Stochastic Network Optimization with Non-Convex Utilities and Costs

Michael J. Neely

Abstract— **This work considers non-convex optimization of time averages of network attributes in a general stochastic network. This includes maximizing a non-concave utility function of the time average throughput vector in a time-varying wireless system, subject to network stability and to an additional collection of time average penalty constraints. We develop a simple algorithm that meets all desired stability and penalty constraints, and, subject to a convergence assumption, yields a time average vector that is a local optimum of the desired utility function. We also consider algorithms that yield "local near optimal" solutions, where the distance to a local optimum can be made as small as desired with a corresponding tradeoff in average delay. Our solution uses Lyapunov optimization with a combination of stochastic dual and primal-dual techniques. We also discuss the relative advantages and disadvantages of these techniques.**

Index Terms— **Queueing analysis, opportunistic scheduling, flow control, wireless networks**

I. Introduction

We consider a queueing network that operates in discrete time $t \in \{0, 1, 2, \ldots\}$. Let $\boldsymbol{Q}(t) = (Q_1(t), \ldots, Q_K(t))$ be the vector of queue backlogs on slot t. The arrival and service of the queues are determined by a *random event* $\omega(t)$ and a *network control action* $\alpha(t)$, chosen in reaction to this event from an abstract set of possible actions $\mathcal{A}_{\omega(t)}$ that possibly depends on $\omega(t)$. The $\alpha(t)$ and $\omega(t)$ values also affect two *attribute vectors* $\boldsymbol{x}(t) = (x_1(t), \ldots, x_M(t))$, $\boldsymbol{y}(t) = (y_0(t), y_1(t), \ldots, y_L(t))$ according to general (possibly non-convex, discontinuous) functions $\hat{x}_m(\cdot)$, $\hat{y}_l(\cdot)$ for $m \in \{1, \ldots, M\}$ and $l \in \{0, 1, \ldots, L\}$:

$$x_m(t) = \hat{x}_m(\alpha(t), \omega(t)) \quad , \quad y_l(t) = \hat{y}_l(\alpha(t), \omega(t))$$

Define $\overline{\boldsymbol{x}} = (\overline{x}_1, \ldots, \overline{x}_M)$ as the limiting time average expectation of the attribute vector $\boldsymbol{x}(t)$ under a particular policy (temporarily assumed to exist):

$$\overline{\boldsymbol{x}} \triangleq \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \mathbb{E}\left\{\boldsymbol{x}(\tau)\right\}$$

Define time average expectations $\overline{y}_l$ similarly. The goal is to determine an algorithm for choosing actions $\alpha(t) \in \mathcal{A}_{\omega(t)}$ over time to solve the following general stochastic network optimization problem:

Michael J. Neely is with the Electrical Engineering department at the University of Southern California, Los Angeles, CA. (web: http://www-rcf.usc.edu/∼mjneely).

This material is supported in part by one or more of the following: the DARPA IT-MANET program grant W911NF-07-0028, the NSF Career grant CCF-0747525, and continuing through participation in the Network Science Collaborative Technology Alliance sponsored by the U.S. Army Research Laboratory.

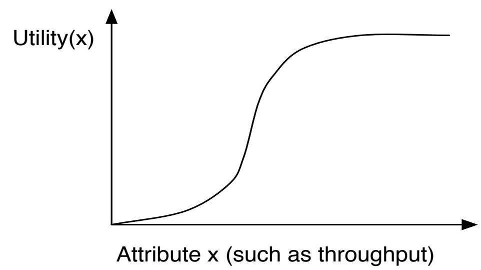

Fig. 1. An example non-concave utility for throughput maximization.

$$\text{Minimize:} \quad \overline{y}_0 + f(\overline{\boldsymbol{x}}) \tag{1}$$

$$\text{Subject to:} \quad \overline{y}_l + g_l(\overline{\boldsymbol{x}}) \leq 0 \;\; \forall l \in \{1, \ldots, L\} \tag{2}$$

$$\overline{\boldsymbol{x}} \in \mathcal{X} \tag{3}$$

$$\text{All queues } Q_k(t) \text{ are mean rate stable} \tag{4}$$

$$\alpha(t) \in \mathcal{A}_{\omega(t)} \;\; \forall t \tag{5}$$

where $g_l(\boldsymbol{x})$ are continuous and convex functions of $\boldsymbol{x} \in \mathcal{R}^M$ for each $l \in \{1, \ldots, L\}$, $f(\boldsymbol{x})$ is a continuous but possibly non-convex function of $\boldsymbol{x} \in \mathcal{R}^M$, and $\mathcal{X}$ is a convex subset of $\mathcal{R}^M$. The definition of "mean rate stable" will be given in Section II.

This non-convex problem has applications in network utility maximization and in other areas of network science. For example, suppose we desire to maximize a non-concave utility function of throughput in a wireless network. An example "sigmoidal-type" utility function is shown in Fig. 1, which has a "flat" area near zero, illustrating that we earn small utility unless throughput crosses a threshold. Constrained optimization of a sum of such utility functions is a difficult problem that can be shown, in some cases, to be at least as difficult as combinatorial bin-packing. Hence, we do not attempt to find a global solution. Such non-convex problems are studied for non-stochastic (static) networks in [1], where difficulties are discussed and a heuristic algorithm is derived that has optimality properties in a limit of a large number of users. Further treatment of non-convex optimization in a static context, including efficient heuristics, are provided in [2].

In this paper, we consider non-convex optimization for *stochastic networks* via the general problem (1)-(5). We show that, under some convergence assumptions, we can find a local optimum. In the special case when $f(\boldsymbol{x})$ is convex, we provide a stronger result that shows the global solution can be approached to within any desired accuracy. Our techniques extend our prior "drift-plus-penalty" or "dual-based" approach to stochastic network optimization in [3], which uses Lyapunov optimization, virtual queues, and auxiliary variables. These techniques alone would allow one to minimize (1) in the case $f(\boldsymbol{x})$ is convex, subject to constraints (2), (4), (5). We introduce an abstract set constraint (3), which we believe is new even in the convex case (we recently also used this in a "universal scheduling" context in [4]). However, if these dual-based techniques were employed for a non-convex function $f(\boldsymbol{x})$, the algorithm would approach a global optimum of the time average of $f(\boldsymbol{x}(t))$, which is not necessarily even a local optimum of $f(\overline{\boldsymbol{x}})$. In this paper, we achieve local optimality by combining these techniques with a "primal-dual based" approach similar to that of [5][6][7].

It is known that, in the convex case, the dual approach is robust to system changes [3][8][4] and provides direct bounds on utility, delay, and convergence time [3][9][10][11]. Unfortunately, such properties are less clear with the primal-dual approach, particularly for the non-convex case, and the proof we provide in this paper for the non-convex case does not specify the time required for convergence. This illustrates some of the challenges involved with non-convex optimization, even if we only want to find a local optimum. We partially address this issue in the special case of maximizing non-convex utility functions that are entrywise non-decreasing. In this case, we briefly describe an alternative 3-phase algorithm that explores the attribute space and ties all stochastic decisions directly to the result of a static non-convex optimization. The advantage here is that the stochastic phases are convex and have strong convergence guarantees, while the non-convex optimization is isolated to a static problem for which any static solver can be employed.

The method of Lyapunov drift to solve queue stability problems (i.e., problems that involve only the constraint (4)) was pioneered by Tassiulas and Ephremides in [12][13], where it was shown that greedy actions that minimize a drift expression $\Delta(t)$ every slot guarantee stability whenever possible. In [9][3][10][11], we extended this result to provide joint stability and performance optimization (or "penalty minimization"), using a greedy action to minimize a drift-plus-penalty metric $\Delta(t) + V \cdot penalty(t)$ every slot, where V is a non-negative parameter that affects a performance-delay tradeoff. Choosing $V = 0$ reduces to the original Tassiulas-Ephremides approach, while increasing V maintains network stability while pushing the time average penalty of the problem to within $O(1/V)$ of optimal, with a tradeoff in average queue backlog that is $O(V)$. This method can be applied to multi-hop networks with general constraints, and is known to be related to dual subgradient algorithms when applied to static convex programs [3]. A related method, also of the dual type, is presented for a wireless downlink in [14], although the fluid model arguments there do not prove the $[O(1/V), O(V)]$ performance-delay properties.

A different "primal-dual" approach is considered for the special case of utility-optimal opportunistic scheduling in a wireless downlink in [5][6]. This work assumes all users always have data to send, and hence does not consider randomly arriving traffic or queue stability issues. It also uses an infinite horizon time-average to inform scheduling decisions, which, unlike the dual method, makes it difficult to apply to systems with parameters that might change. Joint queue stability and performance optimization is treated for multi-hop networks with this primal-dual approach in [7]. This work proves that the utility of a related fluid network is optimal, and conjectures that the utility of the actual network will also be close to optimal. Specifically, the infinite horizon time average of [5][6] is replaced by an exponentially weighted average in [7], and it is conjectured that utility approaches optimality when the exponential decay parameter is scaled (see Section 4.9 in [7]).

In the next section we describe the system model and the first approach to the non-convex stochastic optimization. Section III analyzes the performance of the algorithm using the drift-plus-penalty technique with a penalty that involves a partial derivative (incorporating the "primal-dual" approach). We also compare against the performance of a pure dual-based algorithm in the special case of convex problems. Section IV develops our second approach to the non-convex problem, for the case of optimizing non-concave utility functions that are entrywise non-decreasing.

II. Solving the Problem

A. Model Assumptions

The queueing dynamics for $k \in \{1, \ldots, K\}$ are given by:

$$Q_k(t+1) = \max[Q_k(t) - b_k(t), 0] + a_k(t) \tag{6}$$

where $\boldsymbol{b}(t) = (b_1(t), \ldots, b_K(t))$, $\boldsymbol{a}(t) = (a_1(t), \ldots, a_K(t))$ are non-negative service and arrival vectors, determined as arbitrary functions of $\alpha(t)$ and $\omega(t)$:

$$a_k(t) = \hat{a}_k(\alpha(t), \omega(t)) \ , \ b_k(t) = \hat{b}_k(\alpha(t), \omega(t))$$

The structure of $a_k(t)$ being outside the $\max[\cdot, 0]$ operator in (6) also allows treatment of *multi-hop networks*, where $a_k(t)$ can be a sum of exogenous and endogenous arrivals [3][15].

