LEARNING STATISTICALLY ACCURATE RESOURCE ALLOCATIONS IN NON-STATIONARY WIRELESS SYSTEMS

Mark Eisen *Konstantinos Gatsis* *George J. Pappas* *Alejandro Ribeiro*

Department of Electrical and Systems Engineering, University of Pennsylvania

ABSTRACT

This paper considers the resource allocation problem in wireless systems over an unknown time-varying non-stationary channel. The goal is to maximize a utility function, such as a capacity function, over a set of wireless nodes while satisfying a set of resource constraints. To bypass the need for a model for channel distribution as it varies over time, samples of the channel are taken at every time epoch to estimate the channel. The resulting stochastic optimization problem is converted in its Lagrange dual problem, where the resulting stochastic optimization problem can viewed equivalently as minimizing a certain empirical risk measure, a well-studied problem in machine learning. The second order Newton's method is used to quickly learn statistically approximated optimal resource allocation policies over the sampled dual function as the channel evolves over time epochs. The quadratic convergence rate of Newton is used to establish, under certain conditions on the sampling size and rate of channel variation, an instantaneous learning and tracking of these policies. Numerical simulations demonstrate the effectiveness of the learning algorithm on a low-dimensional wireless capacity maximization problem.

***Index Terms*—** wireless communications, resource allocation, second order method, non-stationary channel

1. INTRODUCTION

Wireless machine-to-machine communication is increasingly used in cyber-physical applications to connect sensing devices for system monitoring, teams of robotic vehicles, and the Internet-of-Things. Traditionally models of wireless channel conditions and their distributions [1, Ch. 2-3] are required in order to operate these systems reliably and with optimal performance. For example, channel models facilitate to optimize the capacity of the wireless channel [1, Ch. 4] and allocate communication resources for wireless networks [2]. Apart from the common i.i.d. channel model (block fading), time-correlated models such as Markov models are frequently assumed [3–5]. Models also facilitate the allocation of communication resources to optimize closed loop control performance for sensor-actuator systems, e.g., [6–9], and wireless resource allocation problems [10–12].

Real world wireless systems operate under unpredictable channel conditions that may vary over time. Both a reliable model for the model and how it varies over time are often not known in practice, but can be observed via collected channel quality samples. Sampling can be used to learn a channel model directly through samples to employ a model-based approach. To bypass the need for modeling altogether, we may directly optimize system performance in a non-stationary environment through sampling [13, 14]. To achieve this we make a connection between the model-based design approach and an empirical risk minimization (ERM) problem, typical in machine learning. Furthermore, we employ a fast second order method to quickly adapt to time-varying channel models. This is motivated by the use of Newton's method to solve large scale ERM problems through adaptive sampling in [15, 16]. The authors previously applied a similar Newton-based learning approach to control systems [17].

In this paper the goal is to design resource allocation policies of a wireless communications system to maximize some utility function subject the constraints of the system, such as total resource availability. 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 demonstrate in Section 2 that the resource allocation problem can be converted in the Lagrangian dual problem, which is itself a stochastic program. By taking samples of the unknown distribution, the problem is equivalent to a empirical risk minimization problem (ERM). Because the optimal resource allocations change as the distribution changes, we present a second order learning method that can learn statistically accurate policies with a single iteration (Section 3). This convergence properties are analyzed formally in Section 4 and given an empirical demonstration in Section 5.

2. RESOURCE ALLOCATION AS EMPIRICAL RISK MINIMIZATION

We consider a generic setting in which we have a wireless network of m nodes indexed by $i = 1, \dots, m$. Given channel states $\mathbf{h} := [h^1, \dots, h^m] \in \mathcal{H}$, instantaneous transmission powers $\mathbf{p}(\mathbf{h}) := [p^1(h^1), \dots, p^m(h^m)]$ are selected for each node with which it will attempt to transmit a message. Further, we track the ergodic averages over time with variables $\mathbf{y} = [y^1, \dots, y^m]$, which are limited by the expectation of a vector function $\mathbf{F}(\mathbf{h}, \mathbf{p}(\mathbf{h}))$ of the channel states and instantaneous transmission powers. This function may represent, for example, the probability of successful transmission. The goal is to maximize a concave utility, or capacity, function $C(\mathbf{y})$ over the system, subject to the limit on ergodic averages and a second function concave $G(\mathbf{y})$ that captures total resource usage—see, e.g. [2] for details on this generic model. This can be formulated with the following stochastic optimization problem

$$[\mathbf{p}^*(\mathbf{h}), \mathbf{y}^*] := \operatorname*{argmax}_{\mathbf{p}, \mathbf{y} \in \mathbb{R}^m} C(\mathbf{y}) \quad \text{s.t.} \quad \mathbf{y} \leq \mathbb{E}_{\mathbf{h}} \left\{ \mathbf{F}(\mathbf{h}, \mathbf{p}(\mathbf{h})) \right\}, \quad 0 \leq G(\mathbf{y}). \tag{1}$$

