Time-Average Stochastic Optimization with Non-convex Decision Set and its Convergence

Sucha Supittayapornpong, Michael J. Neely
Department of Electrical Engineering
University of Southern California
Los Angeles, California
Email: supittay@usc.edu, mjneely@usc.edu

Abstract—This paper considers time-average stochastic optimization, where a time average decision vector, an average of decision vectors chosen in every time step from a time-varying (possibly non-convex) set, minimizes a convex objective function and satisfies convex constraints. This formulation has applications in networking and operations research. In general, time-average stochastic optimization can be solved by a Lyapunov optimization technique. This paper shows that the technique exhibits a transient phase and a steady state phase. When the problem has a unique vector of Lagrange multipliers, the convergence time can be improved. By starting the time average in the steady state, the convergence times become $O(1/\epsilon)$ under a locally-polyhedral assumption and $O(1/\epsilon^{1.5})$ under a locally-non-polyhedral assumption, where $\epsilon$ denotes the proximity to the optimal objective cost.

I. INTRODUCTION

Stochastic network optimization can be used to design dynamic algorithms that optimally control communication networks [1]. The technique has several unique properties which do not exist in a traditional convex optimization setting. In particular, the technique allows for a time-varying and possibly non-convex decision set. For example, it can treat a packet scheduling decisions, or a wireless system with varying channels and decision sets.

This paper considers time-average stochastic optimization, which is useful for example problems of network utility maximization [2]–[5], energy minimization [6], [7], and quality of information maximization [8].

Time $t \in \{0, 1, 2, \ldots \}$ is slotted. Define $\Omega$ to be a finite or countably infinite sample space of random states. Let $\omega(t) \in \Omega$ denote a random state at time $t$. Random state $\omega(t)$ is assumed to be independent and identically distributed (i.i.d.) across time slots. The steady state probability of $\omega \in \Omega$ is denoted by $\pi_\omega$. Let $I$ and $J$ be any positive integers. Each slot $t$, decision vector $x(t) = (x_1(t), \ldots, x_J(t))$ is chosen from a decision set $X_\omega(t)$. For any positive integer $T$, define $\overline{\pi}(T)$ as

$$\overline{\pi}(T) \triangleq \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}[x(t)].$$

The goal is to make decisions over time to solve:

\[
\begin{align*}
\text{Minimize} & \quad \limsup_{T \to \infty} f(\overline{\pi}(T)) \\
\text{Subject to} & \quad \limsup_{T \to \infty} g_j(\overline{\pi}(T)) \leq 0, \quad j \in \{1, \ldots, J\} \\
& \quad x(t) \in X_\omega(t), \quad t \in \{0, 1, 2, \ldots \}.
\end{align*}
\]

Here it is assumed that $X_\omega$ is a compact subset of $\mathbb{R}^J$ for each $\omega \in \Omega$. Assume $\cup_{\omega \in \Omega} X_\omega$ is bounded, and let $C$ be a compact set that contains it. The functions $f$ and $g_j$ are convex functions from $C$ to $\mathbb{R}$, where $\overline{A}$ denotes a convex hull of set $A$. Results in [1] imply that the optimal point can be achieved with an ergodic policy for which the limiting time average expectation exists.

An example of formulation (1) can be a resource allocation problem of a stochastic wireless uplink network. The goal is to achieve:

\[
\begin{align*}
\text{Minimize} & \quad \limsup_{T \to \infty} -\sum_{i=1}^{3} \log (\overline{\pi}_i(T)) \\
\text{Subject to} & \quad \limsup_{T \to \infty} \pi_i(T) \leq \pi_i^{(\text{min})}, \quad i \in \{1, 2, 3\} \\
& \quad x(t) \in X_\omega(t), \quad t \in \{0, 1, 2, \ldots \},
\end{align*}
\]

where $\pi_i^{(\text{min})}$ is the minimum rate for user $i$. In this example, $\Omega = \{1, 2\}$, $\pi_1 = 0.3$, $\pi_2 = 0.7$, $X_1 = \{(0, 0, 0), (2, 1, 0), (0, 2, 2)\}$, $X_2 = \{(0, 0, 0), (0, 1, 2), (1, 1, 1)\}$.

Solving formulation (1) using the stochastic network optimization framework does not require any statistical knowledge of the random states. However, if the steady state probabilities are known, the optimal objective cost of formulation (1) is identical the optimal cost of the following problem:

\[
\begin{align*}
\text{Minimize} & \quad f(x) \\
\text{Subject to} & \quad g_j(x) \leq 0, \quad j \in \{1, \ldots, J\} \\
& \quad x \in \overline{X},
\end{align*}
\]

where $\overline{X} \triangleq \sum_{\omega \in \Omega} \pi_\omega X_\omega$. Note that, for any $a, b \in \mathbb{R}$ and any sets $A$ and $B$, notation $aA+bB = \{aa+\beta b : a \in A, b \in B\}$.

Formulation (2) is convex; however, its optimal solution may not be in any of the sets $X_\omega$. In fact, determining whether $x$ is a member of $\overline{X}$ may already be a difficult task. This illustrates that traditional and state-of-the-art techniques for solving convex optimization cannot be applied directly to solve problem (1). Nevertheless, their convergence times are compelling to be mentioned for a purpose of comparison.

