki/acc2.pdf

[3] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[4] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. Sastry, "Foundations of control and estimation over lossy networks," *Proc. IEEE*, vol. 95, no. 1, pp. 163–187, 2007.

[5] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proc. IEEE*, vol. 95, no. 1, pp. 138–162, 2007.

[6] V. Gupta, B. Hassibi, and R. M. Murray, "Optimal LQG control across packet-dropping links," *Syst. Control Lett.*, vol. 56, no. 6, pp. 439–446, 2007.

[7] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Trans. Autom. Control*, vol. 49, no. 7, pp. 1056–1068, Jul. 2004.

[8] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link— Part I: Scalar systems," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3369–3395, Aug. 2006.

[9] N. Elia, "Remote stabilization over fading channels," *Syst. Control Lett.*, vol. 54, no. 3, pp. 237–249, 2005.

[10] O. C. Imer, S. Yüksel, and T. Başar, "Optimal control of LTI systems over unreliable communication links," *Automatica*, vol. 42, no. 9, pp. 1429–1439, 2006.

[11] M. B. Cloosterman, L. Hetel, N. Van de Wouw, W. Heemels, J. Daafouz, and H. Nijmeijer, "Controller synthesis for networked control systems," *Automatica*, vol. 46, no. 10, pp. 1584–1594, 2010.

[12] K. Gatsis, A. Ribeiro, and G. J. Pappas, "Optimal power management in wireless control systems," *IEEE Trans. Autom. Control*, vol. 59, no. 6, pp. 1495–1510, Jun. 2014.

[13] D. E. Quevedo, A. Ahlén, A. S. Leong, and S. Dey, "On Kalman filtering over fading wireless channels with controlled transmission powers," *Automatica*, vol. 48, no. 7, pp. 1306–1316, 2012.

[14] M. Rabi and K. H. Johansson, "Scheduling packets for event-triggered control," in *Proc. Eur. Control Conf*, 2009, pp. 3779–3784.

[15] M. Mazo and P. Tabuada, "Decentralized event-triggered control over wireless sensor/actuator networks," *IEEE Trans. Autom. Control*, vol. 56, no. 10, pp. 2456–2461, Oct. 2011.

[16] W. Heemels, K. H. Johansson, and P. Tabuada, "An introduction to event-triggered and self-triggered control," in *Proc. 51st Annu. Conf. Decis. Control*, 2012, pp. 3270–3285.

[17] J. Araújo, M. Mazo, A. Anta, P. Tabuada, and K. H. Johansson, "System architectures, protocols and algorithms for aperiodic wireless control systems," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 175–184, Feb. 2014.

[18] M. H. Mamduhi, D. Tolić, A. Molin, and S. Hirche, "Event-triggered scheduling for stochastic multi-loop networked control systems with packet dropouts," in *Proc. IEEE 53rd Annu. Conf. Decis. Control*, 2014, pp. 2776–2782.

[19] K. Gatsis, M. Pajic, A. Ribeiro, and G. J. Pappas, "Opportunistic control over shared wireless channels," *IEEE Trans. Autom. Control*, vol. 60, no. 12, pp. 3140–3155, Dec. 2015.

[20] K. Gatsis, A. Ribeiro, and G. J. Pappas, "Random access design for wireless control systems," *Automatica*, vol. 91, pp. 1–9, May 2018.

[21] T. Chen, A. Mokhtari, X. Wang, A. Ribeiro, and G. B. Giannakis, "Stochastic averaging for constrained optimization with application to online resource allocation," *IEEE Trans. Signal Process.*, vol. 65, no. 12, pp. 3078–3093, Jun. 2017.

[22] A. Mokkadem and M. Pelletier, "A generalization of the averaging procedure: The use of two-time-scale algorithms," *SIAM J. Control Optim.*, vol. 49, no. 4, pp. 1523–1543, 2011.

[23] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learning Representations*, 2015.

[24] O. Besbes, Y. Gur, and A. Zeevi, "Non-stationary stochastic optimization," *Oper. Res.*, vol. 63, no. 5, pp. 1227–1244, 2015.

[25] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[26] A. Ribeiro, "Optimal resource allocation in wireless communication and networking," *EURASIP J. Wireless Commun. Netw.*, vol. 2012, no. 1, 2012, Art. no. 272.

[27] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[28] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proc. COMPSTAT*, 2010, pp. 177–186.

[29] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," *J. Mach. Learn. Res.*, vol. 12, no. Jul, pp. 2121–2159, 2011.

[30] A. Defazio, F. Bach, and S. Lacoste-Julien, "Saga: A fast incremental gradient method with support for non-strongly convex composite objectives," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 1646–1654.

[31] O. Bousquet and L. Bottou, "The tradeoffs of large scale learning," in *Proc. Adv. Neural Inf. Process. Syst.*, 2008, pp. 161–168.