Note that while $C(\mathbf{y})$ and $G(\mathbf{y})$ are concave, $\mathbf{F}(\mathbf{h}, \mathbf{p}(\mathbf{h})$ is not necessarily so. The problem in (1) is a constrained optimization problem over an infinite dimensional variable $\mathbf{p}(\mathbf{h})$ but can converted into the Lagrangian dual problem using a well-known construction. For notational convenience, define the augmented variables $\bar{\mathbf{F}}(\mathbf{h}, \mathbf{p}(\mathbf{h}), \mathbf{y}) := [\mathbf{F}(\mathbf{h}, \mathbf{p}(\mathbf{h})); G(\mathbf{y})]$ and $\tilde{\mathbf{y}} := [\mathbf{y}; 0]$. We form the Lagrangian as $\mathcal{L}(\mathbf{p}(\mathbf{h}), \mathbf{y}, \boldsymbol{\mu}) := C(\mathbf{y}) + \boldsymbol{\mu}^T [\mathbb{E}_{\mathbf{h}} \bar{\mathbf{F}}(\mathbf{h}, \mathbf{p}(\mathbf{h}), \mathbf{y}) - \tilde{\mathbf{y}}]$ using a dual variable $\boldsymbol{\mu} \in \mathbb{R}^{m+1}$. Define $\mathbf{p}(\mathbf{h}, \boldsymbol{\mu}) := \operatorname{argmax}_{\mathbf{p}} \mathcal{L}(\mathbf{p}(\mathbf{h}), \mathbf{y}, \boldsymbol{\mu})$ and $\mathbf{y}(\boldsymbol{\mu}) := \operatorname{argmax}_{\mathbf{y}} \mathcal{L}(\mathbf{p}(\mathbf{h}), \mathbf{y}, \boldsymbol{\mu})$ to be the optimal power allocations and ergodic averages for a given dual variable $\boldsymbol{\mu}$[1]. The dual function

Supported by ARL DCIST CRA W911NF-17-2-0181 and Intel Science and Technology Center for Wireless Autonomous Systems.

[1] Note than solving for $\mathbf{p}(\mathbf{h}, \boldsymbol{\mu})$ doesn't require $\mathbf{y}$ and visa versa.

$L(\boldsymbol{\mu})$ can then be written as

$$\boldsymbol{\mu}^* := \operatorname*{argmin}_{\boldsymbol{\mu}\geq 0} L(\boldsymbol{\mu}) := \operatorname*{argmin}_{\boldsymbol{\mu}\geq \mathbf{0}} \mathbb{E}_{\mathbf{h}} f(\boldsymbol{\mu}, \mathbf{h}), \tag{2}$$
$$\text{where } f(\boldsymbol{\mu}, \mathbf{h}) := C(\mathbf{y}(\boldsymbol{\mu})) + \boldsymbol{\mu}^T \left[\tilde{\mathbf{F}}(\mathbf{h}, \mathbf{p}(\mathbf{h}, \boldsymbol{\mu}), \mathbf{y}(\boldsymbol{\mu})) - \tilde{\mathbf{y}}(\boldsymbol{\mu})\right].$$

It can be shown that the problem in (1), while not strongly convex, nonetheless exhibits zero duality gap [2], thus implying that the optimal solution to (1) can be recovered exactly from (2) as $\mathbf{p}^*(\mathbf{h}) = \operatorname{argmax} \mathcal{L}(\mathbf{p}(\mathbf{h}), \boldsymbol{\mu}^*)$. We stress that $\mathbf{p}^*(\mathbf{h})$ is the optimal power allocation policy for a *fixed* channel distribution $\mathcal{H}$, meaning the problem in (2) must be continuously solved as the channel varies over time. This fact necessitates the need for fast solution methods to solve (2), which we describe in the following section.

Observe that the dual function $L(\boldsymbol{\mu})$ is a statistical loss function, specifically the expectation of a function $f(\boldsymbol{\mu}, \mathbf{h})$ over random variable $\mathbf{h}$. Problems of the form are common in machine learning, where a challenge exists in solving $L(\boldsymbol{\mu})$ when the distribution $\mathcal{H}$ is not known. This is indeed the case in most wireless resource allocation problems, where $\mathcal{H}$ represents the channel distribution. Rather than assume a specific model for $\mathcal{H}$, we may instead replace the *statistical* loss $L(\boldsymbol{\mu})$ with an *empirical* loss $\hat{L}(\boldsymbol{\mu})$. In empirical risk minimization (ERM) problems, the expectation over an unknown distribution is bypassed by taking an empirical average over a set of N samples. If we have N channel samples labelled $\mathbf{h}^1, \mathbf{h}^2, \ldots \mathbf{h}^N$, the empirical loss function is defined as

$$\hat{L}(\boldsymbol{\mu}) := \frac{1}{N} \sum_{l=1}^{N} f(\boldsymbol{\mu}, \mathbf{h}^l). \tag{3}$$