The convergence time of an algorithm is usually measured as a function of an $O(\epsilon)$-approximation to the optimal solution. For a convex optimization problem, several techniques utilizing a time average solution [9]–[11] have $O(1/\epsilon^2)$ convergence time. For unconstrained optimization without a more restrictive “strong convexity” property, the optimal first-order method...
average starts in the steady state phase, a solution converges in a small number of iterations on the structure of a dual function. The first case is when the drift-plus-penalty algorithm is shown to have a transient phase and a steady state phase under the locally-polyhedral assumption are provided in Section V. Simulations of the time-average stochastic problem (1) are preserved in problem (5). Let \( f^\text{opt} \) be the optimal objective cost of problem (1).

**Theorem 1:** The time-average stochastic problem (1) and the auxiliary problem (5) have the same optimal cost, \( f^\text{opt} = f^\text{aux} \).

**Proof:** Please see the full proof in [22].

### B. Lyapunov optimization

The auxiliary problem (5) can be solved by the Lyapunov optimization technique [1]. Define \( W_j(t) \) and \( Z_i(t) \) to be virtual queues of the first and second constraints of problem (5) with update dynamics:

\[
W_j(t + 1) = [W_j(t) + g_j(y(t))]_+, \quad j \in \{1, \ldots, J\} \tag{6}
\]

\[
Z_i(t + 1) = Z_i(t) + x_i(t) - y_i(t) \quad i \in \{1, \ldots, I\}, \tag{7}
\]

where operator \([\cdot]_+\) is the projection to a corresponding nonnegative orthant.

For ease of notations, let \( W(t) \triangleq (W_1(t), \ldots, W_J(t)) \), \( Z(t) \triangleq (Z_1(t), \ldots, Z_I(t)) \), and \( g(y) \triangleq (g_1(y), \ldots, g_J(y)) \) respectively be the vectors of virtual queues \( W_j(t) \), \( Z_i(t) \), and functions \( g_j(y) \).

Define Lyapunov function (8) and Lyapunov drift (9) as

\[
L(t) \triangleq \frac{1}{2} \|W(t)\|^2 + \|Z(t)\|^2 \tag{8}
\]

\[
\Delta(t) \triangleq L(t + 1) - L(t). \tag{9}
\]
Let notation $A^\top$ denote the transpose of vector $A$. Define $C \triangleq \sup_{x \in C, y \in \mathcal{Y}} \left[ \|g(y)^\top\|^2 + \|x - y\|^2 \right]/2$, which is finite.

**Lemma 1:** For every $t \in \{0, 1, 2, \ldots\}$, the Lyapunov drift is upper bounded by

$$
\Delta(t) \leq C + W(t)^\top g(y(t)) + Z(t)^\top [x(t) - y(t)].
$$

*Proof:* The proof is similar to the one in [18] and [22].

Let $V > 0$ be any positive real number representing a parameter of an algorithm solving problem (5). The drift-plus-penalty term is defined as $\Delta(t) + Vf(y(t))$. Applying Lemma 1, the drift-plus-penalty term is bounded for every time $t$ by

$$
\Delta(t) + Vf(y(t)) \leq C + W(t)^\top g(y(t)) + Z(t)^\top [x(t) - y(t)] + Vf(y(t)).
$$

(11)

**C. Drift-plus-penalty algorithm**

Let $W^0$ and $Z^0$ be the initial condition of $W(0)$ and $Z(0)$ respectively. Every time step, the Lyapunov optimization technique observes the current realization of random state $\omega(t)$ before choosing decisions $x(t) \in \mathcal{X}_\omega(t)$ and $y(t) \in \mathcal{Y}$ that minimize the right-hand-side of (11). The drift-plus-penalty algorithm is summarized in Algorithm 1.

| Initialize $V, W(0) = W^0, Z(0) = Z^0$. |
| for $t \in \{0, 1, 2, \ldots\}$ do |
| Observe $\omega(t)$ |
| $x(t) = \text{argmin}_{x \in \mathcal{X}_\omega(t)} Z(t)^\top x$ |
| $y(t) = \text{argmin}_{y \in \mathcal{Y}} [Vf(y) + W(t)^\top g(y) - Z(t)^\top y]$ |
| $W(t + 1) = [W(t) + g(y(t))]_+$ |
| $Z(t + 1) = Z(t) + x(t) - y(t)$ |
| end |

**Algorithm 1:** Drift-plus-penalty algorithm solving (5).

III. BEHAVIORS OF DRIFT-PLUS-PENALTY ALGORITHM

Starting from $(W(0), Z(0))$, Algorithm 1 reaches the steady state when vector $(W(t), Z(t))$ concentrates around a specific set (defined in Section III-A). The transient phase is the period before this concentration. Note that this behavior is different from the deterministic case in [18], where $(W(t), Z(t))$ in the steady state is contained in a specific set.

A. Embedded Formulation

A convex optimization problem, called embedded formulation, is considered. This idea is inspired by [19].