We assume the random event process $\omega(t)$ is i.i.d. over slots, with some (possibly unknown) distribution. Second moments of the arrival, service, and attribute components are assumed to be finite under all possible control actions. Specifically, for all (possibly randomized) choices of $\alpha \in \mathcal{A}_{\omega(t)}$, made based on knowledge of $\omega(t)$, there is a finite constant $\sigma^2 > 0$ such that:

$$\mathbb{E}\left\{\hat{a}_k(\alpha(t), \omega(t))^2\right\} \leq \sigma^2 \ , \ \mathbb{E}\left\{\hat{b}_k(\alpha(t), \omega(t))^2\right\} \leq \sigma^2$$
$$\mathbb{E}\left\{\hat{x}_m(\alpha(t), \omega(t))^2\right\} \leq \sigma^2 \ , \ \mathbb{E}\left\{\hat{y}_l(\alpha(t), \omega(t))^2\right\} \leq \sigma^2$$

These second moment bounds also imply the first moments of network attributes lie within a bounded region, even if the actual realizations of these attributes are unbounded.

The set $\mathcal{X}$ in (3) is assumed to be a compact and convex subset of $\mathbb{R}^M$. We assume the functions $g_l(\boldsymbol{x})$ are convex (not necessarily differentiable) for all $l \in \{1, \ldots, L\}$, and that they are Lipschitz continuous, in the sense that for any $\boldsymbol{x} = (x_1, \ldots, x_M)$, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_M)$ we have:

$$|g_l(\boldsymbol{x}) - g_l(\boldsymbol{\gamma})| \leq \sum_{m=1}^{M} \nu_{l,m} |x_m - \gamma_m| \tag{7}$$

where $\nu_{l,m}$ are finite non-negative constants. The function $f(\boldsymbol{x})$ is possibly non-convex, but is assumed to be bounded, continuously differentiable, and to have bounded partial derivatives, so that:

$$|f(\boldsymbol{x}) - f(\boldsymbol{\gamma})| \leq \sum_{m=1}^{M} \beta_m |x_m - \gamma_m| \quad (8)$$

for some finite constants $\beta_m \geq 0$. Let f_{min}, f_{max} be finite constants such that for all t we have $f_{min} \leq f(\boldsymbol{x}(t)) \leq f_{max}$.

B. Queue Stability

We first define two types of queue stability. Let $Q(t)$ be a discrete time process defined over slots $t \in \{0, 1, 2, \ldots\}$, assumed to take (possibly negative) real values. Negative queue processes will be useful to meet certain constraints.

Definition 1: $Q(t)$ is *mean rate stable* if:

$$\lim_{t\rightarrow\infty} \frac{\mathbb{E}\{|Q(t)|\}}{t} = 0$$

Definition 2: $Q(t)$ is *strongly stable* if:

$$\limsup_{t\rightarrow\infty} \frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\{|Q(\tau)|\} < \infty$$

Under mild technical assumptions it can be shown that strong stability implies mean rate stability, and that strong stability is equivalent to the existence of a stationary distribution in cases when $Q(t)$ is defined over a countably infinite state Markov chain.

C. Auxiliary Variables

To enable the abstract set constraint (3) to be met and to allow optimization over the convex but possibly non-linear functions $g_l(\boldsymbol{x})$, we introduce a vector of *auxiliary variables* $\boldsymbol{\gamma}(t) = (\gamma_1(t), \ldots, \gamma_M(t))$. This is an additional collection of decision variables that must be chosen every slot, subject only to the constraint:

$$\boldsymbol{\gamma}(t) \in \mathcal{X} \;\; \forall t$$

We transform the problem (1)-(4) into:

$$\text{Minimize:} \quad \overline{y}_0 + f(\overline{\boldsymbol{x}}) \quad (9)$$
$$\text{Subject to:} \quad \overline{y}_l + \overline{g_l(\boldsymbol{\gamma})} \leq 0 \;\; \forall l \in \{1, \ldots, L\} \quad (10)$$
$$\overline{\gamma}_m = \overline{x}_m \;\; \forall m \in \{1, \ldots, M\} \quad (11)$$
$$\text{All queues } Q_k(t) \text{ are mean rate stable} \quad (12)$$
$$\alpha(t) \in \mathcal{A}_{\omega(t)} \;\; \forall t \;\; , \;\; \boldsymbol{\gamma}(t) \in \mathcal{X} \;\; \forall t \quad (13)$$

where notation $\overline{g_l(\boldsymbol{\gamma})}$ is defined:

$$\overline{g_l(\boldsymbol{\gamma})} \triangleq \lim_{t\rightarrow\infty}\frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\{g_l(\boldsymbol{\gamma}(\tau))\}$$

The motivation for this new problem is the observation that, assuming we (temporarily) restrict to algorithms that have well defined limits, the new problem is equivalent to (1)-(5). Indeed, suppose that $\alpha^*(t)$ is a policy that meets all constraints of the problem (1)-(5), and that yields a cost metric in (1) of $\overline{y}_0^* + f(\overline{\boldsymbol{x}}^*)$, where $\overline{y}_l^*$ and $\overline{\boldsymbol{x}}^*$ represent time averages associated with this policy. Then this policy *also* satisfies the constraints of problem (9)-(13), and yields the same cost value in (9), if we define constant auxiliary variable decisions $\boldsymbol{\gamma}^*(t) = \overline{\boldsymbol{x}}^*$ for all t (these satisfy the required constraints (13) because of (3)). It follows that the *infimum* cost metric of the problem (1)-(5) over all policies that meet the constraints is greater than or equal to the infimum associated with problem (9)-(13). Conversely, by Jensen's inequality, it can be shown that any algorithm that solves the constraints in the problem (9)-(13) must *also* satisfy the constraints of the problem (1)-(5), and produces a cost (1) that is the same.

For simplicity of exposition, we have chosen to write the stochastic network optimization problems (1)-(5) and (9)-(13) with notation that implicitly assumes limits are well defined. However, a solution need not have well defined limits, and we can more precisely define the problem using a $\limsup$, as done in future sections (for example, the constraint (2) is written more generally as (16)).

D. Virtual Queues

To ensure the inequality constraints (10) and equality constraints (11) are met, we introduce *virtual queues* $Z_l(t)$ and $H_m(t)$ for $l \in \{1, \ldots, L\}$, $m \in \{1, \ldots, M\}$, with update equation:

$$Z_l(t+1) = \max[Z_l(t) + y_l(t) + g_l(\boldsymbol{\gamma}(t)), 0] \quad (14)$$
$$H_m(t+1) = H_m(t) + \gamma_m(t) - x_m(t) \quad (15)$$

The $Z_l(t)$ queues are always non-negative. The $H_m(t)$ queues have a different structure because they enforce an equality constraint, and can possibly be negative.

Lemma 1: If $\mathbb{E}\{Z_l(0)\} < \infty$ and $\mathbb{E}\{|H_m(0)|\} < \infty$ for all $l \in \{1, \ldots, L\}$ and $m \in \{1, \ldots, M\}$, and if $Z_l(t)$ and $H_m(t)$ are mean rate stable, then:

$$\limsup_{t\rightarrow\infty}[\overline{y}_l(t) + g_l(\overline{\boldsymbol{x}}(t))] \leq 0 \;\; \forall l \in \{1, \ldots, L\} \quad (16)$$
$$\lim_{t\rightarrow\infty}|\overline{\gamma}_m(t) - \overline{x}_m(t)| = 0 \;\; \forall m \in \{1, \ldots, M\} \quad (17)$$
$$\lim_{t\rightarrow\infty}\text{dist}(\overline{\boldsymbol{x}}(t), \mathcal{X}) = 0 \quad (18)$$

where for each $t > 0$, $\overline{\boldsymbol{x}}(t)$ and $\overline{\boldsymbol{\gamma}}(t)$ are defined:

$$\overline{\boldsymbol{x}}(t) \triangleq \frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\{\boldsymbol{x}(\tau)\} \;\; , \;\; \overline{\boldsymbol{\gamma}}(t) \triangleq \frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\{\boldsymbol{\gamma}(\tau)\} \quad (19)$$

where $\overline{y}_l(t)$ is defined similarly, and where $\text{dist}(\overline{\boldsymbol{x}}(t), \mathcal{X})$ is the distance between the vector $\overline{\boldsymbol{x}}(t)$ and the set $\mathcal{X}$, being zero if and only if $\overline{\boldsymbol{x}}(t)$ is in the (closed) set $\mathcal{X}$.

The proximity of the time averages to their required values for all time $t > 0$ is provided as a function of the virtual queue sizes in (20) and (21) of the proof.

Proof: (Lemma 1) Fix $t > 0$. Note by summing (15) over $\tau \in \{0, \ldots, t-1\}$ we have:

$$H_m(t) - H_m(0) = \sum_{\tau=0}^{t-1}\gamma_m(\tau) - \sum_{\tau=0}^{t-1}x_m(\tau)$$

Dividing both sides by t and taking expectations yields:

$$\frac{\mathbb{E}\{H_m(t)\}}{t} - \frac{\mathbb{E}\{H_m(0)\}}{t} = \overline{\gamma}_m(t) - \overline{x}_m(t) \quad (20)$$

The above holds for all $t > 0$. Assuming that $H_m(t)$ is mean rate stable, we can take a limit of (20) as $t \to \infty$ to yield (17).[1] Now note that $\boldsymbol{\gamma}(t) \in \mathcal{X}$ for all t, and hence (because $\mathcal{X}$ is convex), $\overline{\boldsymbol{\gamma}}(t) \in \mathcal{X}$ for all $t > 0$. This together with (17) implies (18).