The goal of learning in the setting of wireless communication then becomes to solve, with some approximation, resource allocation problems through the sampling of an unknown channel to solve the statistical loss function derived from Lagrangian duality. The difference between the empirical loss $\hat{L}(\boldsymbol{\mu})$ and the dual loss $L(\boldsymbol{\mu})$ is well-studied in machine learning literature [18] and is captured in a term V, known as the *statistical accuracy*. The constant V_N is, in particular, a bound on the point-wise difference between the N-samples empirical and statistical loss with high probability, i.e. $\sup_{\boldsymbol{\mu}} \|\hat{L}(\boldsymbol{\mu}) - L(\boldsymbol{\mu})\| \leq V_N$. Common bounds for the statistical accuracy V_N can be obtained in the order of $\mathcal{O}(1/\sqrt{N})$ or, in some stronger cases, $\mathcal{O}(1/N)$ [18, 19]. In this work, we assume this constant is given or can be estimated easily. To employ a fast converging second order method that can be used in non-stationary settings, there are two issues inherent in the Lagrangian dual model to be corrected. The first is that the dual function $L(\boldsymbol{\mu})$—and $\hat{L}(\boldsymbol{\mu})$—is not strongly convex. The second is that the problem in (2) includes a non-negativity constraint $\boldsymbol{\mu} \geq 0$ due to the inequality constraints. We solve both of these problems through the use of regularizers. More specifically, we add the regularization term $\alpha V_N/2\|\boldsymbol{\mu}\|^2$ to the empirical risk in (3) to make the problem strongly convex. We also remove the non-negativity constraint by adding a logarithmic barrier regularizer. To preserve smoothness for small $\boldsymbol{\mu}$, however, we specifically use an ϵ-thresholded log function, defined as

$$\log_\epsilon(\boldsymbol{\mu}) := \begin{cases} \log(\boldsymbol{\mu}) & \boldsymbol{\mu} \geq \epsilon \\ \ell_{2,\epsilon}(\boldsymbol{\mu} - \epsilon) & \boldsymbol{\mu} < \epsilon, \end{cases} \tag{4}$$

where $\ell_{2,\epsilon}(\boldsymbol{\mu})$ is a second order Taylor series expansion of $\log(\boldsymbol{\mu})$ centered at ϵ for some small $0 < \epsilon < 1$. The second regularizer $-\beta V_N \mathbf{1}^T \log_\epsilon \boldsymbol{\mu}$ is then added to obtain a regularized empirical risk function

$$\hat{R}(\boldsymbol{\mu}) := \frac{1}{N} \sum_{l=1}^{N} f(\boldsymbol{\mu}, \mathbf{h}^l) + \frac{\alpha V_N}{2} \|\boldsymbol{\mu}\|^2 - \beta V_N \mathbf{1}^T \log_\epsilon \boldsymbol{\mu}. \tag{5}$$

The regularized function in (5) includes two additional terms compared to (3), each scaled by the statistical accuracy V_N and constants α and β. While the presence of such terms will modify the solution $\hat{R}^* := \min_{\boldsymbol{\mu}} \hat{R}(\boldsymbol{\mu})$ as compared to $\hat{L}^* = \min_{\boldsymbol{\mu}\geq 0} \hat{L}(\boldsymbol{\mu})$, we stress here that the value of interest is the original loss function minimizer $L^* = \min_{\boldsymbol{\mu}\geq \mathbf{0}} L(\boldsymbol{\mu})$. From the definition of statistical accuracy we have that $|L^* - \hat{L}^*| \leq V_N$, so any additional bias of order $\mathcal{O}(V_N)$ provides a negligible additional error. Both the quadratic and log-barrier regularizers, when scaled by V_N, are known to introduce biases of this order—see, e.g., [18,20] for details. Indeed, regularizers of this form are common in statistical learning problems of this form for this reason [18]. It then suffices to find a minimizer for $\hat{R}(\boldsymbol{\mu})$ rather than $\hat{L}(\boldsymbol{\mu})$, as the former is both strongly convex and can be minimizes with imposing the non-negativity constraint on $\boldsymbol{\mu}$.