**Minimize** $f(y)$

Subject to $g_j(y) \leq 0, j \in \{1, \ldots, J\}$

$y = \sum_{\omega \in \Omega} \pi_\omega x^\omega$

$y \in \mathcal{Y}, x^\omega \in \mathcal{X}_\omega, \omega \in \Omega$.

Note that formulation (12) contains multiple sets $\mathcal{X}_\omega$ and is more complex than its deterministic version in [18].

This formulation has a dual problem, whose properties are used in convergence analysis. Let $w \in \mathbb{R}_+^J$ and $z \in \mathbb{R}^I$ be the vectors of dual variables associated with the first and second constraints of problem (12). The Lagrangian is defined as

$$
\Gamma(\{x^\omega\}_{\omega \in \Omega}, y, w, z) = \sum_{\omega \in \Omega} \pi_\omega \left[ f(y) + w^\top g(y) + z^\top (x^\omega - y) \right].
$$

The dual function of problem (12) is

$$
d(w, z) = \inf_{y \in \mathcal{Y}, x \in \mathcal{X}_\omega} \Gamma(\{x^\omega\}_{\omega \in \Omega}, y, w, z)
= \sum_{\omega \in \Omega} \pi_\omega d_\omega(w, z),
$$

(13)

where $d_\omega(w, z)$ is defined in (14) and all of the minimizing solutions $y$ take the same value.

$$
d_\omega(w, z) \triangleq \inf_{y \in \mathcal{Y}, x \in \mathcal{X}_\omega} \left[ f(y) + w^\top g(y) + z^\top (x - y) \right].
$$

(14)

Define the solution to the infimum in (14) as

$$
y^\star(z) \triangleq \arg\inf_{y \in \mathcal{Y}} \left[ f(y) + w^\top g(y) - z^\top y \right],
$$

(15)

$$
x^\star_\omega(z) \triangleq \arg\inf_{x \in \mathcal{X}_\omega} z^\top x.
$$

(16)

Finally, the dual problem of formulation (12) is

**Maximize** $d(w, z)$

**Subject to** $(w, z) \in \mathbb{R}_+^J \times \mathbb{R}^I$.

Problem (17) has an optimal solution that may not be unique. A set of these optimal solutions, which are vectors of Lagrange multipliers, can be used to analyze the transient time. However, to simplify the proofs and notations, an uniqueness assumption is defined below. Define $\lambda^\star(w, z)$ as a concatenation vector of $w$ and $z$.

**Assumption 3:** Dual problem (17) has a unique vector of Lagrange multipliers denoted by $\lambda^\star(w^\star, z^\star)$.

This assumption is assumed throughout Section IV and Section V. Note that this is a mild assumption when practical systems are considered, e.g., [5], [19]. Furthermore, simulation results in [18] evince that this assumption may not be needed.

To prove the main result of this section, a useful property of $d_\omega(w, z)$ is derived. Define $h(x, y) \triangleq (g(y), x - y)$.

**Lemma 2:** For any $\lambda = (w, z) \in \mathbb{R}_+^J \times \mathbb{R}^I$ and $\omega \in \Omega$, it holds that

$$
d_\omega(\lambda) \leq d_\omega(\lambda^\star) + h(x^\star_\omega(z), y^\star(z))^\top [\lambda^\star - \lambda].
$$

(18)

*Proof:* From (14), it follows, for any $\lambda = (w, z) \in \mathbb{R}_+^J \times \mathbb{R}^I$ and $(x, y) \in \mathcal{X}_\omega \times \mathcal{Y}$, that

$$
d_\omega(\lambda) \leq f(y) + h(x, y)^\top \lambda^\star
= f(y) + h(x, y)^\top \lambda + h(x, y)^\top [\lambda^\star - \lambda].
$$

Setting $(x, y) = (x^\star_\omega(w, z), y^\star_\omega(z))$, as defined in (15) and (16), and using (14) proves the lemma.
The following lemma ties the virtual queues of Algorithm 1 to the Lagrange multipliers. Given the generated $W(t)$ and $Z(t)$ of Algorithm 1, define $Q(t) \triangleq (W(t), Z(t))$ as a concatenation of these vectors. The queue dynamics (6) and (7) are equivalent to

$$Q(t + 1) = \mathcal{P}[Q(t) + h(x(t), y(t))]$$

where $\mathcal{P}[(W, Z)]$ denotes the projection of the concatenated vector $(W, Z)$ onto the set $\mathbb{R}_+^t \times \mathbb{R}_+^t$.