Finally, from (14) we have for all slots τ:

$$Z_l(\tau+1) \geq Z_l(\tau) + y_l(\tau) + g_l(\boldsymbol{\gamma}(\tau))$$

Fixing $t > 0$ and summing over $\tau \in \{0, \ldots, t-1\}$ yields:

$$Z_l(t) - Z_l(0) \geq \sum_{\tau=0}^{t-1} y_l(\tau) + \sum_{\tau=0}^{t-1} g_l(\boldsymbol{\gamma}(\tau))$$

Taking expectations, dividing by t and using Jensen's inequality yields:

$$\begin{aligned} \frac{\mathbb{E}\{Z_l(t)\}}{t} - \frac{\mathbb{E}\{Z_l(0)\}}{t} &\geq \overline{y}_l(t) + g_l(\overline{\boldsymbol{\gamma}}(t)) \\ &\geq \overline{y}_l(t) + g_l(\overline{\boldsymbol{x}}(t)) \\ &- \sum_{m=1}^{M} \nu_{l,m} |\overline{x}_m(t) - \overline{\gamma}_m(t)| \end{aligned} \quad (21)$$

where the last inequality is due to (7). Assuming $Z_l(t)$ is mean rate stable, we can take the $\limsup$ of both sides of (21) and use (17) to yield (16). □

E. Lyapunov Optimization

Define $\boldsymbol{\Theta}(t) \triangleq [\boldsymbol{Q}(t); \boldsymbol{Z}(t); \boldsymbol{H}(t); \boldsymbol{x}_{av}(t)]$ as the collective queue vector together with the current running time-average of attribute vector $\boldsymbol{x}(\tau)$:

$$\boldsymbol{x}_{av}(t) \triangleq \frac{1}{t} \sum_{\tau=0}^{t-1} \boldsymbol{x}(\tau) \quad (22)$$

Define the following non-negative function, called a Lyapunov function:

$$L(\boldsymbol{\Theta}(t)) \triangleq \frac{1}{2} \sum_{k=1}^{K} Q_k(t)^2 + \frac{1}{2} \sum_{l=1}^{L} Z_l(t)^2 + \frac{1}{2} \sum_{m=1}^{M} H_m(t)^2 \quad (23)$$

Define the *conditional Lyapunov drift* $\Delta(\boldsymbol{\Theta}(t))$ by:[2]

$$\Delta(\boldsymbol{\Theta}(t)) \triangleq \mathbb{E}\{L(\boldsymbol{\Theta}(t+1)) - L(\boldsymbol{\Theta}(t)) | \boldsymbol{\Theta}(t)\} \quad (24)$$

Lemma 2: Under any algorithm for choosing $\alpha(t) \in \mathcal{A}_{\omega(t)}$ and $\boldsymbol{\gamma}(t) \in \mathcal{X}$ for all t, we have for all t and all possible $\boldsymbol{\Theta}(t)$:

$$\Delta(\boldsymbol{\Theta}(t)) \leq RHS(\alpha(t), \boldsymbol{\gamma}(t), \boldsymbol{\Theta}(t)) \quad (25)$$

where

$$\begin{aligned} &RHS(\alpha(t), \boldsymbol{\gamma}(t), \boldsymbol{\Theta}(t)) \triangleq \\ &B + \sum_{k=1}^{K} Q_k(t) \mathbb{E}\left\{\hat{a}_k(\alpha(t), \omega(t)) - \hat{b}_k(\alpha(t), \omega(t)) | \boldsymbol{\Theta}(t)\right\} \\ &+ \sum_{l=1}^{L} Z_l(t) \mathbb{E}\{\hat{y}_l(\alpha(t), \omega(t)) + g_l(\boldsymbol{\gamma}(t)) | \boldsymbol{\Theta}(t)\} \\ &+ \sum_{m=1}^{M} H_m(t) \mathbb{E}\{\gamma_m(t) - \hat{x}_m(\alpha(t), \omega(t)) | \boldsymbol{\Theta}(t)\} \end{aligned}$$

where B is a finite constant that satisfies the following for all t and all possible $\boldsymbol{\Theta}(t)$:

$$\begin{aligned} B \geq & \sum_{k=1}^{K} \mathbb{E}\left\{b_k(t)^2 + a_k(t)^2 | \boldsymbol{\Theta}(t)\right\} \\ &+ \sum_{l=1}^{L} \mathbb{E}\left\{(y_l(t) + g_l(\boldsymbol{\gamma}(t)))^2 | \boldsymbol{\Theta}(t)\right\} \\ &+ \sum_{m=1}^{M} \mathbb{E}\left\{(\gamma_m(t) - x_m(t))^2 | \boldsymbol{\Theta}(t)\right\} \end{aligned} \quad (26)$$

Such a finite constant B exists by the system boundedness assumptions.

Proof: The proof follows by squaring the equations (6), (14), (15) and is omitted for brevity (see [3]). □

Our algorithm seeks to take control actions $\alpha(t) \in \mathcal{A}_{\omega(t)}$, $\boldsymbol{\gamma}(t) \in \mathcal{X}$ every slot t in reaction to the observed $\omega(t)$ and to the current queue states and $\boldsymbol{x}_{av}(t)$ values in $\boldsymbol{\Theta}(t)$, to minimize:

$$\begin{aligned} &RHS(\alpha(t), \boldsymbol{\gamma}(t), \boldsymbol{\Theta}(t)) + V\mathbb{E}\{\hat{y}_0(\alpha(t), \omega(t)) | \boldsymbol{\Theta}(t)\} \\ &\quad + \sum_{m=1}^{M} \frac{V \partial f(\boldsymbol{x}_{av}(t))}{\partial x_m} \mathbb{E}\{\hat{x}_m(\alpha(t), \omega(t)) | \boldsymbol{\Theta}(t)\} \end{aligned} \quad (27)$$

where $\boldsymbol{x}_{av}(t)$ is defined in (22), and where we define $\boldsymbol{x}_{av}(0)$ as an initial sample $\hat{\boldsymbol{x}}(\alpha(-1), \omega(-1))$ under any decision $\alpha(-1) \in \mathcal{A}_{\omega(-1)}$ taken at some preliminary slot. The parameter V is chosen as a non-negative constant that will affect proximity to the optimal solution, with a corresponding tradeoff in average queue congestion, as in [3].

In the special case when there is no non-convex function, so that $f(\boldsymbol{x}) = 0$, we do not require the time averages $\boldsymbol{x}_{av}(t)$. In this case the above algorithm reduces to one that is similar to our work in [3], with the exception that we are considering an abstract set constraint $\mathcal{X}$ and the convex functions $g_l(\boldsymbol{x})$ are not necessarily monotonic. In the case when $f(\boldsymbol{x}) \neq 0$, our algorithm includes a derivative multiplied by a running time average, as in the primal-dual algorithms that do not consider queue stability in [5][6]. This primal-dual component is also similar to fluid limit works in [7][16], which do consider queue stability, with the exception that our algorithm includes auxiliary variables and does not use an exponential weighted average.[3] An advantage of our approach is that it is compatible with non-convex objectives, ensures all convex constraints are satisfied as $t \to \infty$, and provides explicit bounds on constraint violation for all $t > 0$.

The algorithm is specified as follows: Every slot t, observe $\omega(t)$, $\boldsymbol{\Theta}(t)$ (where $\boldsymbol{\Theta}(t)$ includes all current queue states and the current $\boldsymbol{x}_{av}(t)$), and perform the following:

- (Auxiliary Variables) Choose $\boldsymbol{\gamma}(t) \in \mathcal{X}$ to minimize:

$$\sum_{l=1}^{L} Z_l(t) g_l(\boldsymbol{\gamma}(t)) + \sum_{m=1}^{M} H_m(t) \gamma_m(t)$$

[1] Note that if $\mathbb{E}\{|H_m(t)|\}/t \to 0$ then $\mathbb{E}\{H_m(t)\}/t \to 0$.

[2] Strictly speaking, correct notation should be $\Delta(\boldsymbol{\Theta}(t), t)$, as the drift may be non-stationary, although we use the simpler notation $\Delta(\boldsymbol{\Theta}(t))$ as a formal representation of the right hand side of (24).

[3] Using an exponential weighted average would be advantagous for robustness to system changes, but it is not clear if the algorithm would converge properly in this case.

- (Choosing $\alpha(t)$) Choose $\alpha(t) \in \mathcal{A}_{\omega(t)}$ to minimize:

$$\sum_{k=1}^{K} Q_k(t)[\hat{a}_k(\alpha(t),\omega(t)) - \hat{b}_k(\alpha(t),\omega(t))]$$
$$+ \sum_{l=1}^{L} Z_l(t)\hat{y}_l(\alpha(t),\omega(t)) - \sum_{m=1}^{M} H_m(t)\hat{x}_m(\alpha(t),\omega(t))$$
$$+V\hat{y}_0(\alpha(t),\omega(t)) + \sum_{m=1}^{M} \frac{V\partial f(\boldsymbol{x}_{av}(t))}{\partial x_m}\hat{x}_m(\alpha(t),\omega(t))$$

- (Queue Updates) Update the virtual queues $Z_l(t)$ and $H_m(t)$ according to (14), (15), and update the actual queues $Q_k(t)$ via (6). Also update $\boldsymbol{x}_{av}(t)$ by (22).

III. PERFORMANCE ANALYSIS

A. Optimality via ω-Only Policies

We say that the problem (1)-(5) is *feasible* if there is an algorithm for choosing $\alpha(t) \in \mathcal{A}_{\omega(t)}$ over time so that all constraints (2)-(5) are satisfied. We assume throughout that the problem is feasible. Define $y_0^{opt} + f^{opt}$ as the infimum value of (1) over all feasible policies.