Recall that the regularized empirical loss $\hat{R}(\boldsymbol{\mu})$ is derived from samples from a single channel distribution $\mathcal{H}$, and thus its minimizer will only recover an (approximately) optimal power allocation $\mathbf{p}^i(h)$ for this particular channel. In a non-stationary setting, we may instead consider that the distribution is indexed by a time epoch k (where we assume a time epoch is small enough where the channel $\mathcal{H}_k$ is fixed over that period). Then, at each k we can draw N samples $\mathbf{h}_k^1, \ldots, \mathbf{h}_k^N$ from $\mathcal{H}_k$ and construct a corresponding regularized empirical loss $\hat{R}_k(\boldsymbol{\mu})$. We are interested in recovering the optimal dual parameter $\boldsymbol{\mu}_k^*$ as

$$\boldsymbol{\mu}_k^* := \operatorname*{argmin}_{\boldsymbol{\mu}} \hat{R}_k(\boldsymbol{\mu}). \tag{6}$$

Observe that (6) would need to be solved at each time k, which is not necessarily feasible if the channel distribution is changing faster than the time it takes to find $\boldsymbol{\mu}_k^*$. In the following section, we develop a learning method that uses second order information to find approximate solutions to $\boldsymbol{\mu}^* k$ with a single iteration, thus making problems of this form practical. We conclude with a brief remark regarding the sampling on a non-stationary channel.

Remark 1 Observe that forming $\hat{R}_k$ requires drawing N new samples at each time epoch to achieve accuracy V_N. Reaching the desired accuracy may be hindered by the ability to draw many samples quickly or easily. One may consider, if the consecutive distributions are close, to alternatively keep $(M-1)N/M$ samples previously drawn the window $\mathcal{H}_{k-M+1}, \ldots, \mathcal{H}_{k-1}$ and draw only N/M new samples from $\mathcal{H}_k$. The exact bounds on the statistical accuracy achieved by $\hat{L}_k$ in this non-i.i.d. case are not well studied, so are no considered in this work, although this sampling approach can be used in practice to reduce the sampling burden at each time epoch.

3. SECOND ORDER LEARNING METHOD

We propose the use of second order information through Newton's method to find an approximate solution to (6) at each time epoch k with a single iteration. Recall that the exact solution $\boldsymbol{\mu}_k^*$ only solves the original problem of interest up to within the statistical accuracy V_N. As a consequence, there is no need to solve for $\boldsymbol{\mu}_k^*$ exactly as long as a solution $\boldsymbol{\mu}_k$ can be found that is also V_N. Consider updating $\boldsymbol{\mu}_k$ over each epoch $k = 0, 1, \ldots$ using the Newton update on the (6) as follows. First define the gradient $\nabla \hat{R}_k(\boldsymbol{\mu})$ and Hessian $\nabla^2 \hat{R}_{k+1}(\boldsymbol{\mu})$ as

$$\nabla \hat{R}_k(\boldsymbol{\mu}) = \frac{1}{N} \sum_{l=1}^{N} \left(\tilde{\mathbf{F}}(\mathbf{h}_k, \mathbf{p}(\mathbf{h}_k, \boldsymbol{\mu}), \mathbf{y}(\boldsymbol{\mu})) - \tilde{\mathbf{y}}(\boldsymbol{\mu})\right) + \alpha V_N \boldsymbol{\mu} - \beta V_N \boldsymbol{\mu}^{-1}. \tag{7}$$

$$\nabla^2 \hat{R}_k(\boldsymbol{\mu}) = \frac{1}{N} \sum_{l=1}^{N} \nabla_{\boldsymbol{\mu}} \left(\tilde{\mathbf{F}}(\mathbf{h}_k, \mathbf{p}(\mathbf{h}_k, \boldsymbol{\mu}), \mathbf{y}(\boldsymbol{\mu})) - \tilde{\mathbf{y}}(\boldsymbol{\mu})\right) + \alpha V_N \mathbf{I} + \beta V_N \text{diag}\{\boldsymbol{\mu}^{-2}\}. \tag{8}$$

Algorithm 1 Learning via Newton's Method

1: **Parameters:** Sample size increase constants $N_0 \geq 1$ backtracking params $0 < \delta < 1, \alpha, \beta$.
2: **Input:** Initial sample size $N = N_0$ and argument $\boldsymbol{\mu}_0$ $\|\nabla \hat{R}_0(\boldsymbol{\mu}_0)\| < (\sqrt{2\alpha})V_N$
3: **for** $k = 0, 1, 2, \ldots$ **do** {main loop}
4: Reset factor $N = N_0$.
5: **repeat** {sample size backtracking loop}
6: Draw N samples from $\mathcal{H}_{k+1}$.
7: Gradient $\nabla \hat{R}_{k+1}(\boldsymbol{\mu}_k)$, Hessian $\nabla^2 \hat{R}_{k+1}(\boldsymbol{\mu}_k)$ [cf. (7), (8)]:
8: Update [cf. (9)]: $\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k - \nabla^2 \hat{R}_{k+1}^{-1}(\boldsymbol{\mu}_k) \nabla \hat{R}_{k+1}(\boldsymbol{\mu}_k)$
9: Determine power allocation, ergodic averages:

$$\mathbf{p}(\mathbf{h}_{k+1}, \boldsymbol{\mu}_{k+1}) = \underset{p}{\operatorname{argmax}} \left\{ \boldsymbol{\mu}_{k+1}^T [\frac{1}{N} \sum_l \tilde{\mathbf{F}}(\mathbf{h}_{k+1}^l, \mathbf{p}(\mathbf{h}^l), \mathbf{y})] \right\}$$

$$\mathbf{y}(\boldsymbol{\mu}_{k+1}) = \underset{\mathbf{y}}{\operatorname{argmax}} \sum_{i=1}^{m} \mu_{k+1}^i y^i + \mu_{k+1}^{m+1} G(\mathbf{y})$$

