ASYNCHRONOUS OPTIMIZATION OVER WEAKLY COUPLED RENEWAL SYSTEMS

By Xiaohan Wei[*] and Michael J. Neely[*]

University of Southern California
E-mail: xiaohanw@usc.edu; mikejneely@gmail.com

This paper considers optimization over multiple renewal systems coupled by time average constraints. These systems act asynchronously over variable length frames. When a particular system starts a new renewal frame, it chooses an action from a set of options for that frame. The action determines the duration of the frame, the penalty incurred during the frame (such as energy expenditure), and a vector of performance metrics (such as instantaneous number of job services). The goal is to minimize the time average penalty subject to time average overall constraints on the corresponding metrics. This problem has applications to task processing networks and coupled Markov decision processes (MDPs). We propose a new algorithm so that each system can make its own decision after observing a global multiplier that is updated every slot. We show that this algorithm satisfies the desired constraints and achieves $O(\epsilon)$ near optimality with $O(1/\epsilon^2)$ convergence time.

1. Introduction. Consider N renewal systems that operate over a slotted timeline ($t \in \{0, 1, 2, \ldots\}$). The timeline for each system $n \in \{1, \ldots, N\}$ is segmented into back-to-back intervals of time slots called *renewal frames*. The start of each renewal frame for a particular system is called a *renewal time* or simply a *renewal* for that system. The duration of each renewal frame is a random positive integer with distribution that depends on a control action chosen by the system at the start of the frame. The decision at each renewal frame also determines the penalty and a vector of performance metrics during this frame. The systems are coupled by time average constraints placed on these metrics over all systems. The goal is to design a decision strategy for each system so that overall time average penalty is minimized subject to time average constraints.

We use $k = 0, 1, 2, \cdots$ to index the renewals. Let t_k^n be the time slot corresponding to the k-th renewal of the n-th system with the convention that $t_0^n = 0$. Let $\mathcal{T}_k^n$ be the set of all slots from t_k^n to $t_{k+1}^n - 1$. At time t_k^n, the n-th system chooses a possibly random decision α_k^n in a set $\mathcal{A}^n$. This action determines the distributions of the following random variables:

- The duration of the k-th renewal frame $T_k^n := t_{k+1}^n - t_k^n$, which is a positive integer.
- A vector of performance metrics at each slot of that frame $\mathbf{z}^n[t] := (z_1^n[t], \ z_2^n[t], \ \cdots, \ z_L^n[t])$, $t \in \mathcal{T}_k^n$.
- A penalty incurred at each slot of the frame $y^n[t]$, $t \in \mathcal{T}_k^n$.

We assume each system has the *renewal property* that given $\alpha_k^n = \alpha^n \in \mathcal{A}^n$, the random variables T_k^n, $\mathbf{z}^n[t]$ and $y^n[t]$, $t \in \mathcal{T}_k^n$ are independent of the information of all systems from the slots before t_k^n with the following *known* conditional expectations $\mathbb{E}(T_k^n | \alpha_k^n = \alpha^n)$, $\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} y^n[t] \,\middle|\, \alpha_k^n = \alpha^n\right)$

---

and $\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} \mathbf{z}^n[t] \,\middle|\, \alpha_k^n = \alpha^n\right)$.

In addition, we have an uncontrollable external i.i.d. random process $\{\mathbf{d}[t]\}_{t=0}^{\infty} \subseteq \mathbb{R}^L$ which can be observed during each time slot. Let $d_l := \mathbb{E}(d_l[t])$. The goal is to minimize the total time average penalty of these N renewal systems subject to L total time average constraints on the performance metrics related to the external i.i.d. process, i.e. we aim to solve the following optimization problem:

$$(1) \qquad\qquad \min \quad \limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} \mathbb{E}(y^n[t])$$

$$(2) \qquad\qquad \text{s.t.} \quad \limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} \mathbb{E}(z_l^n[t]) \le d_l, \quad l \in \{1, 2, \cdots, L\}.$$

This problem is challenging because these N systems are weakly coupled by the time average constraints (2), yet each of them operates over its own renewal frames. The renewals of different systems do not have to be synchronized and they do not have to occur at the same rate. Fig. 1 plots a sample timeline of three parallel renewal systems. In Section 2, we will develop an algorithm that does not need the knowledge of $d_l = \mathbb{E}(d_l[t])$ with a provable performance guarantee.
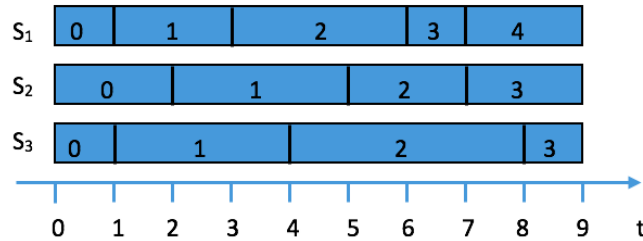