It turns out that this infimum can be characterized by *ω-only policies*. Specifically, we say that a policy $\alpha^*(t)$ is *ω-only* if for all t it chooses $\alpha^*(t) \in \mathcal{A}_{\omega(t)}$ as a stationary and randomized function of the observed $\omega(t)$, i.i.d. over all slots for which the same $\omega(t)$ value is observed. By the law of large numbers, all time averages resulting from an ω-only policy are well defined and are equal to their expectations over one slot. In particular, we have for all t:

$$\overline{y}_l = \mathbb{E}\left\{\hat{y}_l(\alpha^*(t),\omega(t))\right\} \ , \ \overline{\boldsymbol{x}} = \mathbb{E}\left\{\hat{\boldsymbol{x}}(\alpha^*(t),\omega(t))\right\}$$

For a given $\epsilon > 0$, define the set $\mathcal{Y}_\epsilon$ as the set of all vectors $(\boldsymbol{y},\boldsymbol{x}) = (\overline{y}_0, \overline{y}_1, \ldots, \overline{y}_L, \overline{x}_1, \ldots, \overline{x}_M)$ that can be achieved as time averages by some ω-only policy $\alpha^*(t) \in \mathcal{A}_{\omega(t)}$ that satisfies:

$$\overline{y}_l + g_l(\overline{\boldsymbol{x}}) \leq \epsilon \ \forall l \in \{1,\ldots,L\}$$
$$\text{dist}(\overline{\boldsymbol{x}},\mathcal{X}) \leq \epsilon$$
$$\mathbb{E}\left\{\hat{a}_k(\alpha^*(t),\omega(t)) - \hat{b}_k(\alpha^*(t),\omega(t))\right\} \leq \epsilon$$

The set $\mathcal{Y}_\epsilon$ is bounded by the boundedness assumptions, and can be shown to be convex and to satisfy $\mathcal{Y}_{\epsilon_1} \subseteq \mathcal{Y}_{\epsilon_2}$ whenever $\epsilon_1 \leq \epsilon_2$. Define $Cl(\mathcal{Y}_\epsilon)$ as the closure of $\mathcal{Y}_\epsilon$, and define $\mathcal{Y}$ as the limit of $Cl(\mathcal{Y}_\epsilon)$ as $\epsilon \rightarrow 0$:

$$\mathcal{Y} \triangleq \cap_{i=1}^{\infty} Cl(\mathcal{Y}_{1/i})$$

It can be shown that $\mathcal{Y}$ is non-empty whenever the problem (1)-(5) is feasible, and is bounded, closed, and convex.

Theorem 1: (Optimality over the set $\mathcal{Y}$) The set $\mathcal{Y}$ represents the set of all vectors $(\boldsymbol{y},\boldsymbol{x})$ that can be achieved as a limit point of a trajectory $(\overline{\boldsymbol{y}}(t), \overline{\boldsymbol{x}}(t))$ generated by a feasible policy implemented over $t \in \{0,1,2,\ldots\}$. Further, the infimum value $y_0^{opt} + f^{opt}$ is the solution to the following problem:

$$\text{Minimize:} \quad y_0 + f(\boldsymbol{x})$$
$$\text{Subject to:} \quad (y_0, y_1, \ldots, y_L, x_1, \ldots, x_M) \in \mathcal{Y}$$

Proof: Omitted for brevity (see, for example, [9][10]). □

Theorem 2: For any vector $(\boldsymbol{y}^*,\boldsymbol{x}^*) \in \mathcal{Y}$ and for any $\epsilon > 0$, there is an ω-only policy $\alpha^*(t)$ such that:

$$|\mathbb{E}\left\{\hat{y}_l(\alpha^*(t),\omega(t))\right\} - y_l^*| \leq \epsilon \ \forall l \in \{0,1,\ldots,L\} \quad (28)$$
$$|\mathbb{E}\left\{\hat{x}_m(\alpha^*(t),\omega(t))\right\} - x_m^*| \leq \epsilon \ \forall m \in \{1,\ldots,M\} \quad (29)$$
$$\mathbb{E}\left\{\hat{y}_l(\alpha^*(t),\omega(t))\right\} + g_l(\mathbb{E}\left\{\hat{\boldsymbol{x}}(\alpha^*(t),\omega(t))\right\}) \leq \epsilon \quad (30)$$
$$\text{dist}\left(\mathbb{E}\left\{\hat{\boldsymbol{x}}(\alpha^*(t),\omega(t))\right\},\mathcal{X}\right) \leq \epsilon \quad (31)$$
$$\mathbb{E}\left\{\hat{a}_k(\alpha^*(t),\omega(t))\right\} \leq \mathbb{E}\left\{\hat{b}_k(\alpha^*(t),\omega(t))\right\} + \epsilon \ \forall k \quad (32)$$

Proof: The proof follows by observing that any point in $\mathcal{Y}$ can be achieved arbitrarily closely by an ω-only policy that satisfies the required constraints to within ϵ. □

B. The Performance Theorem

For a given constant $C \geq 0$, we say that an algorithm is a *C-approximation* if every slot t, given the existing $\boldsymbol{\Theta}(t)$ on slot t, it chooses $\alpha(t) \in \mathcal{A}_{\omega(t)}$, $\boldsymbol{\gamma}(t) \in \mathcal{X}$ to achieve a value in (27) that is within C of the infimum over all possible decisions given $\boldsymbol{\Theta}(t)$. A 0-approximation achieves the exact infimum and is given by the algorithm in Section II. Using $C > 0$ allows for approximate implementations.

Theorem 3: Suppose $V \geq 0$, $\omega(t)$ is i.i.d. over slots, the problem (1)-(5) is feasible, and any C-approximate algorithm is used every slot t. For simplicity, assume that all queues are initially zero, so that $\boldsymbol{\Theta}(0) = \boldsymbol{0}$. Then:

(a) All queues are mean rate stable, and hence all required constraints are satisfied. In particular, the equalities and inequalities (16)-(18) are ensured. Bounds on expected queue sizes (and hence, constraint violations) for all $t > 0$ are provided in inequality (35) of the proof.

(b) For all $t > 0$ and for any $(\boldsymbol{y}^*,\boldsymbol{x}^*) \in \mathcal{Y}$ (including the one that optimizes the infimum cost), we have:

$$\overline{y}_0(t) + \frac{1}{t}\sum_{\tau=0}^{t-1}\sum_{m=1}^{M}\mathbb{E}\left\{\frac{x_m(\tau)\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\} \leq$$
$$y_0^* + \frac{1}{t}\sum_{\tau=0}^{t-1}\sum_{m=1}^{M}\mathbb{E}\left\{\frac{x_m^*\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\} + \frac{B+C}{V}$$

where B is defined by (26).

(c) If all time averages converge, so that there are constant vectors $\overline{\boldsymbol{x}}$, $\overline{\boldsymbol{y}}$ such that $\boldsymbol{x}_{av}(t) \rightarrow \overline{\boldsymbol{x}}$ with probability 1, $\overline{\boldsymbol{x}}(t) \rightarrow \overline{\boldsymbol{x}}$, and $\overline{y}_m(t) \rightarrow \overline{y}_m$, then the achieved limiting point is a *near local optimum*, in the sense that for any $(\boldsymbol{y}^*,\boldsymbol{x}^*) \in \mathcal{Y}$ (including any local or global optimum):

$$\overline{y}_0 + \sum_{m=1}^{M}\frac{\overline{x}_m\partial f(\overline{\boldsymbol{x}})}{\partial x_m} \leq y_0^* + \sum_{m=1}^{M}\frac{x_m^*\partial f(\overline{\boldsymbol{x}})}{\partial x_m} + \frac{B+C}{V}$$

The result of part (c) can be viewed as a near-local optimum result because of the "fudge-factor" $(B+C)/V$, which can be made arbitrarily small as the V parameter is increased. Indeed, the above result shows that the achieved cost, as measured by a *linearized* version of the $f(\boldsymbol{x})$ function about our achieved $\overline{\boldsymbol{x}}$ vector, is no more than $(B+C)/V$ above the cost associated with any other $(\boldsymbol{y}^*,\boldsymbol{x}^*) \in \mathcal{Y}$, when measured against the same linearized function. In particular, suppose we consider moving from our achieved $(\overline{\boldsymbol{y}},\overline{\boldsymbol{x}})$ vector towards any other $(\boldsymbol{y}^*,\boldsymbol{x}^*) \in$

$\mathcal{Y}$ by taking a small step $\epsilon(\boldsymbol{y}^* - \overline{\boldsymbol{y}}, \boldsymbol{x}^* - \overline{\boldsymbol{x}})$ in this direction (for some small value $\epsilon > 0$). If ϵ is small, then the cost differential $\Delta_{cost}(\epsilon)$ incurred satisfies:

$$\frac{\Delta_{cost}(\epsilon)}{\epsilon} \approx (y_0^* - \overline{y}_0) + \sum_{m=1}^{M} \frac{(x_m^* - \overline{x}_m)\partial f(\overline{\boldsymbol{x}})}{\partial x_m} \geq -\frac{(B+C)}{V}$$

where the inequality follows by part (c) of the above theorem. It follows that such a step cannot decrease cost by more than approximately $\epsilon(B+C)/V$ (where the approximation is due to the linearization), which can be made arbitrarily small with a suitably large choice of V. Hence, this is a *near local optimum*.

The value of V affects the coefficient in the decay of $\mathbb{E}\{|H_m(t)|\}/t$, $\mathbb{E}\{Z_l(t)\}/t$, $\mathbb{E}\{Q_k(t)\}/t$ derived in (35) of the proof, which also affects the decay in constraint violation via (20), (21). It can be shown that if a mild *Slater-type* condition holds for the constraints (2)-(5) of the problem, then all queues are *strongly stable* and have average backlog that is $O(V)$ (see [3][10][9] for similar results).