10: Backtrack sample draw $N = \delta N$.
11: **until** $\|\nabla \hat{R}_{k+1}(\boldsymbol{\mu}_{k+1})\| < (\sqrt{2\alpha})V_N$
12: **end for**

Each dual parameter $\boldsymbol{\mu}_{k+1}$ is then updated from the previous parameter $\boldsymbol{\mu}_k$ using the Newton update as

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k - \nabla^2 \hat{R}_{k+1}(\boldsymbol{\mu}_k)^{-1} \nabla \hat{R}_{k+1}(\boldsymbol{\mu}_k). \tag{9}$$

Observe in the update in (9) that the gradients and Hessian are computed at the sampled loss function at time $k+1$, built from samples from the current distribution $\mathcal{H}_{k+1}$. It therefore requires the sampling of N samples from each distribution at each time epoch before performing the update (or using other sampling schemes, as discussed in Remark 1). The key observation in this update is that, if successive channel distributions $\mathcal{H}_k$ and $\mathcal{H}_{k+1}$ are sufficiently similar, i.e. the channel varies slowly, then the optimal solutions $\boldsymbol{\mu}_k^*$ and $\boldsymbol{\mu}_{k+1}^*$ will also be close together. Therefore, a V_N-approximate solution $\boldsymbol{\mu}_k$ to the loss function $\hat{R}_k(\boldsymbol{\mu})$ will be close to a V_N-approximate solution to $\hat{R}_{k+1}(\boldsymbol{\mu})$. We can consider the current iterate $\boldsymbol{\mu}_k$ as a "soft" start to finding a point $\boldsymbol{\mu}_{k+1}$ that approximately minimizes $\hat{R}_{k+1}(\boldsymbol{\mu})$, from which we find the a point that approximately minimizers $\hat{R}_{k+1}(\boldsymbol{\mu})$, and so forth. Assuming we start from a V_N-optimal point, at each subsequent iteration k we learn a near-optimal power allocation of the wireless channel for the current channel distribution.

The complete second order 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 $\mathcal{H}_{k+1}$ to form $\hat{R}_{k+1}$. The gradient $\nabla \hat{R}_{k+1}$ and Hessian $\mathbf{H}_{k+1}$ of the regularized dual loss function are computed in Step 7, after which the Newton step is taken to update $\boldsymbol{\mu}_{k+1}$ in Step 8. In Step 9, the optimal resource allocation variables $\mathbf{p}(\mathbf{h}, \boldsymbol{\mu}_{k+1})$ and $\mathbf{y}(\boldsymbol{\mu}_{k+1})$ are computed using the updated dual variable. Because the function and channel system parameters that are not known in practice, we include a backtracking step for the sample draw N in Step 10 to ensure the new iterate $\boldsymbol{\mu}_{k+1}$ is within the statistical accuracy V_N of $\hat{R}_{k+1}$.

4. CONVERGENCE ANALYSIS

We develop in this section an analysis of the second order learning method for the ERM problem in (6). Our primary result gives establishes conditions on the functions and channel distributions that allow for statistically accurate solutions to (6) for each time k using a single iteration of Newton's method. We first state two assumptions on the dual loss functions f.

Assumption 1 *The statistical loss function gradients $\nabla f(\boldsymbol{\mu}, \mathbf{z})$ 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 $\mathbf{h}$.*

Assumption 1 provides a smoothness condition for the dual functions, which can also be obtained by assuming strong concavity of the primal objective function $C(\mathbf{y})$. Because the dual function is always convex [20], it follows then that the regularized empirical loss gradients $\nabla \hat{R}_k$ are Lipschitz continuous with constant $\Delta + cV_N$ where $c := \alpha + \beta/\epsilon^2$ and the function $\hat{R}_k$ is strongly convex with constant αV_N. Assumption 2 establishes self concordance of the loss functions, which is a customary assumption in the analysis of second-order methods. This implies self concordance of $\hat{R}_{k+1}$ because both the quadratic and thresholded log regularizers are self-concordant. We include two additional assumptions.

