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 be 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 quickly learn statistically approximated policies with a single iteration (Section 3). This convergence properties are analyzed formally in Section 4 and given an empirical demonstration in Section 5.

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 an 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 $i = 1, \ldots, m$. Given channel states $h := [h^1, \ldots, h^m] \in H$, instantaneous transmission powers $p(h) := \{p^h(h^1), \ldots, p^m(h^m)\}$ are selected for each node with which it will attempt to transmit a message. Further, we track the ergodic averages over time with variables $y = [y^1, \ldots, y^m]$, which are limited by the expectation of a vector function $F(h, p(h))$ 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 $C(y)$ over the system, subject to the limit on ergodic averages and a second function concave $G(y)$ 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

$$[p^*(h), y^*] := \arg\max_{p, y} C(y)$$

subject to 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 an 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.

Supported by ARL DCIST CRA W911NF-17-2-0181 and Intel Science and Technology Center for Wireless Autonomous Systems.
\( L(\mu) \) can then be written as
\[
\mu^* := \arg\min_{\mu \geq 0} L(\mu) := \arg\min_{\mu \geq 0} E_{hf}(\mu,\mu), \quad (2)
\]
where \( f(\mu, h) := C(y(\mu)) + \alpha V_{N\mu}^2 - \beta V_{N\mu-1} \).

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 \( p^*(h) = \arg\max \mathcal{L}(p(h), \mu^*) \). We stress that \( p^*(h) \) is the optimal power allocation policy for a fixed channel distribution \( \mathcal{H} \), meaning the problem (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 \( L(\mu) \) is a statistical loss function, specifically the expected value of a function \( f(\mu, h) \) over random variable \( h \). Problems of the form are common in machine learning, where a challenge exists in solving \( L(\mu) \) 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 \( L(\mu) \) with an empirical loss \( \hat{L}(\mu) \). 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 \( h^1, h^2, \ldots, h^N \), the empirical loss function is defined as
\[
\hat{L}(\mu) := \frac{1}{N} \sum_{l=1}^{N} f(\mu, h^l). \quad (3)
\]

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 \( \hat{L}(\mu) \) and the dual loss \( L(\mu) \) is well-studied in machine learning literature [18] and is captured in a term \( V_N \), 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. \( \sup_{\mu} \| L(\mu) - \hat{L}(\mu) \| \leq V_N \).

Common bounds for the statistical accuracy \( V_N \) can be obtained in the order of \( O(1/\sqrt{N}) \) or, in some stronger cases, \( O(1/N) \) [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 \( L(\mu) \)—and \( \hat{L}(\mu) \)—is not strongly convex. The second is that the problem in (2) includes a non-negativity constraint \( \mu \geq 0 \) due to the inequality constraints. We solve both of these problems through the use of regularizers. More specifically, we add the regularization term \( \alpha V_N \mu \) 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 \( \mu \), however, we specifically use an \( \epsilon \)-thresholded log function, defined as
\[
\log_{\epsilon}(\mu) := \begin{cases} 
\log(\mu) & \mu \geq \epsilon \\
\ell_2(\mu - \epsilon) & \mu < \epsilon,
\end{cases}
\]
where \( \ell_2(\mu - \epsilon) \) is a second order Taylor series expansion of \( \log(\mu) \) centered at \( \epsilon \) for some small \( 0 < \epsilon < 1 \). The second regularizer \(-\beta V_N \mu \) is then added to obtain a regularized empirical risk function
\[
\hat{R}(\mu) := \frac{1}{N} \sum_{l=1}^{N} f(\mu, h^l) + \frac{\alpha V_N}{2} \| \mu \|^2 - \beta V_N \mu^T \log_{\epsilon}(\mu). \quad (5)
\]
The regularized function in (5) includes two additional terms compared to (3), each scaled by the statistical accuracy \( V_N \) and constants \( \alpha \) and \( \beta \). While the presence of such terms will modify the solution \( \hat{R}^* := \min_{\mu} \hat{R}(\mu) \) as compared to \( L^* = \min_{\mu \geq 0} L(\mu) \), we stress here that the value of interest is the original loss function minimizer \( L^* = \min_{\mu \geq 0} L(\mu) \). From the definition of statistical accuracy we have that \( |L^* - \hat{L}^*| \leq V_N \), so any additional bias of order \( O(V_N) \) 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 \( \hat{R}(\mu) \) rather than \( L(\mu) \), as the former is both strongly convex and can be minimized with imposing the non-negativity constraint on \( \mu \).

Recall that the regularized empirical loss \( \hat{R}(\mu) \) is derived from samples from a single channel distribution \( \mathcal{H} \), and thus its minimizer will only recover an (approximately) optimal power allocation \( p^*(h) \) 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 new samples \( h^1_k, \ldots, h^N_k \) from \( \mathcal{H}_k \) and construct a corresponding regularized empirical loss \( \hat{R}_k(\mu) \). We are interested in recovering the optimal dual parameter \( \mu_k^* \) as
\[
\mu_k^* := \arg\min_{\mu} \hat{R}_k(\mu).\quad (6)
\]

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 \( \mu_k^* \). In the following section, we develop a learning method that uses second order information to find approximate solutions to \( \mu_k \) 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 \( L_k \) in this non-i.i.d. case are not well studied, so are not 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 \( \mu_k^* \) 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 \( \mu_k^* \) exactly as long as a solution \( \mu_k \) can be found that is also \( V_N \). Consider updating \( \mu_k \) over epoch \( k = 0, 1, \ldots \) using the Newton update on the (6) as follows.