Proof: (Theorem 3 part (a)) By simply adding the same thing to both sides of (25) in Lemma 2, we have:

$$\begin{aligned}
&\Delta(\boldsymbol{\Theta}(t)) + V\mathbb{E}\{y_0(t)|\boldsymbol{\Theta}(t)\} \\
&+ \sum_{m=1}^{M} \frac{V\partial f(\boldsymbol{x}_{av}(t))}{\partial x_m}\mathbb{E}\{x_m(t)|\boldsymbol{\Theta}(t)\} \leq B \\
&\quad + \sum_{k=1}^{K} Q_k(t)\mathbb{E}\{a_k(t) - b_k(t)|\boldsymbol{\Theta}(t)\} \\
&\quad + \sum_{l=1}^{L} Z_l(t)\mathbb{E}\{y_l(t) + g_l(\boldsymbol{\gamma}(t))|\boldsymbol{\Theta}(t)\} \\
&\quad + \sum_{m=1}^{M} H_m(t)\mathbb{E}\{\gamma_m(t) - x_m(t)|\boldsymbol{\Theta}(t)\} \\
&\quad + V\mathbb{E}\{y_0(t)|\boldsymbol{\Theta}(t)\} + \sum_{m=1}^{M} \frac{V\partial f(\boldsymbol{x}_{av}(t))}{\partial x_m}\mathbb{E}\{x_m(t)|\boldsymbol{\Theta}(t)\}
\end{aligned}$$

Because, given the existing $\boldsymbol{\Theta}(t)$ on slot t, our C-approximation minimizes the right-hand-side of the above expression to within C of the infimum over all policies for choosing $\alpha(t) \in \mathcal{A}_{\omega(t)}$, $\boldsymbol{\gamma}(t) \in \mathcal{X}$, we have:

$$\begin{aligned}
&\Delta(\boldsymbol{\Theta}(t)) + V\mathbb{E}\{y_0(t)|\boldsymbol{\Theta}(t)\} + \\
&\sum_{m=1}^{M} \frac{V\partial f(\boldsymbol{x}_{av}(t))}{\partial x_m}\mathbb{E}\{x_m(t)|\boldsymbol{\Theta}(t)\} \leq B + C \\
&\quad + \sum_{k=1}^{K} Q_k(t)\mathbb{E}\{a_k^*(t) - b_k^*(t)|\boldsymbol{\Theta}(t)\} \\
&\quad + \sum_{l=1}^{L} Z_l(t)\mathbb{E}\{y_l^*(t) + g_l(\boldsymbol{\gamma}^*(t))|\boldsymbol{\Theta}(t)\} \\
&\quad + \sum_{m=1}^{M} H_m(t)\mathbb{E}\{\gamma_m^*(t) - x_m^*(t)|\boldsymbol{\Theta}(t)\} \\
&\quad + V\mathbb{E}\{y_0^*(t)|\boldsymbol{\Theta}(t)\} + \sum_{m=1}^{M} \frac{V\partial f(\boldsymbol{x}_{av}(t))}{\partial x_m}\mathbb{E}\{x_m^*(t)|\boldsymbol{\Theta}(t)\}
\end{aligned}$$

where $\alpha^*(t) \in \mathcal{A}_{\omega(t)}$, $\boldsymbol{\gamma}^*(t) \in \mathcal{X}$ are any alternative control actions, and for all $k \in \{1, \ldots, K\}$, $l \in \{0, 1, \ldots, L\}$:

$$a_k^*(t) \triangleq \hat{a}_k(\alpha^*(t), \omega(t)) \;,\; b_k^*(t) \triangleq \hat{b}_k(\alpha^*(t), \omega(t))$$
$$y_l^*(t) \triangleq \hat{y}_l(\alpha^*(t), \omega(t))$$

For any $(\boldsymbol{y}^*, \boldsymbol{x}^*) \in \mathcal{Y}$ and any $\epsilon > 0$, Theorem 2 ensures there is an ω-only algorithm $\alpha^*(t)$ that satisfies (28)-(32). For simplicity, assume that (28)-(32) hold for $\epsilon = 0$.[4] Now plug this algorithm into the right hand side of the above drift expression, and note that, because it is ω-only, it makes decisions independent of backlog (so that the conditional expectation given $\boldsymbol{\Theta}(t)$ is the same as the unconditional expectation). Further, plug decisions $\boldsymbol{\gamma}^*(t) = \mathbb{E}\{\hat{\boldsymbol{x}}(\alpha^*(t), \omega(t))\} = \boldsymbol{x}^*$ (this satisfies $\boldsymbol{\gamma}(t) \in \mathcal{X}$ as required because of (31) with $\epsilon = 0$ and noting that $\mathcal{X}$ is closed). This simplifies the right-hand-side of the above drift bound to:

$$\begin{aligned}
&\Delta(\boldsymbol{\Theta}(t)) + V\mathbb{E}\{y_0(t)|\boldsymbol{\Theta}(t)\} + \\
&\sum_{m=1}^{M} \frac{V\partial f(\boldsymbol{x}_{av}(t))}{\partial x_m}\mathbb{E}\{x_m(t)|\boldsymbol{\Theta}(t)\} \leq B + C + Vy_0^* \\
&\quad + \sum_{m=1}^{M} \frac{Vx_m^*\partial f(\boldsymbol{x}_{av}(t))}{\partial x_m}
\end{aligned} \tag{33}$$

Because of the boundedness assumptions on the set $\mathcal{X}$, the first and second moments of the attribute vectors, and the function $f(\boldsymbol{x})$ and its derivatives, the above drift bound simplifies to:

$$\Delta(\boldsymbol{\Theta}(t)) \leq D \tag{34}$$

where D is any finite constant that satisfies for all t, $\boldsymbol{\Theta}(t)$:

$$\begin{aligned}
D \geq\ & B + C + V\left[y_0^* - \mathbb{E}\{y_0(t)|\boldsymbol{\Theta}(t)\}\right] \\
&+ \sum_{m=1}^{M} V\beta_m\left[x_m^* - \mathbb{E}\{x_m(t)|\boldsymbol{\Theta}(t)\}\right]
\end{aligned}$$

Taking an expectation of both sides and using the law of iterated expectations yields:

$$\mathbb{E}\{L(\boldsymbol{\Theta}(t+1))\} - \mathbb{E}\{L(\boldsymbol{\Theta}(t))\} \leq D$$

The above holds for all t. Summing over $\tau \in \{0, \ldots, t-1\}$ and using the fact that $L(\boldsymbol{\Theta}(0)) = 0$ yields:

$$\mathbb{E}\{L(\boldsymbol{\Theta}(t))\} \leq Dt$$

By definition of $L(\boldsymbol{\Theta}(t))$ in (23) we have:

$$\mathbb{E}\left\{Q_k(t)^2\right\}, \mathbb{E}\left\{Z_l(t)^2\right\}, \mathbb{E}\left\{H_m(t)^2\right\} \leq 2Dt$$

Because $\mathbb{E}\{|X|\}^2 \leq \mathbb{E}\left\{X^2\right\}$ for any random variable X, we have:

$$\mathbb{E}\{Q_k(t)\}^2, \mathbb{E}\{Z_l(t)\}^2, \mathbb{E}\{|H_m(t)|\}^2 \leq 2Dt$$

Dividing by t^2 and taking square roots shows that for all $t > 0$:

$$\frac{\mathbb{E}\{Q_k(t)\}}{t}, \frac{\mathbb{E}\{Z_l(t)\}}{t}, \frac{\mathbb{E}\{|H_m(t)|\}}{t} \leq \sqrt{\frac{2D}{t}} \tag{35}$$

Taking a limit as $t \to \infty$ proves mean rate stability of all queues. □

[4]The same result (33) can be derived without this assumption by taking a limit as $\epsilon \to 0$.

Proof: (Theorem 3 parts (b) and (c)) We start with the drift bound in (33) derived in the proof of part (a), which holds for all $t \geq 0$. Consider this inequality for some slot τ. Taking expectations of both sides and using the law of iterated expectations and the definition of $\Delta(\boldsymbol{\Theta}(\tau))$ in (24), we have:

$$\begin{aligned}\mathbb{E}\left\{L(\boldsymbol{\Theta}(\tau+1))\right\} - \mathbb{E}\left\{L(\boldsymbol{\Theta}(\tau))\right\} + V\mathbb{E}\left\{y_0(\tau)\right\} \\ + \sum_{m=1}^{M}\mathbb{E}\left\{\frac{Vx_m(\tau)\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\} \\ \leq B + C + Vy_0^* + \sum_{m=1}^{M}\mathbb{E}\left\{\frac{Vx_m^*\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\} \end{aligned} \quad (36)$$

Fix any slot $t > 0$. Summing the above over $\tau \in \{0, \ldots, t-1\}$ and using telescoping sums yields:

$$\begin{aligned}\mathbb{E}\left\{L(\boldsymbol{\Theta}(t))\right\} - \mathbb{E}\left\{L(\boldsymbol{\Theta}(0))\right\} + V\sum_{\tau=0}^{t-1}\mathbb{E}\left\{y_0(\tau)\right\} \\ + \sum_{\tau=0}^{t-1}\sum_{m=1}^{M}\mathbb{E}\left\{\frac{Vx_m(\tau)\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\} \\ \leq t(B + C + Vy_0^*) + \\ \sum_{\tau=0}^{t-1}\sum_{m=1}^{M}\mathbb{E}\left\{\frac{Vx_m^*\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\}\end{aligned}$$

Dividing by Vt and using the fact that $L(\boldsymbol{\Theta}(0)) = 0$ and $L(\boldsymbol{\Theta}(t)) \geq 0$ yields:

$$\begin{aligned}\overline{y}_0(t) + \frac{1}{t}\sum_{\tau=0}^{t-1}\sum_{m=1}^{M}\mathbb{E}\left\{\frac{x_m(\tau)\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\} \leq \\ \frac{B+C}{V} + y_0^* + \frac{1}{t}\sum_{\tau=0}^{t-1}\sum_{m=1}^{M}\mathbb{E}\left\{\frac{x_m^*\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\}\end{aligned}$$

This proves part (b). Part (c) is an immediate consequence, noting boundedness and continuity of derivatives of $f(\boldsymbol{x})$. □

C. Variable V for Local Optimality

As in [17], we can use an algorithm that gradually increases the value of V while maintaining mean rate stability. This eliminates the "fudge factor" $(B+C)/V$ in Theorem 3 part (c). A disadvantage is that it precludes strong stability of queues even when a Slater condition is satisfied.