**Lemma 3:** The following holds for every $t \in \{0, 1, 2, \ldots\}$:

$$\mathbb{E} \left[ ||Q(t + 1) - V \lambda^*||^2 | Q(t) \right] \leq ||Q(t) - V \lambda^*||^2 + 2C + 2V[d(Q(t)/V) - d(\lambda^*)].$$

**Proof:** The non-expansive projection [21] implies that

$$||Q(t + 1) - V \lambda^*||^2 \leq ||Q(t) + h(x(t), y(t)) - V \lambda^*||^2$$

$$= ||Q(t) - V \lambda^*||^2 + ||h(x(t), y(t))||^2$$

$$+ 2h(x(t), y(t))^\top [Q(t) - V \lambda^*]$$

$$\leq ||Q(t) - V \lambda^*||^2 + 2C + 2h(x(t), y(t))^\top [Q(t) - V \lambda^*].$$

(19)

From (14), when $\lambda = Q(t)/V$, we have

$$d_{\omega(t)}(Q(t)/V) = \inf_{y \in \mathcal{Y}, x \in \mathcal{T}_{\omega(t)}}\left[ f(y) + \frac{W(t)}{V}^\top g(y) + \frac{Z(t)}{V}^\top (x - y) \right],$$

so $y^*(W(t)/V, Z(t)/V) = y(t)$ and $x^*_{\omega(t)}(Z(t)/V) = x(t)$ where $(y^*(W(t)/V, Z(t)/V), x^*_{\omega(t)}(Z(t)/V))$ is defined in (15) and (16), and $(y(t), x(t))$ is the decision from Algorithm 1. Therefore, property (18) implies that

$$h(x(t), y(t))^\top [Q(t) - V \lambda^*] \leq V [d_{\omega(t)}(Q(t)/V) - d_{\omega(t)}(\lambda^*)].$$

Applying the above inequality on the last term of (19) gives

$$||Q(t + 1) - V \lambda^*||^2 \leq ||Q(t) - V \lambda^*||^2 + 2C + 2V[d_{\omega(t)}(Q(t)/V) - d_{\omega(t)}(\lambda^*)].$$

Taking a conditional expectation given $Q(t)$ proves the lemma:

$$\mathbb{E} \left[ ||Q(t + 1) - V \lambda^*||^2 | Q(t) \right] \leq ||Q(t) - V \lambda^*||^2 + 2C + 2V \sum_{\omega \in \Omega} \pi_\omega [d_{\omega}(Q(t)/V) - d_{\omega}(\lambda^*)].$$

B. T-slot convergence

For any positive integer $T$ and any starting time $t_0$, define the T-slot average starting at $t_0$ as

$$\overline{x}(t_0, T) \triangleq \frac{1}{T} \sum_{t=t_0}^{t_0+T-1} x(t).$$

This average leads to the following convergence bounds.

**Theorem 2:** Let $(Q(t))_{t=0}^\infty$ be a sequence generated by Algorithm 1. For any positive integer $T$ and any starting time $t_0$, the objective cost converges as

$$\mathbb{E} \left[ f(\overline{x}(t_0, T)) \right] - f^{(\text{opt})} \leq \frac{M_f}{T} \mathbb{E} \left[ ||Z(t_0 + T) - Z(t_0)|| \right]$$

$$+ \frac{1}{2TV} \mathbb{E} \left[ ||Q(t_0)||^2 - ||Q(t_0 + T)||^2 \right] + C.$$

(20)

and the constraint violation for every $j \in \{1, \ldots, J\}$ is

$$\mathbb{E} [g_j(\overline{x}(t_0, T))] \leq \frac{1}{T} \mathbb{E} [W_j(t_0 + T) - W_j(t_0)]$$

$$+ \frac{M_{g_j}}{T} \mathbb{E} [||Z(t_0 + T) - Z(t_0)||].$$

(21)

**Proof:** Please see the full proof in [22].

To interpret Theorem 2, the following concentration bound is provided. It is proven in [23].

C. Concentration bound

**Theorem 3:** Let $K(t)$ be a real random process over $t \in \{0, 1, 2, \ldots\}$ satisfying

$$|K(t + 1) - K(t)| \leq \delta$$

and

$$\mathbb{E} [K(t + 1) - K(t)] \leq \left\{ \begin{array}{ll} \delta & , K(t) < \gamma, \\ -\beta & , K(t) \geq \gamma, \end{array} \right.$$

for some positive real-valued $\delta, \gamma$, and $0 < \beta \leq \delta$.

Suppose $K(0) = k_0$ (with probability 1) for some $k_0 \in \mathbb{R}$. Then for every time $t \in \{0, 1, 2, \ldots\}$, the following holds:

$$\mathbb{E} \left[ e^{cK(t)} \right] \leq D + (e^{c_{k_0}} - D) \rho^t$$

where $0 < \rho < 1$ and constants $r, \rho$, and $D$ are:

$$r \triangleq \frac{\beta}{(\delta^2 + \delta \beta/3)}, \quad \rho \triangleq 1 - \frac{r \beta}{2}$$

$$D \triangleq \frac{(e^{r \delta} - \rho)e^{r \gamma}}{1 - \rho}.$$

In this paper, random process $K(t)$ is defined to be the distance between $Q(t)$ and the vector of Lagrange multipliers, so $K(t) \triangleq ||Q(t) - V \lambda^*||$ for every $t \in \{0, 1, 2, \ldots\}$.

**Lemma 4:** It holds for every $t \in \{0, 1, 2, \ldots\}$ that

$$\mathbb{E} [K(t + 1) - K(t)] \leq \sqrt{2C}$$

$$\mathbb{E} [K(t + 1) - K(t)|K(t)] \leq \sqrt{2C}.$$

**Proof:** The first part is proven in two cases.
i) If $K(t+1) \geq K(t)$, the non-expansive projection implies

$$|K(t+1) - K(t)| = K(t+1) - K(t) \leq \|Q(t) - Q(t+1) + V\lambda^* - K(t+1)\| \leq \|h(x(t), y(t))\| \leq 2\sqrt{C}.$$ 

Therefore, $|K(t+1) - K(t)| \leq \sqrt{2C}$. Using $K(t+1) - K(t) \leq |K(t+1) - K(t)|$ proves the second part.