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

$$\sup_{\boldsymbol{\mu}} \|\nabla L_k(\boldsymbol{\mu}) - \nabla \hat{L}_k(\boldsymbol{\mu})\| \leq V_N^{1/2}, \qquad \textit{w.h.p.} \tag{10}$$

Assumption 4 *The difference between two successive expected loss $L_k(\boldsymbol{\mu}) = \mathbb{E}_{h_k} f(\boldsymbol{\mu}, \mathbf{h}_k)$ and $L_{k+1}(\boldsymbol{\mu}) = \mathbb{E}_{h_{k+1}} f(\boldsymbol{\mu}, \mathbf{h}_{k+1})$ and the difference between gradients are bounded respectively by a bounded sequence of constants $\{D_k\}, \{\bar{D}_k\} \geq 0$ for all $\boldsymbol{\mu}$,*

$$\sup_{\boldsymbol{\mu}} |L_k(\boldsymbol{\mu}) - L_{k+1}(\boldsymbol{\mu})| \leq D_k, \tag{11}$$

$$\sup_{\boldsymbol{\mu}} \|\nabla L_k(\boldsymbol{\mu}) - \nabla L_{k+1}(\boldsymbol{\mu})\| \leq \bar{D}_k. \tag{12}$$

In Assumption 3, we bound the difference between gradients of the expected loss and the empirical loss with N samples by $V_N^{1/2}$, which can be derived from the law of large numbers. Assumption 4 bounds the difference in the expected loss functions and gradients at epochs k and $k+1$ by constants D_k and $\bar{D}_k$. This effectively provides a limit on the rate at which the channel evolves between epochs, and is necessary to establish relative closeness of $\hat{R}_k$ and $\hat{R}_{k+1}$.

Our goal is to find conditions on the sampling and the non-stationarity parameters under which a single step of Newton's method generates a statistically accurate minimizer to $\hat{R}_{k+1}$ when starting from a statistically accurate minimizer to $\hat{R}_k$. This is done in two steps. We first derive a condition under which V_N-accurate solution to $\hat{R}_k$, labelled $\boldsymbol{\mu}_k$, is in the local region of $\hat{R}_{k+1}$ where convergence is quadratic with Newton's method. To characterize this region, we use a quantity called the Newton decrement, defined as $\lambda_{k+1}(\boldsymbol{\mu}) := \|\nabla^2 \hat{R}_{k+1}(\boldsymbol{\mu})^{-1/2} \nabla \hat{R}_{k+1}(\boldsymbol{\mu})\|$. We say the dual iterate $\boldsymbol{\mu}$ is in the quadratic convergence region of $\hat{R}_{k+1}$ when $\lambda_{k+1}(\boldsymbol{\mu}) < 1/4$—see [20, Chapter 9.6.4] for details. The conditions necessary to for this to hold are established in the following lemma.

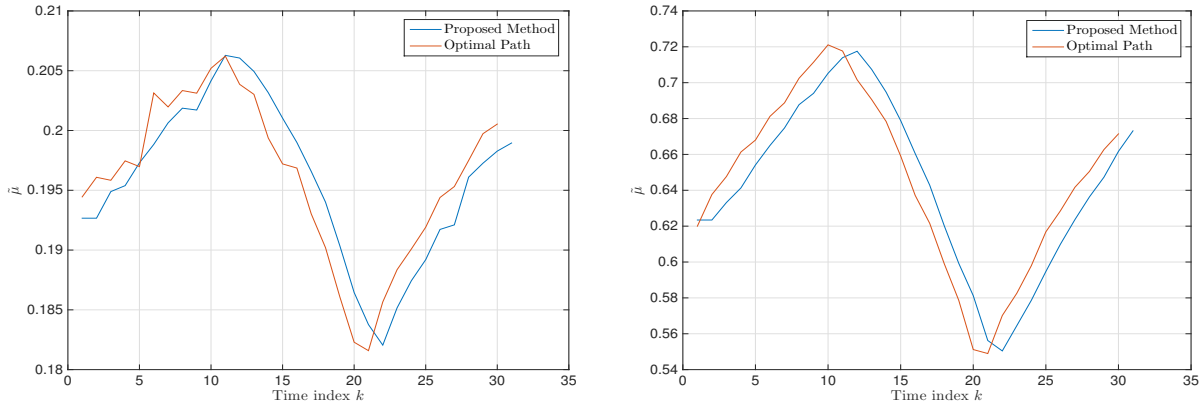

Fig. 1: Convergence paths of optimal values vs. values generated by the proposed second order learning method for time-varying $\mathcal{H}_k$ for dual parameter in a resource-constrained capacity maximization problem with $m = 4$ (left) and $m = 20$ (right) nodes. Newton's method is able to find an approximately optimal value for the dual variable at each iteration.