Theorem 4: Suppose that we use $V(t)$ instead of V in the drift expression (27), and we use any C-approximation, i.e., any decisions that come within C of minimizing (27) under this variable $V(t)$ setting on all slots t. Suppose $\omega(t)$ is i.i.d. over slots, the problem (1)-(5) is feasible, and all queues are initially empty. Further assume that we use the following increasing $V(t)$ function:

$$V(t) \triangleq V_0(1+t)^d$$

where $V_0 > 0$ and d satisfies $0 < d < 1$. Then:

(a) All queues are mean rate stable as before, and hence the equalities and inequalities in (16)-(18) are satisfied.

(b) If all time averages converge, so that there are vectors $\overline{\boldsymbol{x}}$, $\overline{\boldsymbol{y}}$ such that $\boldsymbol{x}_{av}(t) \to \overline{\boldsymbol{x}}$ with probability 1, $\overline{\boldsymbol{x}}(t) \to \overline{\boldsymbol{x}}$, and $\overline{y}_m(t) \to \overline{y}_m$, then the achieved point is a local optimum, in the sense that for any $(\boldsymbol{y}^*, \boldsymbol{x}^*) \in \mathcal{Y}$ (including any local or global optimum):

$$\overline{y}_0 + \sum_{m=1}^{M}\frac{\overline{x}_m\partial f(\overline{\boldsymbol{x}})}{\partial x_m} \leq y_0^* + \sum_{m=1}^{M}\frac{x_m^*\partial f(\overline{\boldsymbol{x}})}{\partial x_m}$$

That is, the achieved vector $(\boldsymbol{y}^*, \boldsymbol{x}^*)$ *globally optimizes the linearized cost function*. In particular, if we start at our achieved $(\overline{\boldsymbol{y}}, \overline{\boldsymbol{x}})$ and move in the direction of any other $(\boldsymbol{y}^*, \boldsymbol{x}^*) \in \mathcal{Y}$ by taking a small step $\epsilon(\boldsymbol{y}^* - \overline{\boldsymbol{y}}, \boldsymbol{x}^* - \overline{\boldsymbol{x}})$, then:

$$\lim_{\epsilon \to 0^+}\frac{\Delta_{cost}(\epsilon)}{\epsilon} = (y_0^* - \overline{y}_0) + \sum_{m=1}^{M}\frac{(x_m^* - \overline{x}_m)\partial f(\overline{\boldsymbol{x}})}{\partial x_m} \geq 0$$

Thus, such a small step cannot improve cost.

While the theorem holds for any $V_0 > 0$ and $0 < d < 1$, these values affect the manner in which the utility and virtual queue values converge, as shown in the proof.

Proof: (Theorem 4) The proof of part (a) is similar to [17] and is omitted for brevity. Here we prove part (b). Following the proof of Theorem 3 almost exactly, we find that (36) translates to the following for any slot $\tau \geq 0$:

$$\begin{aligned}\mathbb{E}\left\{L(\boldsymbol{\Theta}(\tau+1))\right\} - \mathbb{E}\left\{L(\boldsymbol{\Theta}(\tau))\right\} + V(\tau)\mathbb{E}\left\{y_0(\tau)\right\} \\ + \sum_{m=1}^{M}\mathbb{E}\left\{\frac{V(\tau)x_m(\tau)\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\} \\ \leq B + C + V(\tau)y_0^* + \sum_{m=1}^{M}\mathbb{E}\left\{\frac{V(\tau)x_m^*\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\}\end{aligned}$$

Dividing everything by $V(\tau)$ yields:

$$\begin{aligned}\frac{\mathbb{E}\left\{L(\boldsymbol{\Theta}(\tau+1))\right\}}{V(\tau)} - \frac{\mathbb{E}\left\{L(\boldsymbol{\Theta}(\tau))\right\}}{V(\tau)} + \mathbb{E}\left\{y_0(\tau)\right\} \\ + \sum_{m=1}^{M}\mathbb{E}\left\{\frac{x_m(\tau)\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\} \leq \\ \frac{B+C}{V(\tau)} + y_0^* + \sum_{m=1}^{M}\mathbb{E}\left\{\frac{x_m^*\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\}\end{aligned}$$

Fix $t > 0$, and sum the above over $\tau \in \{0, \ldots, t-1\}$. Collecting terms yields:

$$\begin{aligned}\frac{\mathbb{E}\left\{L(\boldsymbol{\Theta}(t))\right\}}{V(t-1)} + \sum_{\tau=1}^{t-1}\mathbb{E}\left\{L(\boldsymbol{\Theta}(\tau))\right\}\left[\frac{1}{V(\tau-1)} - \frac{1}{V(\tau)}\right] \\ + \sum_{\tau=0}^{t-1}\mathbb{E}\left\{y_0(\tau)\right\} + \sum_{\tau=0}^{t-1}\sum_{m=1}^{M}\mathbb{E}\left\{\frac{x_m(\tau)\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\} \leq \\ \sum_{\tau=0}^{t-1}\frac{B+C}{V(\tau)} + ty_0^* + \sum_{\tau=0}^{t-1}\sum_{m=1}^{M}\mathbb{E}\left\{\frac{x_m^*\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\}\end{aligned}$$

Because $V(\tau)$ is non-decreasing, we have for all τ:

$$\frac{1}{V(\tau-1)} - \frac{1}{V(\tau)} \geq 0$$

Using this, dividing everything by t, and using non-negativity of $L(\cdot)$ yields for all $t > 0$:

$$\overline{y}_0(t) + \frac{1}{t}\sum_{\tau=0}^{t-1}\sum_{m=1}^{M}\mathbb{E}\left\{\frac{x_m(\tau)\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\} \leq$$

$$\frac{1}{t}\sum_{\tau=0}^{t-1}\frac{(B+C)}{V(\tau)} + y_0^* + \frac{1}{t}\sum_{\tau=0}^{t-1}\sum_{m=1}^{M}\mathbb{E}\left\{\frac{x_m^*\partial f(\boldsymbol{x}_{av}(\tau))}{\partial x_m}\right\}$$

However, we have:

$$\frac{1}{t}\sum_{\tau=0}^{t-1}\frac{(B+C)}{V(\tau)} = \frac{(B+C)}{V_0}\frac{1}{t}\sum_{\tau=0}^{t-1}(t+1)^{-d} = O(t^{-d}) \rightarrow 0$$

Thus, taking limits of the above drift bound and using the assumption that time averages converge yields the result. □

D. Convex Problems and the Pure Dual Approach

A significant limitation of the primal-dual component in the above algorithms is that the running time average $\boldsymbol{x}_{av}(t)$ in (22) precludes adaptation when system conditions change. An additional limitation is that our theorem requires vectors $\boldsymbol{x}_{av}(t)$, $\overline{\boldsymbol{x}}(t)$, $\overline{\boldsymbol{y}}(t)$ to converge, but lacks a proof of convergence and a measure of how long such convergence would take.[5] For comparison, here we show that these limitations can be solved by the drift-plus-penalty algorithm without using partial derivatives, via the pure "dual-based" approach, when $f(\boldsymbol{x})$ is convex. These advantages are exploited in an alternative algorithm for the non-convex problem in Section IV. The analysis in this subsection is similar to [3][17], with the exception that we include the abstract set constraint $\mathcal{X}$.

Assume here that the function $f(\boldsymbol{x})$ is continuous and convex (possibly non-differentiable). Define finite bounds f_{min} and f_{max} to satisfy for all t and all (possibly randomized) decisions $\alpha(t) \in \mathcal{A}_{\omega(t)}$:

$$f_{min} \leq f(\mathbb{E}\left\{\hat{\boldsymbol{x}}(\alpha(t), \omega(t))\right\}) \leq f_{max}$$

The following algorithm is motivated by seeking to solve a modified version of the problem (9)-(13), with the cost metric (9) changed to minimizing $\overline{y}_0 + \overline{f(\boldsymbol{\gamma})}$, which can be shown to still be equivalent to the original problem (1)-(5) by Jensen's inequality and the fact that $f(\boldsymbol{\gamma})$ is convex.

Using the same virtual queues and the same Lyapunov function $L(\boldsymbol{\Theta}(t))$ as before (with the exception that $\boldsymbol{\Theta}(t)$ no longer includes $\boldsymbol{x}_{av}(t)$ information), we have by adding the same thing to both sides of the drift inequality of Lemma 2:

$$\begin{aligned}\Delta(\boldsymbol{\Theta}(t)) + V\mathbb{E}\left\{y_0(t) + f(\boldsymbol{\gamma}(t))|\boldsymbol{\Theta}(t)\right\} \leq B \\ + \sum_{k=1}^{K} Q_k(t)\mathbb{E}\left\{a_k(t) - b_k(t)|\boldsymbol{\Theta}(t)\right\} \\ + \sum_{l=1}^{L} Z_l(t)\mathbb{E}\left\{y_l(t) + g_l(\boldsymbol{\gamma}(t))|\boldsymbol{\Theta}(t)\right\} \\ + \sum_{m=1}^{M} H_m(t)\mathbb{E}\left\{\gamma_m(t) - x_m(t)|\boldsymbol{\Theta}(t)\right\} \\ + V\mathbb{E}\left\{y_0(t) + f(\boldsymbol{\gamma}(t))|\boldsymbol{\Theta}(t)\right\} \end{aligned} \quad (37)$$

[5]Under an additional Slater condition, the policy is a Markov chain with all queues strongly stable, and so convergence would follow if we had a countably infinite state space.