Lemma 4 prepares $K(t)$ for Theorem 3. The only constants left to be specified are $\beta$ and $\gamma$, which depend on properties of dual function (13).

IV. LOCALLY-POLYHEDRAL DUAL FUNCTION

This section analyzes the transient time and a convergence result in the steady state. Dual function (13) in this section is assumed to satisfy a locally-polyhedral property, introduced in [19]. This property is illustrated in Figure 1. It holds when $f$ and each $g_j$, for $j \in \{1, \ldots, J\}$ are either linear or piece-wise linear.

Assumption 4: Let $\lambda^*$ be the unique Lagrange multiplier vector. There exists $L_P > 0$ such that dual function (13) satisfies, for any $\lambda \in \mathbb{R}_+^d \times \mathbb{R}^J$

$$d(\lambda^*) \geq d(\lambda) + L_P\|\lambda - \lambda^*\|.$$  

(22)

Note that, by concavity of the dual function, if inequality (22) holds locally about $\lambda^*$, it must also hold globally. The subscript P denotes “polyhedral.”

A. Transient time

The progress of $Q(t)$ at each step can be analyzed. Define

$$B_P \triangleq \max \left[ \frac{L_P}{2}, \frac{2C}{r_P} \right].$$  

(23)

Lemma 5: Under Assumptions 3 and 4, whenever $\|Q(t) - V\lambda^*\| \geq B_P$, the following holds

$$\mathbb{E}[\|Q(t+1) - V\lambda^*\| |Q(t)] \leq \|Q(t) - V\lambda^*\| - \frac{L_P}{2}.$$  

Proof: If condition

$$2C + 2V[d(Q(t)/V) - d(\lambda^*)] \leq -L_P\|Q(t) - V\lambda^*\| + L_P^2/4.$$  

(24)

is true, Lemma 3 implies

$$\mathbb{E} \left[ \|Q(t+1) - V\lambda^*\|^2 |Q(t)\right] \leq \|Q(t) - V\lambda^*\|^2 - L_P\|Q(t) - V\lambda^*\| + L_P^2/4 = \|Q(t) - V\lambda^*\| - L_P/2.$$  

Applying Jensen’s inequality [21] on the left-hand-side yields

$$\mathbb{E}[\|Q(t+1) - V\lambda^*\| |Q(t)] \leq \|Q(t) - V\lambda^*\| - L_P/2.$$  

(25)

It requires to show that condition (24) holds whenever $\|Q(t) - V\lambda^*\| \geq B_P$. Assumption 4 implies that

$$2C + 2V[d(Q(t)/V) - d(\lambda^*)] \leq 2C - 2L_P\|Q(t) - V\lambda^*\|.$$  

From the definition of $B_P$ in (23), $\|Q(t) - V\lambda^*\| \geq B_P$ implies $L_P\|Q(t) - V\lambda^*\| \geq 2C$ and

$$2C + 2V[d(Q(t)/V) - d(\lambda^*)] \leq -L_P\|Q(t) - V\lambda^*\|,$$  

which implies (24).

Lemma 5 implies that, in expectation, $Q(t)$ proceeds closer to $V\lambda^*$ in the next step when the distance between them is at least $B_P$. This implication means that $Q(t)$ concentrates around $V\lambda^*$ in the steady state.

B. Convergence time in a steady state

Define constants $r_P, r_P, D_P, U_P, U_P'$ as

$$r_P \triangleq \frac{3L_P}{12C + L_P\sqrt{2C}}, \quad r_P \triangleq 1 - \frac{r_P L_P}{4}, \quad r_P \triangleq \frac{e^{r_P B_P}(e^{r_P \sqrt{2C}} - r_P)}{1 - r_P}, \quad U_P \triangleq \frac{\log(D_P + 1)}{r_P}, \quad U_P' \triangleq \frac{2(D_P + 1)}{r_P^2}.$$  

(26, 27, 28)

Given the initial condition $Q^0 \triangleq (W^0, Z^0)$, define

$$T_P \triangleq \left[ \frac{r_P \|Q^0 - V\lambda^*\|}{\log(1/r_P)} \right].$$  

(29)

where constants $r_P$ and $\rho_P$ are defined in (26). The value $T_P$ is $O(V)$. The next lemma shows $T_P$ can be interpreted as the transient time, so that desirable “steady state” bounds hold after this time.

Lemma 6: Suppose Assumptions 3 and 4 hold. Given the initial condition $Q^0 \triangleq (W^0, Z^0)$, for any time $t \geq T_P$ when $T_P$ is defined in (29), the following holds

$$\mathbb{E}[\|Q(t) - V\lambda^*\|] \leq U_P$$  

(30)

$$\mathbb{E}[\|Q(t) - V\lambda^*\|^2] \leq U_P'.$$  

(31)

where constants $U_P$ and $U_P'$ are defined in (28).