Lemma 1 *Define $\boldsymbol{\mu}_k$ as a V_N-accurate optimal solution of the loss $\hat{R}_k$, i.e., $\hat{R}_k(\boldsymbol{\mu}_k) - \hat{R}_k(\boldsymbol{\mu}_k^*) \leq V_N$. If Assumptions 1-4 hold, then Newton's method at point $\boldsymbol{\mu}_k$ is in the quadratic convergence phase for the objective function $\hat{R}_{k+1}$, i.e., $\lambda_{k+1}(\boldsymbol{\mu}_k) < 1/4$, if we have*

$$\left(\frac{2(\Delta + cV_N)V_N}{\alpha V_N}\right)^{1/2} + \frac{2V_N^{1/2} + \bar{D}_k}{(\alpha V_N)^{1/2}} < \frac{1}{4}. \quad \textit{w.h.p.} \tag{13}$$

The second step is to establish conditions under which a point in the quadratic convergence region of $\hat{R}_{k+1}$ will reach its statistical accuracy V_N with a single Newton step as given in (9). We state this condition in the proceeding lemma.

Lemma 2 *Consider $\boldsymbol{\mu}_k$ to be in the quadratic neighborhood of the loss $\hat{R}_{k+1}$, i.e., $\lambda_{k+1}(\boldsymbol{\mu}_k) \leq 1/4$. Recall the definition of the variable $\boldsymbol{\mu}_{k+1}$ in* (9) *as the updated variable using Newton's method. If Assumptions 1-3 hold, then the difference $\hat{R}_{k+1}(\boldsymbol{\mu}_{k+1}) - \hat{R}_{k+1}^*$ is upper bounded by*

$$\hat{R}_{k+1}(\boldsymbol{\mu}_{k+1}) - \hat{R}_{k+1}^* \leq 144(5V_N + 2D_k)^2. \tag{14}$$

With the previous two Lemmata we establish a conditions for a i.) current iterate $\boldsymbol{\mu}_k$being in the quadratic convergence region of subsequent empirical loss function $\hat{R}_{k+1}$ and ii.) an upper bound on the suboptimality of the updated iterate $\boldsymbol{\mu}_{k+1}$ from within this quadratic region. From these two results, the conditions necessary for reaching statistical accuracy with single updates follows, as presented in our primary theorem.

Theorem 1 *Consider Newton's method defined in* (9)*. Define V_N to be the statistical accuracy of $\hat{L}_k$ after taking N samples. Further consider the variable $\boldsymbol{\mu}_k$ as a V_N-optimal solution of the loss $\hat{R}_k$, and suppose Assumptions 1-4 hold. If the following conditions*

$$\left(\frac{2(\Delta + cV_N)V_N}{\alpha V_N}\right)^{1/2} + \frac{2V_N^{1/2} + \bar{D}_k}{(\alpha V_N)^{1/2}} < \frac{1}{4} \tag{15}$$

$$144(5V_N + 2D_k)^2 \leq V_N \tag{16}$$

are satisfied, then the variable $\boldsymbol{\mu}_{k+1}$ computed from (9) *has the suboptimality of V_N with high probability, i.e.,*

$$\hat{R}_{k+1}(\boldsymbol{\mu}_{k+1}) - \hat{R}_{k+1}^* \leq V_N, \qquad \textit{w.h.p.} \tag{17}$$

The expressions in (15) and (16) give conditions on the statistical accuracy V_N (controlled by sampling rate N) and bounds on nonstationarity D_k and $\bar{D}_k$ provide us V_N-accurate updates. We conclude by stressing that this is a theoretical results, and these parameters may not be known in real applications. In practice, one can use a backtracking step (as done in Algorithm 1) to control parameters such as N and c to achieve statistical accuracy.

5. SIMULATION RESULTS

We simulate the performance of our second order learning method on a simple wireless capacity maximization problem with resource constraints. Consider the capacity function for the ith node $C_i(p^i(h)) = u\log(1 + hp^i(h)/v)$ for some positive constants u, v. The total capacity to be maximized is then written as

$$C(\mathbf{p}(\mathbf{h})) := \sum_{i=1}^{N} u\log(1 + hp^i(h)/v). \tag{18}$$

Note that here we forego the ergodic averages for simplicity and seek to maximize $\mathbb{E}_{\mathbf{h}}C(\mathbf{p}(\mathbf{h}))$ over the instantaneous power allocation variables. The channel states at epoch k are drawn from an exponential distribution with mean w_k. To model a time-varying channel, we slowly vary w_k for different epochs k. We draw $N = 1000$ samples at each epoch and set $u = 1$ and $v = 10$. Furthermore, we impose the resource constraint $\mathbb{E}_{\mathbf{h}} \sum_i p^i(h^i) \leq p_{\max}$ for some power budget $p_{\max}$.

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 (6) with the defined capacity function $C(\cdot)$, channel distributions $\mathcal{H}_k$, with previously mentioned resource constraint. In Figure 1 we show the path of Newton's method at each epoch k for the dual variable μ_k for a small number of nodes $m = 4$ and large number of nodes $m = 20$. The red line of each figure plots the optimal values for the current distribution parameter u_k as it changes with k. The blue line plots the values generated by the proposed method over epochs. The channel evolves at over epochs by a fixed rate $u_{k+1} = u_k \pm r$ for some rate r. Within some small error, Newton's method is indeed able to quickly and approximately find each new solution as the channel varies over time with single iterations for both the small and larger wireless networks.