First define the gradient \( \nabla \hat{R}_k(\mu) \) and Hessian \( \nabla^2 \hat{R}_k(\mu) \) as
\[
\nabla \hat{R}_k(\mu) = \frac{1}{N} \sum_{l=1}^{N} \left( \mathbf{F}(h_k, p(h_k, \mu), y(\mu)) - \hat{y}(\mu) \right) + \alpha V_N \mu - \beta V_N \mu^{-1}.\quad (7)
\]
\[
\nabla^2 \hat{R}_k(\mu) = \frac{1}{N} \sum_{l=1}^{N} \nabla_{\mu} \left( \mathbf{F}(h_k, p(h_k, \mu), y(\mu)) - \hat{y}(\mu) \right) + \alpha V_N \mathbf{I} + \beta V_N \text{diag}(\mu^{-2}).\quad (8)
\]
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 $\mu_0$.
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 $H_{k+1}$.
7: Gradient $\nabla \hat{R}_{k+1}(\mu_k)$, Hessian $\nabla^2 \hat{R}_{k+1}(\mu_k)$ [cf. (7), (8)];
8: Update [cf. (9)]: $\mu_{k+1} = \mu_k - \nabla^2 \hat{R}_{k+1}(\mu_k)\nabla \hat{R}_{k+1}(\mu_k)$
9: Determine power allocation, ergodic averages:

$$p(h_{k+1}, \mu_{k+1}) = \arg\max_p \left\{ \mu_{k+1}^T \left[ \frac{1}{N} \sum_i \hat{F}(h_{k+1}, p(h_i), y) \right] \right\}$$

$$y(\mu_{k+1}) = \arg\max_y \left\{ \sum_{i=1}^m \mu_{k+1} y_i + \mu_0^{m+1} G(y) \right\}$$
10: Backtrack sample draw $N = \delta N$.
11: until $|\nabla \hat{R}_{k+1}(\mu_{k+1})| < (\sqrt{2\alpha})V_N$
12: end for

Each dual parameter $\mu_{k+1}$ is then updated from the previous parameter $\mu_k$ using the Newton update as

$$\mu_{k+1} = \mu_k - \nabla^2 \hat{R}_{k+1}(\mu_k)^{-1} \nabla \hat{R}_{k+1}(\mu_k).$$

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 $H_{k+1}$. 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 $H_k$ and $H_{k+1}$ are sufficiently similar, i.e., the channel varies slowly, then the optimal solutions $\mu_k$ and $\mu_{k+1}$ will also be close together. Therefore, a $V_N$-approximate solution $\mu_k$ to the loss function $R_k(\mu)$ will be close to a $V_N$-approximate solution to $R_{k+1}(\mu)$. We can consider the current iterate $\mu_k$ as a "soft" start to finding a point $\mu_{k+1}$ that approximately minimizes $\hat{R}_{k+1}(\mu)$, from which we find the a point that approximately minimizers $\hat{R}_{k+1}(\mu)$, 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 $H_{k+1}$ to form $\hat{R}_{k+1}$. The gradient $\nabla \hat{R}_{k+1}$ and Hessian $H_{k+1}$ of the regularized dual loss function are computed in Step 7, after which the Newton step is taken to update $\mu_{k+1}$ in Step 8. In Step 9, the optimal resource allocation variables $p(h, \mu_{k+1})$ and $y(\mu_{k+1})$ 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 $\mu_{k+1}$ 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 $\nabla f(\mu, \chi)$ are Lipschitz continuous with constant $\Delta$.

**Assumption 2** The loss functions $f(\mu, \chi)$ are self-concordant with respect to $\mu$ for all $\chi$.

Assumption 1 provides a smoothness condition for the dual functions, which can also be obtained by assuming strong concavity of the primal objective function $C(y)$. Because the dual function is always convex [20], it follows then that the regularized empirical loss gradients $\nabla R_k$ are Lipschitz continuous with constant $\Delta + cV_N$ where $c := \alpha + \beta/c^2$ and the function $R_k$ is strongly convex with constant $\alpha V_N$.

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 $L_k$ and the statistical average loss $\hat{L}_k$ is bounded by $\sqrt{\lambda N}$ for all $\mu$ and $k$ with high probability,

$$\sup_\mu \|\nabla L_k(\mu) - \nabla \hat{L}_k(\mu)\| \leq \sqrt{\lambda N}, \text{ w.h.p.}$$

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

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

$$\sup_\mu \|\nabla L_k(\mu) - \nabla L_{k+1}(\mu)\| \leq \hat{D}_k.$$
Proposed Method
Optimal Path
Time index $k$
0 5 10 15 20 25 30 35

Fig. 1: Convergence paths of optimal values vs. values generated by the proposed second order learning method for time-varying $H_k$ 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.

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 $i$th node $C(p(h)) = u \log(1 + hp(h)/v)$ for some positive constants $u, v$. The total capacity to be maximized is then written as

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

Note that here we forego the ergodic averages for simplicity and seek to maximize $E_k C(p(h))$ 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 $\sum_{i=1}^{N} p_i(h) \leq p_{\max}$ 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 $H_k$, with previously mentioned resource constraint. In Figure 1 we show the path of Newton’s method at each epoch $k$ for the dual variable $\mu_k$ 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 $u_{k+1} = u_k + r$ 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 methods to find statistically approximated optimal resource allocation policies to use in a non-stationary 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