Proof: Recall that $K(t) \triangleq \|Q(t) - V\lambda^*\|$. From Lemmas 4 and 5, constants in Theorem 3 are $\delta = \sqrt{2C}, \gamma = B_P$, and $\beta = L_P/2$. Theorem 3 implies, for any $t \geq 0$, that

$$\mathbb{E}[e^{r_P K(t)}] \leq D_P + e^{r_P \rho_P} \leq D_P + e^{r_P \rho_P}.$$  

(32)
where \( k_0 = K(0) = \|Q^0 - V\lambda^*\| \) and constants \( r_p, \rho_p, D_p \) are defined in (26) and (27). We then show that
\[
e^{r_p k_0} \rho_p^T \leq 1 \quad \forall t \geq T_p.
\]
(33)

Inequality \( e^{r_p k_0} \rho_p^T \leq 1 \) is equivalent to \( t \geq \frac{r_p \|Q^0 - V\lambda^*\|}{\log(1/\rho_p)} \) by arithmetic and the fact that \( \log(1/\rho_p) > 0 \). From the definition of \( T_p \) in (29), it holds that \( T_p \geq \frac{r_p \|Q^0 - V\lambda^*\|}{\log(1/\rho_p)} \), and the result (33) follows.

From (33), inequality (32) becomes
\[
E \left[ e^{r_p K(t)} \right] \leq D_p + 1 \quad \forall t \geq T_p.
\]
(34)

Jensen’s inequality implies that \( e^{r_p E[K(t)]} \leq E \left[ e^{r_p K(t)} \right] \), and we have \( e^{r_p E[K(t)]} \leq D_p + 1 \). Taking logarithm and dividing by \( r_p \) proves (30).

Chernoff bound (see for example in [24]) implies that, for any \( m \in \mathbb{R}_+ \),
\[
\mathbb{P} [K(t) \geq m] \leq e^{-rm} E \left[ e^{r_p K(t)} \right] \leq e^{-rm} (D_p + 1) \quad \forall t \geq T_p
\]
(35)
where the last inequality uses (34). Since \( K(t)^2 \) is always non-negative, it can be shown that \( E[K(t)^2] = 2 \int_0^\infty m^2 \mathbb{P} [K(t) \geq m] dm \) by the integration by parts. Using (35), we have
\[
E \left[ K(t)^2 \right] \leq 2(D_p + 1) \int_0^\infty me^{-rm} dm.
\]
Performing the integration by parts proves (31).

The above lemma implies that, in the steady state, the expected distance and square distance between \( Q(t) \) and the vector of Lagrange multipliers are bounded by constants that do not depend on \( V \). This phenomenon leads to an improved convergence time when the average is performed in the steady state. A useful result is derived below the main theorem.

**Lemma 7:** For any times \( t_1 \) and \( t_2 \), it holds that
\[
E \left[ \|Q(t_1)\|^2 - \|Q(t_2)\|^2 \right] \leq E \left[ \|Q(t_1) - V\lambda^*\|^2 \right] + 2\|V\lambda^*\| E \left\{ \|Q(t_1) - V\lambda^*\| + \|Q(t_2) - V\lambda^*\| \right\}.
\]
**Proof:** It holds for any \( Q \in \mathbb{R}_+^4 \times \mathbb{R}^J \) that
\[
\|Q\|^2 = \|Q - V\lambda^*\|^2 + 2\|V\lambda^*\|^2 + 2(Q - V\lambda^*)^\top (V\lambda^*).
\]
Using the above equality with \( Q_1, Q_2 \in \mathbb{R}_+^4 \times \mathbb{R}^J \) leads to
\[
\|Q_1\|^2 - \|Q_2\|^2
\leq \|Q_1 - V\lambda^*\|^2 - \|Q_2 - V\lambda^*\|^2 + 2(Q_1 - Q_2)^\top (V\lambda^*)
\leq \|Q_1 - V\lambda^*\|^2 + 2\|Q_1 - Q_2\| \|V\lambda^*\|
\leq \|Q_1 - V\lambda^*\|^2 + 2\|V\lambda^*\| \|Q_1 - V\lambda^*\| + \|Q_2 - V\lambda^*\|.
\]
Taking an expectation proves the lemma.

Finally, the convergence in the steady state is analyzed.

**Theorem 4:** Suppose Assumptions 3 and 4 hold. For any time \( t_0 \geq T_p \) and positive integer \( T \), the objective cost converges as
\[
E \left[ f(\pi(t_0, T)) \right] - f^{(\text{opt})} \leq \frac{2M_f U_p}{T} + \frac{U_p + 4VU_p\|\lambda^*\|}{2TV} + \frac{C}{V}
\]
(36)
and the constraint violation is upper bounded by
\[
E \left[ g_j(\pi(t_0, T)) \right] \leq \frac{2U_p}{T} + \frac{2M_gj U_p}{T}.
\]
(37)