6. CONCLUSION

In this paper we develop second order learning method to find statistically approximated optimal resource allocation policies to use in a nonstationary wireless network. We apply Lagrangian duality to derive a parametrized model for such policies that reduces to minimizing a statistical loss function. Because the channel distribution is unknown and varies, we collect channel samples to formulate the resource allocation problem as an empirical risk minimization (ERM) problem. We establish conditions under which a Newton update can, with a single iteration, find approximately optimal resource allocations as the channel distribution changes. This is further exemplified with a numerical simulation.

7. REFERENCES

[1] Andrea Goldsmith, *Wireless communications*, Cambr. Univ. Press, 2005.

[2] Alejandro Ribeiro, "Ergodic stochastic optimization algorithms for wireless communication and networking," *IEEE Transactions on Signal Processing*, vol. 58, no. 12, pp. 6369–6386, 2010.

[3] Mordechai Mushkin and Israel Bar-David, "Capacity and coding for the gilbert-elliott channels," *IEEE Transactions on Information Theory*, vol. 35, no. 6, pp. 1277–1290, 1989.

[4] Sahand Haji Ali Ahmad, Mingyan Liu, Tara Javidi, Qing Zhao, and Bhaskar Krishnamachari, "Optimality of myopic sensing in multichannel opportunistic access," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 4040–4050, 2009.

[5] Wenzhuo Ouyang, Atilla Eryilmaz, and Ness B Shroff, "Asymptotically optimal downlink scheduling over markovian fading channels," in *INFOCOM, 2012 Proceedings IEEE*. IEEE, 2012, pp. 1224–1232.

[6] Konstantinos Gatsis, Alejandro Ribeiro, and George J. Pappas, "Optimal power management in wireless control systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1495–1510, 2014.

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

[8] José Araújo, Manuel Mazo, Adolfo Anta, Paulo Tabuada, and Karl H Johansson, "System architectures, protocols and algorithms for aperiodic wireless control systems," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 175–184, 2014.

[9] Mohammad H Mamduhi, Domagoj Tolić, Adam Molin, and Sandra Hirche, "Event-triggered scheduling for stochastic multi-loop networked control systems with packet dropouts," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*. IEEE, 2014, pp. 2776–2782.

[10] Guocong Song and Ye Li, "Cross-layer optimization for ofdm wireless networks-part i: theoretical framework," *IEEE transactions on wireless communications*, vol. 4, no. 2, pp. 614–624, 2005.

[11] Guocong Song and Ye Li, "Cross-layer optimization for ofdm wireless networks-part ii: algorithm development," *IEEE Transactions on Wireless Communications*, vol. 4, no. 2, pp. 625–634, 2005.

[12] Jianwei Huang, Vijay G Subramanian, Rajeev Agrawal, and Randall Berry, "Joint scheduling and resource allocation in uplink ofdm systems for broadband wireless access networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 2, pp. 226–234, 2009.

[13] Konstantinos Gatsis, Miroslav Pajic, Alejandro Ribeiro, and George J. Pappas, "Opportunistic control over shared wireless channels," *IEEE Transactions on Automatic Control*, vol. 60, no. 12, pp. 3140–3155, December 2015.

[14] Konstantinos Gatsis, Alejandro Ribeiro, and George J. Pappas, "Random access design for wireless control systems," 2016, Submitted. Available on Arxiv.

[15] Aryan Mokhtari, Hadi Daneshmand, Aurelien Lucchi, Thomas Hofmann, and Alejandro Ribeiro, "Adaptive newton method for empirical risk minimization to statistical accuracy," in *Advances in Neural Information Processing Systems*, 2016, pp. 4062–4070.

[16] Mark Eisen, Aryan Mokhtari, and Alejandro Ribeiro, "Large scale empirical risk minimization via truncated adaptive newton method," *arXiv preprint arXiv:1705.07957*, 2017.

[17] Mark Eisen, Konstantinos Gatsis, George J. Pappas, and Alejandro Ribeiro, "Learning in non-stationary wireless control systems via newton's method," 2017, Available at http://www.seas.upenn.edu/ $\sim$ maeisen/wiki/wirelesscontrol.pdf.

[18] Olivier Bousquet and Léon Bottou, "The tradeoffs of large scale learning," in *Advances in neural information processing systems*, 2008, pp. 161–168.

[19] Vladimir Vapnik, *The nature of statistical learning theory*, Springer science & business media, 2013.

[20] Stephen Boyd and Lieven Vandenberghe, *Convex optimization*, Cambridge university press, 2004.