Our policy is now to observe $\boldsymbol{\Theta}(t)$ and $\omega(t)$ for each slot t, and to take actions $\alpha(t) \in \mathcal{A}_{\omega(t)}$, $\boldsymbol{\gamma}(t) \in \mathcal{X}$ to minimize the right hand side of the above drift bound. For a given constant $C \geq 0$, we define a *C-approximation* for this algorithm to be one that, every slot t and given $\boldsymbol{\Theta}(t)$, chooses $\alpha(t) \in \mathcal{A}_{\omega(t)}$, $\boldsymbol{\gamma}(t) \in \mathcal{X}$ to come within C of the infimum of the right hand side of (37). A 0-approximation would make decisions every slot t as follows:

- (Auxiliary Variables) Choose $\boldsymbol{\gamma}(t) \in \mathcal{X}$ to minimize:

$$Vf(\boldsymbol{\gamma}(t)) + \sum_{m=1}^{M} H_m(t)\gamma_m(t) + \sum_{l=1}^{L} Z_l(t)g_l(\boldsymbol{\gamma}(t))$$

- (Choosing $\alpha(t)$) Choose $\alpha(t) \in \mathcal{A}_{\omega(t)}$ to minimize:

$$\begin{aligned}\sum_{k=1}^{K} Q_k(t)[\hat{a}_k(\alpha(t), \omega(t)) - \hat{b}_k(\alpha(t), \omega(t))] \\ + \sum_{l=1}^{L} Z_l(t)\hat{y}_l(\alpha(t), \omega(t)) - \sum_{m=1}^{M} H_m(t)\hat{x}_m(\alpha(t), \omega(t)) \\ + V\hat{y}_0(\alpha(t), \omega(t))\end{aligned}$$

- (Queue Update) Update virtual queues $Z_l(t)$, $H_m(t)$ by (14), (15), and update actual queues $Q_k(t)$ by (6).

Theorem 5: ("Pure Dual" Performance) Suppose problem (1)-(5) is feasible, $\omega(t)$ is i.i.d. over slots, and the function $f(\boldsymbol{x})$ is convex. For simplicity, assume all initial queue values are 0. If we use any C-approximation on each slot t, then:

(a) All queues are mean rate stable, and hence all required constraints are satisfied. In particular, the equalities and inequalities (16)-(18) are ensured. Further, bounds on the expected queue size for all $t > 0$ are provided in inequality (39) of the proof.

(b) For all slots $t > 0$, the achieved time average expected cost satisfies:

$$\begin{aligned}\overline{y}_0(t) + f(\overline{\boldsymbol{x}}(t)) \leq y_0^{opt} + f^{opt} + \frac{B+C}{V} \\ + \sum_{m=1}^{M}\beta_m\frac{\mathbb{E}\left\{|H_m(t)|\right\}}{t}\end{aligned}$$

where we recall that $\overline{\boldsymbol{x}}(t)$ is defined in (19), and $\overline{y}_0(t)$ is defined similarly. The term $\mathbb{E}\left\{|H_m(t)|\right\}/t$ decays to zero at least as fast as $O(1/\sqrt{t})$, as shown in (39) of the proof.

(c) Suppose we use any C-approximation applied to a *variable-V* version of (37), with $V(t) = V_0(1+t)^d$ for any constants $V_0 > 0$ and $0 < d < 1$. Then this maintains mean rate stability of all queues, ensures all constraints (16)-(18), and yields *exact cost optimality*:

$$\lim_{t\rightarrow\infty}[\overline{y}_0(t) + f(\overline{\boldsymbol{x}}(t))] = y_0^{opt} + f^{opt}$$

We note that if an additional Slater condition is satisfied, then under the fixed V algorithm all queues can be shown to be strongly stable with averages $O(V)$ (see related results in [3][10][9]). In many flow control problems, the queues can be shown to be *deterministically bounded* by a constant of size $O(V)$, even without the Slater condition [10][4]. In the variable-V scenario, queues will not be strongly stable even if the Slater condition is satisfied, and hence the cost of achieving exact optimality is having infinite average backlog and delay.

Proof: (Theorem 5) Because every slot τ the C-approximation comes within C of the infimum in the right hand side of (37), we have:

$$\begin{aligned}\Delta(\mathbf{\Theta}(t)) + V\mathbb{E}\left\{y_0(t) + f(\boldsymbol{\gamma}(t))|\mathbf{\Theta}(t)\right\} \leq B + C \\ + \sum_{k=1}^{K} Q_k(t)\mathbb{E}\left\{a_k^*(t) - b_k^*(t)|\mathbf{\Theta}(t)\right\} \\ + \sum_{l=1}^{L} Z_l(t)\mathbb{E}\left\{y_l^*(t) + g_l(\boldsymbol{\gamma}^*(t))|\mathbf{\Theta}(t)\right\} \\ + \sum_{m=1}^{M} H_m(t)\mathbb{E}\left\{\gamma_m^*(t) - x_m^*(t)|\mathbf{\Theta}(t)\right\} \\ + V\mathbb{E}\left\{y_0^*(t) + f(\boldsymbol{\gamma}^*(t))|\mathbf{\Theta}(t)\right\}\end{aligned}$$

where $a_k^*(t)$, $b_k^*(t)$, $y_l^*(t)$, $\boldsymbol{\gamma}^*(t)$ are associated with any alternative decisions $\alpha^*(t) \in \mathcal{A}_{\omega(t)}$, $\boldsymbol{\gamma}^*(t) \in \mathcal{X}$. Fix $(\boldsymbol{y}^{opt}, \boldsymbol{x}^{opt}) \in \mathcal{Y}$ as the optimal solution in Theorem 1, yielding cost $f(\boldsymbol{x}^{opt}) = f^{opt}$. Plug the policy of (28)-(32), and for simplicity again assume $\epsilon = 0$. Also plug $\boldsymbol{\gamma}^*(t) = \boldsymbol{x}^{opt}$. Because these decisions are independent of $\mathbf{\Theta}(t)$, we have:

$$\begin{aligned}\Delta(\mathbf{\Theta}(t)) + V\mathbb{E}\left\{y_0(t) + f(\boldsymbol{\gamma}(t))|\mathbf{\Theta}(t)\right\} \leq B + C \\ + Vy_0^{opt} + Vf^{opt}\end{aligned} \quad (38)$$

Thus, $\Delta(\mathbf{\Theta}(t)) \leq \hat{D}$, where $\hat{D}$ is defined:

$$\hat{D} \triangleq B + C + V(y_0^{opt} - y_0^{min}) + V(f^{opt} - f_{min})$$

where we define y_0^{min} as a lower bound on $\mathbb{E}\left\{y_0(t)|\mathbf{\Theta}(t)\right\}$ over all possible $\alpha(t)$. This has the same structure as (34), and hence we obtain for all $t > 0$ (compare with (35)):

$$\frac{\mathbb{E}\left\{Q_k(t)\right\}}{t}, \frac{\mathbb{E}\left\{Z_l(t)\right\}}{t}, \frac{\mathbb{E}\left\{|H_m(t)|\right\}}{t} \leq \sqrt{\frac{2\hat{D}}{t}} \quad (39)$$

Taking a limit of (39) as $t \to \infty$ shows that all queues are mean rate stable, proving part (a).

To prove (b), take expectations of both sides of (38) and use iterated expectations to obtain:

$$\begin{aligned}\mathbb{E}\left\{L(\mathbf{\Theta}(t+1))\right\} - \mathbb{E}\left\{L(\mathbf{\Theta}(t))\right\} + V\mathbb{E}\left\{y_0(t) + f(\boldsymbol{\gamma}(t))\right\} \leq \\ B + C + V(y_0^{opt} + f^{opt})\end{aligned}$$

Fix $t > 0$, sum the above over $\tau \in \{0, \ldots, t-1\}$, and divide by t to yield:

$$\begin{aligned}\frac{\mathbb{E}\left\{L(\mathbf{\Theta}(t))\right\} - \mathbb{E}\left\{L(\mathbf{\Theta}(0))\right\}}{t} + V\overline{y}_0(t) + Vf(\overline{\boldsymbol{\gamma}}(t)) \leq \\ B + C + V(y_0^{opt} + f^{opt})\end{aligned}$$

where we have used Jensen's inequality in the convex function $f(\boldsymbol{\gamma})$. Noting that $L(\mathbf{\Theta}(0)) = 0$, $L(\mathbf{\Theta}(t)) \geq 0$, and dividing by V, we have:

$$\overline{y}_0(t) + f(\overline{\boldsymbol{\gamma}}(t)) \leq \frac{B+C}{V} + y_0^{opt} + f^{opt} \quad (40)$$

However, note from (8) that:

$$\begin{aligned} f(\overline{\boldsymbol{x}}(t)) &\leq f(\overline{\boldsymbol{\gamma}}(t)) + \sum_{m=1}^{M} \beta_m |\overline{x}_m(t) - \overline{\gamma}_m(t)| \\ &\leq f(\overline{\boldsymbol{\gamma}}(t)) + \sum_{m=1}^{M} \beta_m \mathbb{E}\left\{|H_m(t)|\right\}/t \quad (41)\end{aligned}$$

where (41) follows from (20). Using (41) and (40) proves the result of part (b). Part (c) is similar to Theorem 4. □

Remark 1: While Theorems 3-5 assume $\omega(t)$ is i.i.d. over slots, this assumption is not crucial, and the same policies can be shown to yield similar results by using a T-slot Lyapunov drift argument, under the assumption that $\omega(t)$ is ergodic with a "decaying memory property," so that averages over T slots are close to the steady state average [18][19][20][3].

Remark 2: We can remove auxiliary variables $\boldsymbol{\gamma}(t)$ and set $H_m(t) = 0$ for all t if: (i) We use the primal-dual method, there is no abstract set constraint (3), and $g_l(\boldsymbol{x}) = 0$ for all l, or (ii) We use the pure dual method, there is no abstract set constraint (3), and $g_l(\boldsymbol{x}) = f(\boldsymbol{x}) = 0$ for all l.

IV. An Alternative Search Method

Here we provide an alternative algorithm for the non-convex problem in the special case of utility maximization with entrywise non-decreasing utility functions. It involves three phases: Two phases are *convex* stochastic network optimizations for which we have stronger performance guarantees as in Section III-D. All non-convexity is isolated to the remaining phase that involves a *deterministic* non-convex problem for which any available solver can be employed.