**Proof:** From Theorem 2, the objective cost converges as (20). Since \( T_p \leq t_0 < t_0 + T \), we use results in Lemma 6 to upper bound \( E \left[ \|Q(t) - V\lambda^*\|^2 \right] \) and \( E \left[ \|Q(t) - V\lambda^*\|^2 \right] \) for \( t_0 \) and \( t_0 + T \). Terms in the right-hand-side of (20) are bounded by
\[
E \left[ \|Z(t_0 + T) - Z(t_0)\| \right] \leq E \left[ \|Q(t_0 + T) - Q(t_0)\| \right]
\leq E \left[ K(t_0 + T) + K(t_0) \right] \leq 2U_p.
\]
(38)

Lemma 7 implies that
\[
E \left[ \|Q(t_0)\|^2 - \|Q(t_0 + T)\|^2 \right]
\leq E \left[ K(t_0)^2 + 2\|V\lambda^*\| \|K(t_0) + K(t_0 + T)\| \right]
\leq U_p + 4VU_p\|\lambda^*\|.
\]
(39)
Substituting bounds (38) and (39) into (20) proves (36).

The constraint violation converges as (21) where \( T_p \leq t_0 < t_0 + T \). Using Lemma 6, the last term in the right-hand-side of (21) is bounded in (38). The first term is bounded by
\[
E \left[ W_j(t_0 + T) - W_j(t_0) \right]
\leq E \left[ W_j(t_0 + T) - V\lambda^* \right] + |W_j(t_0) - V\lambda^*|
\leq E \left[ K(t_0 + T) + K(t_0) \right] \leq 2U_p.
\]
Substituting the above bound and (38) into (21) proves (37).

The implication of Theorem 4 is as follows. When the average starts in the steady state, the deviation from the optimal cost is \( O(1/T^1/2) \), and the constraint violation is bounded by \( O(1/T) \). By setting \( V = 1/e \) and \( T = 1/e \), both optimal cost and constrain violation achieve \( O(e) \)-approximation, and the convergence time is \( O(1/e) \) slots. Note that this setting yields \( O(1/e) \) transient time, since \( T_p = O(V) = O(1/e) \).

V. LOCALLY-NON-POLYHEDRAL DUAL FUNCTION

The dual function (13) in Section V is assumed to satisfy a locally-non-polyhedral property, modified from [19]. This property is illustrated in Figure 1.

**Assumption 5:** Let \( \lambda^* \) be the unique Lagrange multiplier vector. There exist \( S > 0 \) and \( L_N > 0 \) such that, whenever \( \lambda \in \mathbb{R}_+^4 \times \mathbb{R}^J \) and \( \|\lambda - \lambda^*\| \leq S \), dual function (13) satisfies
\[
d(\lambda) \geq d(\lambda) + L_N \|\lambda - \lambda^*\|^2.
\]
Note that the subscript \( N \) denotes “non-polyhedral.”

It can be shown that Assumption 5 implies
\[
d(\lambda^*) \geq d(\lambda) + SL_N \|\lambda - \lambda^*\|^2
\]
for all \( \lambda \in \mathbb{R}_+^4 \times \mathbb{R}^J \) and \( \|\lambda - \lambda^*\| > S \).\(^1\)

\(^1\)We would like to thank Hao Yu for noticing this fact.
A. Transient time

The progress of $Q(t)$ at each step can be analyzed. Define

$$B_N(V) \triangleq \max \left\{ \frac{1}{\sqrt{V}}, \sqrt{\left( 1 + \sqrt{1 + 4L_N C} \right)} \right\},$$

$$B'_N \triangleq \max \left\{ \frac{SL_N}{2}, \frac{2C}{SL_N} \right\}.$$  

Lemma 8: Suppose Assumptions 3 and 5 hold. When $V$ is large enough to ensure both $B_N(V) \leq SV$ and $B'_N \leq SV$, the following holds

$$\mathbb{E} \left[ \|Q(t + 1) - V \lambda^*\| \right] \leq \mathbb{E} \left[ \|Q(t) - V \lambda^*\| \right].$$

(40)

Proof: Please see the full proof in [22].

The interpretation of Lemma 8 is similar to Lemma 5 except that $B_N(V)$ and the negative drift (40) are functions of $V$. Nevertheless, Lemma 8 implies that $Q(t)$ concentrates around $V \lambda^*$ in the steady state.

B. Convergence time in a steady state

Define constants $r_N(V), \rho_N(V), D_N(V), U_N(V), U'_N(V)$ as

$$r_N(V) \triangleq \frac{3}{6C\sqrt{V} + 2C}, \quad \rho_N(V) \triangleq 1 - \frac{r_N(V)}{2\sqrt{V}},$$

$$D_N(V) \triangleq \frac{\epsilon^{r_N(V)} B_N(V)}{1 - \rho_N(V)}, \quad U_N(V) \triangleq \frac{\log \left( \frac{D_N(V) + 1}{r_N(V)} \right)}{r_N(V)},$$

$$U'_N(V) \triangleq \frac{2(D_N(V) + 1)}{r_N(V)^2}.\quad (43)$$

Given the initial condition $Q^0 \triangleq (W^0, Z^0)$, define the transient time for a locally-nonthreshold dual function as

$$T_N \triangleq \left[ \frac{r_N(V) \|Q^0 - V \lambda^*\|}{\log \left( 1/\rho_N(V) \right)} \right],$$

where constants $r_N(V)$ and $\rho_N(V)$ are defined in (41). Definition (44) implies that the transient time under the locally-nonthreshold assumption is $O(V^{1.5})$. This $T_N$ can be interpreted as the transient time, so that desirable bounds hold after this time.