[32] V. Vapnik, *The Nature of Statistical Learning Theory*. Berlin, Germany: Springer Science & Business Media, 2013.

[33] R. Johnson and T. Zhang, "Accelerating stochastic gradient descent using predictive variance reduction," in *Proc. Adv. Neural Inf. Process. Syst.*, 2013, pp. 315–323.

[34] A. Mokhtari, H. Daneshmand, A. Lucchi, T. Hofmann, and A. Ribeiro, "Adaptive newton method for empirical risk minimization to statistical accuracy," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 4062–4070.

[35] M. Eisen, A. Mokhtari, and A. Ribeiro, "Large scale empirical risk minimization via truncated adaptive Newton method," in *Proc. 21st Int. Conf. Artificial Intelligence and Statistics*, 2017, pp. 1447–1455.

[36] S. Shalev-Shwartz, O. Shamir, N. Srebro, and K. Sridharan, "Learnability, stability and uniform convergence," *J. Mach. Learn. Res.*, vol. 11, pp. 2635–2670, 2010.

[37] A. B. Owen, "Self-concordance for empirical likelihood," *Can. J. Stat.*, vol. 41, no. 3, pp. 387–397, 2013.

[38] I. Necoara and J. Suykens, "Interior-point Lagrangian decomposition method for separable convex optimization," *J. Optim. Theory Appl.*, vol. 143, no. 3, 2009, Art. no. 567.

[39] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, vol. 87. Berlin, Germany: Springer Science & Business Media, 2013.

Konstantinos Gatsis (S'10–M'16) received the Ph.D. degree in electrical and systems engineering from the University of Pennsylvania, Philadelphia, PA, USA, in 2016.

He is currently a Postdoctoral Researcher with the Department of Electrical and Systems Engineering, University of Pennsylvania. His research interests include safe, secure, autonomous applications in the Internet-of-Things developing control, learning, communication, optimization, and security tools.

Dr. Gatsis was a recipient of the Joseph, D16, and Rosaline Wolf Award for Best Doctoral Dissertation from the Department of Electrical and Systems Engineering, University of Pennsylvania, the 2014 O. Hugo Schuck Best Paper Award, the Student Best Paper Award at the 2013 American Control Conference, and was a Best Paper Award Finalist at the 2014 ACM/IEEE International Conference on Cyber-Physical Systems.

George J. Pappas (F'09) is the Joseph Moore Professor and Chair of the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA, USA. He also holds a secondary appointment with the Departments of Computer and Information Sciences, and Mechanical Engineering and Applied Mechanics. He is member of the GRASP Lab and the PRECISE Center. He has previously served as the Deputy Dean for Research with the School of Engineering and Applied Science. His research focuses on control theory and in particular, hybrid systems, embedded systems, hierarchical and distributed control systems, with applications to unmanned aerial vehicles, distributed robotics, green buildings, and biomolecular networks. He was a recipient of various awards, such as the Antonio Ruberti Young Researcher Prize, the George S. Axelby Award, the O. Hugo Schuck Best Paper Award, the National Science Foundation PECASE, and the George H. Heilmeier Faculty Excellence Award.

Mark Eisen received the B.Sc. degree in electrical engineering from the University of Pennsylvania, Philadelphia, PA, USA, in 2014. He is now currently working toward the Ph.D. degree from the Department of Electrical and Systems Engineering, University of Pennsylvania. His research interests include distributed optimization and machine learning. In the summer of 2013, he was a research intern with the Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN, USA. Mr. Eisen was a recipient of the Outstanding Student Presentation at the 2014 Joint Mathematics Meeting, as well as the recipient of the 2016 Penn Outstanding Undergraduate Research Mentor Award.

Alejandro Ribeiro received the B.Sc. degree in electrical engineering from the Universidad de la Republica Oriental del Uruguay, Montevideo, Uruguay, in 1998 and the M.Sc. and Ph.D. degrees in electrical engineering from the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, USA, in 2005 and 2007, respectively. From 1998 to 2003, he was a member of the technical staff with Bellsouth, Montevideo, Uruguay. After his M.Sc. and Ph.D studies, in 2008, he joined the University of Pennsylvania, Philadelphia, PA, USA, where he is currently the Rosenbluth Associate Professor with the Department of Electrical and Systems Engineering. His research interests include applications of statistical signal processing to the study of networks and networked phenomena. His focus is on structured representations of networked data structures, graph signal processing, network optimization, robot teams, and networked control. Dr. Ribeiro is a Fulbright scholar and a Penn Fellow. He was a recipient of the 2014 O. Hugo Schuck best paper award, the 2012 S. Reid Warren, Jr. Award presented by Penn's undergraduate student body for outstanding teaching, the NSF CAREER Award in 2010, and paper awards at the 2016 SSP Workshop, 2016 SAM Workshop, 2015 Asilomar SSC Conference, ACC 2013, ICASSP 2006, and ICASSP 2005.