To begin, let $f(\boldsymbol{x}) = -\phi(\boldsymbol{x})$, where $\phi(\boldsymbol{x})$ is a utility function to be maximized. Suppose that $\phi(\boldsymbol{x})$ is possibly non-concave, but is entrywise non-decreasing. For simplicity, assume $\hat{y}_l(\cdot) = 0$ for all l, so that the set $\mathcal{Y}$ consists of vectors $(\mathbf{0}, \boldsymbol{x})$. Let $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N$ be a collection of N different *search direction vectors*, being non-negative M-dimensional vectors. Assume for simplicity that for each $\boldsymbol{\theta}_n$, the set $\mathcal{Y}$ contains a point $\boldsymbol{x}$ such that $\boldsymbol{x} = \eta\boldsymbol{\theta}_n$ for some non-negative scalar η (this is trivially true if $\mathcal{Y}$ contains the origin). Define η_n as the largest value of η for which this is true, so that $\boldsymbol{x}_n \triangleq \eta_n\boldsymbol{\theta}_n$ is the largest value of $\boldsymbol{x}$ that we can push in this direction. Our new goal is to find a feasible policy that yields a utility $\phi(\overline{\boldsymbol{x}})$ that is a local maximum over the set $\mathcal{C}$ defined:

$$\mathcal{C} \triangleq Conv(\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\})$$

where $Conv(\cdot)$ denotes the convex hull. Because $\mathcal{Y}$ is convex, any point in this convex hull is achievable. The value of each $\boldsymbol{x}_n$ depends on the value of η_n, which is unknown if the stochastics of the network are unknown. However, suppose we know a bound on η_n, so that $0 \leq \eta_n \leq \eta_n^{max}$. If this is not a true bound, then we simply redefine $\boldsymbol{x}_n \triangleq \min[\eta_n, \eta_n^{max}]\boldsymbol{\theta}_n$.

Our algorithm has 3 phases:

- *Phase 1:* Determine the vectors $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ by solving N different stochastic network optimization problems, each with auxiliary variables $\eta_n(t)$:

$$\begin{aligned} &\text{Maximize:} && \overline{\eta}_n \\ &\text{Subject to:} && \overline{\boldsymbol{x}} = \overline{\eta}_n\boldsymbol{\theta}_n \\ &&& \overline{\boldsymbol{x}} \in \mathcal{X} \\ &&& g_l(\overline{\boldsymbol{x}}) \leq 0 \ \forall l \in \{1, \ldots, L\} \\ &&& \text{All queues } Q_k(t) \text{ are mean rate stable} \\ &&& \alpha(t) \in \mathcal{A}_{\omega(t)} \ , \ 0 \leq \eta_n(t) \leq \eta_n^{max} \ \forall t\end{aligned}$$

Each of these is a simple *convex* stochastic network optimization problem, and we can run the pure dual approach to provide accurate estimates $\tilde{\eta}_n$ of the optimal $\overline{\eta}_n$ value for each n. Note that this can be done using N different "virtual networks" in parallel, reusing the same $\omega(t)$ events that are observed. The equality constraints can either be treated as additional attributes $\boldsymbol{h}(t) = \boldsymbol{x}(t) - \eta_n(t)\boldsymbol{\theta}_n$ with an abstract set constraint $\overline{\boldsymbol{h}} \in \{\boldsymbol{0}\}$, or can simply be treated using virtual queues of the type (15).

- *Phase 2:* Define estimates $\tilde{\boldsymbol{x}}_n \triangleq \tilde{\eta}_n \boldsymbol{\theta}_n$. Use any *non-stochastic* optimizer to find a local minimum of the following *static optimization*:

$$\begin{aligned} \text{Maxmize:} \quad & \phi(\boldsymbol{x}) \\ \text{Subject to:} \quad & \textstyle\sum_{n=1}^{N} p_n \tilde{\boldsymbol{x}}_n = \boldsymbol{x} \\ & p_n \geq 0 \;\; \forall n \in \{1, \ldots, N\} \\ & \textstyle\sum_{n=1}^{N} p_n = 1 \end{aligned}$$

This type of problem can be solved, for example, via methods in [21], including Newton-type methods that do not restrict to small step sizes.

- *Phase 3:* Given the optimal $\boldsymbol{x}^*$ from phase 2, run an algorithm to solve the following stochastic network optimization problem:

$$\begin{aligned} \text{Maximize:} \quad & \overline{\eta} \\ \text{Subject to:} \quad & \overline{\boldsymbol{x}} \in \mathcal{X} \\ & \overline{\boldsymbol{x}} = \overline{\eta} \boldsymbol{x}^* \\ & g_l(\overline{\boldsymbol{x}}) \leq 0 \;\; \forall l \in \{1, \ldots, L\} \\ & \text{All queues } Q_k(t) \text{ are mean rate stable} \\ & \alpha(t) \in \mathcal{A}_{\omega(t)} \;\; , \;\; 0 \leq \eta_n(t) \leq 2 \; \forall t \end{aligned}$$

This is again a basic convex stochastic network optimization. It seeks to maximize the time average attribute $\overline{\boldsymbol{x}}$ in the direction of the $\boldsymbol{x}^*$ vector obtained in phase 2 (noticing that the resulting optimum might push even past the set $\mathcal{C}$). Here we limit our search to doubling $\boldsymbol{x}^*$, with the intent of obtaining a vector $\overline{\boldsymbol{x}}$ with utility $\phi(\overline{\boldsymbol{x}})$ at least as good as $\phi(\boldsymbol{x}^*)$. The $\boldsymbol{x}^*$ vector used in this phase might be periodically changed as a result of periodically running the other two phases in the background.

V. CONCLUSIONS

We have provided two approaches to constrained optimization of non-convex functions of time average attributes in a stochastic network. The first approach combines dual-based and primal-dual operations. It ensures all desired (convex) constraints are satisfied, and, provided that the algorithm converges, it also provides either a local optimum or near-local optimum, depending on whether we use a fixed V or variable V algorithm. While the method provides guarantees on the expected queue backlogs at any time t, it does not provide such guarantees for the utility, and it is not clear how much time is required for convergence to the local min. It also requires an infinite horizon time average that is not robust to system changes. This is in contrast to the "pure dual" approach to convex problems that provide explicit backlog and utility bounds for all time, and which are known to adapt to system changes [4][8]. The second approach uses 3 phases, two of which are stochastic and convex, while the remaining phase is purely deterministic but non-convex. These algorithms significantly extend our ability to optimize dynamic networks.

REFERENCES

[1] J.W. Lee, R. R. Mazumdar, and N. B. Shroff. Non-convex optimization and rate control for multi-class services in the internet. *IEEE/ACM Trans. on Networking*, vol. 13, no. 4, pp. 827-840, Aug. 2005.

[2] M. Chiang. Nonconvex optimization of communication systems. *Advances in Mechanics and Mathematics, Special volume on Strang's 70th Birthday, Springer*, vol. 3, 2008.

[3] L. Georgiadis, M. J. Neely, and L. Tassiulas. Resource allocation and cross-layer control in wireless networks. *Foundations and Trends in Networking*, vol. 1, no. 1, pp. 1-149, 2006.

[4] M. J. Neely. Universal scheduling for networks with arbitrary traffic, channels, and mobility. *ArXiv technical report*, arXiv:1001.0960v1, Jan. 2010.

[5] R. Agrawal and V. Subramanian. Optimality of certain channel aware scheduling policies. *Proc. 40th Annual Allerton Conference on Communication , Control, and Computing, Monticello, IL*, Oct. 2002.

[6] H. Kushner and P. Whiting. Asymptotic properties of proportional-fair sharing algorithms. *Proc. of 40th Annual Allerton Conf. on Communication, Control, and Computing*, 2002.

[7] A. Stolyar. Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm. *Queueing Systems*, vol. 50, pp. 401-457, 2005.

[8] M. J. Neely and R. Urgaonkar. Cross layer adaptive control for wireless mesh networks. *Ad Hoc Networks (Elsevier)*, vol. 5, no. 6, pp. 719-743, August 2007.

[9] M. J. Neely. *Dynamic Power Allocation and Routing for Satellite and Wireless Networks with Time Varying Channels*. PhD thesis, Massachusetts Institute of Technology, LIDS, 2003.

[10] M. J. Neely. Energy optimal control for time varying wireless networks. *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 2915-2934, July 2006.

[11] M. J. Neely, E. Modiano, and C. Li. Fairness and optimal stochastic control for heterogeneous networks. *Proc. IEEE INFOCOM*, March 2005.

[12] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transacations on Automatic Control*, vol. 37, no. 12, pp. 1936-1949, Dec. 1992.

[13] L. Tassiulas and A. Ephremides. Dynamic server allocation to parallel queues with randomly varying connectivity. *IEEE Transactions on Information Theory*, vol. 39, pp. 466-478, March 1993.

[14] A. Eryilmaz and R. Srikant. Fair resource allocation in wireless networks using queue-length-based scheduling and congestion control. *Proc. IEEE INFOCOM*, March 2005.

[15] M. J. Neely and R. Urgaonkar. Opportunism, backpressure, and stochastic optimization with the wireless broadcast advantage. *Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA*, Oct. 2008.

[16] A. Stolyar. Greedy primal-dual algorithm for dynamic resource allocation in complex networks. *Queueing Systems*, vol. 54, pp. 203-220, 2006.

[17] M. J. Neely. Max weight learning algorithms with application to scheduling in unknown environments. *arXiv:0902.0630v1*, Feb. 2009.

[18] L. Tassiulas. Scheduling and performance limits of networks with constantly changing topology. *IEEE Trans. on Inf. Theory*, May 1997.

[19] M. J. Neely, E. Modiano, and C. E Rohrs. Dynamic power allocation and routing for time varying wireless networks. *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 1, pp. 89-103, January 2005.

[20] R. Urgaonkar and M. J. Neely. Opportunistic scheduling with reliability guarantees in cognitive radio networks. *IEEE Transactions on Mobile Computing*, vol. 8, no. 6, pp. 766-777, June 2009.

[21] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 1995.