Lemma 9: Suppose Assumptions 3 and 5 hold. When $V$ is large enough to ensure $B_N(V) \leq SV$, $B'_N \leq SV$, and $\sqrt{V} \geq 2/SL_N$, for any time $t \geq T_N$, the following holds

$$\mathbb{E} \left[ \|Q(t) - V \lambda^*\| \right] \leq U_N(V)$$

$$\mathbb{E} \left[ \|Q(t) - V \lambda^*\|^2 \right] \leq U'_N(V)$$

where $U_N(V)$ and $U'_N(V)$ are defined in (43).

Proof: Please see the full proof in [22].

The convergence results in the steady state are as follows.

Theorem 5: Suppose Assumptions 3 and 5 hold. When $V$ is large enough to ensure $B_N(V) \leq SV$, $B'_N \leq SV$, and $\sqrt{V} \geq 2/SL_N$, then for any time $t_0 \geq T_N$ and any positive integer $T$, the objective cost converges as

$$\mathbb{E} \left[ f(\pi(t_0), T) \right] - f^{(opt)} \leq \frac{2MfU_N(V)}{T} + \frac{U_N(V) + 4VU_N(V)\|\lambda^*\|}{2TV} + \frac{C}{V}$$

(47)

and the constraint violation is upper bounded by

$$\mathbb{E} \left[ g_j(\pi(t_0), T) \right] \leq \frac{2U_N(V)}{T} + \frac{2Mg_jU_N(V)}{T}. \quad (48)$$

Proof: Please see the full proof in [22].

The implication of Theorem 5 is as follows. When the average starts in the steady state, the deviation from the optimal cost is $O(\sqrt{V}/T + 1/V)$, and the constraint violation is bounded by $O(\sqrt{V}/T)$. Note that this can be shown by substituting $B_N(V), r_N(V), \rho_N(V), D_N(V)$ into (47) and (48). By setting $V = 1/\epsilon$ and $T = 1/\epsilon^{1.5}$, this achieves an $O(\epsilon)$-approximation with transient time and convergence time of $O(1/\epsilon^{1.5})$.

VI. SIMULATION

A. Staggered Time Averages

In order to take advantage of the improved convergence times, computation of time averages must be started in the steady state phase shortly after the transient time. To achieve this performance without knowing the end of the transient phase, time averages can be restarted over successive frames whose frame lengths increase geometrically. For example, if one triggers a restart at times $2^k$ for integers $k$, then a restart is guaranteed to occur within a factor of 2 of the time of the actual end of the transient phase.

B. Results

This section illustrates the convergence times of the drift-plus-penalty Algorithm 1 under locally-polyhedron and locally-nonthreshold assumptions. Let $\Omega = \{0, 1, 2\}, \chi_0 = \{(0, 0)\}, \chi_1 = \{(-5, 0), (0, 10)\}, \chi_2 = \{(0, -10), (5, 0)\},$ and $\{\pi_0, \pi_1, \pi_2\} = (0, 1, 0.6, 0.3)$. A formulation is

Minimize $f(T)$

Subject to $\limsup_{T \to \infty} [-2\pi_1(T) - \pi_2(T)] \leq -1.5$

where function $f$ will be given for different cases.

Under locally-polyhedron assumption, let $f(x) = 1.5x_1 + x_2$ be the objective function of problem (49). In this setting, the optimal value is 1.25 where $\lim_{T \to \infty} \pi_1(T) = \lim_{T \to \infty} \pi_2(T) = 0.5$. Figure 2 shows the values of objective and constraint functions of time-average solutions. It is easy to see the improved convergence time $O(1/\epsilon)$ from the staggered time averages compared to the convergence time $O(1/\epsilon^2)$ of Algorithm 1.

Under locally-nonthreshold assumption, let $f(x) = x_1^2 + x_2^2$ be the objective function of problem (49). Note that the
optimal value of this problem is 0.5 where \( \lim_{T \to \infty} \mathcal{J}_1(T) = \lim_{T \to \infty} \mathcal{J}_2(T) = 0.5 \). Figure 3 shows the values of objective and constraint functions of time-average solutions. It can be seen from the plot of constraints that the staggered time averages converge faster than Algorithm 1. This illustrates the different between convergence times \( O(1/e^{1.5}) \) and \( O(1/e^2) \). Additional results of problems without the uniqueness assumption can be found in [22].

VII. CONCLUSION

We consider the time-average stochastic optimization problem with a non-convex decision set. The problem can be solved using the drift-plus-penalty algorithm, which has convergence time \( O(1/e^2) \). After we analyze the transient and steady state phases of the algorithm, the convergence time can be improved by performing time average in the steady state. We prove that the improved convergence time is \( O(1/e) \) under the locally-polyhedral assumption and is \( O(1/e^{1.5}) \) under the locally-non-polyhedral assumption.

REFERENCES