FIG 1. *The sample timelines of three asynchronous parallel renewal systems, where the numbers underneath the figure index time slots and the numbers inside the blocks index the renewals of each system.*

1.1. *Example Applications.*

1.1.1. *Multi-server energy-aware scheduling.* Consider a slotted time system with L classes of jobs and N servers. Job arrivals are Poisson distributed with rates $\lambda_1, \cdots, \lambda_L$, respectively. These jobs are stored in separate queues denoted as $Q_1[t], \cdots, Q_L[t]$ in a router waiting to be served. Assume the system is empty at time $t = 0$ so that $Q_l[0] = 0, \forall l \in \{1, 2, \cdots, L\}$. Let $\lambda_l[t]$ be the precise number of class l job arrivals at slot t, then, we have $\mathbb{E}(\lambda_l[t]) = \lambda_l, \forall l \in \{1, 2, \cdots, L\}$. Let $\mu_l^n[t]$ and $e^n[t]$ be the number of class l jobs served and the energy consumption for server n at time slot t, respectively. Fig. 2 sketches an example architecture of the system with 3 classes of jobs and 10 servers.

Each server makes decisions over renewal frames and the first frame starts at time slot $t = 0$. Successive renewals can happen at different slots for different servers. For the n-th server, at the beginning of the k-th frame ($k \in \mathbb{N}$), it chooses a processing mode m_k^n within the set of all modes $\mathcal{M}^n$. The processing mode m_k^n determines distributions on the number of jobs served, the service time, and the energy expenditure, with conditional expectations:

- $\widehat{T}^n(m_k^n) := \mathbb{E}(T_k^n|\ m_k^n)$. The expected frame size.
- $\widehat{\mu}_l^n(m_k^n) = \mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} \mu_l^n[t]\Big|\ m_k^n\right)$. The expected number of class l jobs served.
- $\widehat{e}^n(m_k^n) = \mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} e^n[t]\Big|\ m_k^n\right)$. The expected energy consumption.

The goal is to minimize the time average energy consumption, subject to the queue stability constraints, i.e.

$$(3) \qquad \min \quad \limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} \mathbb{E}(e^n[t])$$

$$(4) \qquad s.t. \quad \liminf_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} \mathbb{E}(\mu_l^n[t]) \geq \lambda_l, \ \forall l \in \{1, 2, \cdots, L\}.$$

Thus, we have formulated the problem into the form (1)-(2). Note that the external process in this example is the arrival process of L classes of jobs with potentially unknown arrival rates λ_l.

Previously, [12] treats a special case of this problem where all energy and service quantities are deterministic functions of the processing modes. The newly developed algorithm in the current paper can be used to solve this problem with considerably more general stochastic assumptions.
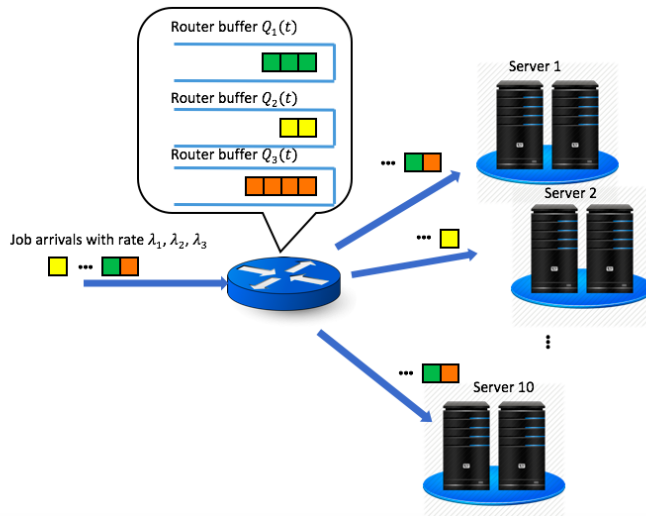

FIG 2. *Illustration of an energy-aware scheduling system with 3 classes of jobs and 10 parallel servers.*

1.1.2. *Coupled ergodic MDPs.* Consider N discrete time Markov decision processes (MDPs) over an infinite horizon. Each MDP consists of a finite state space $\mathcal{S}^n$, and an action space $\mathcal{U}^n$ at each state $s \in \mathcal{S}^n$.[1] For each state $s \in \mathcal{S}$, we use $P_u^n(s, s')$ to denote the transition probability from $s \in \mathcal{S}^n$ to $s' \in \mathcal{S}^n$ when taking action $u \in \mathcal{U}^n$, i.e.

$$P_u^n(s, s') = Pr(s[t+1] = s' \mid s[t] = s, \ u[t] = u),$$

where $s[t]$ and $u[t]$ are state and action at time slot t.

---

[1]To simplify the notation, we assume each state has the same action space $\mathcal{A}^n$. All our analysis generalizes trivially to states with different action spaces.

At time slot t, after observing the state $s[t] \in \mathcal{S}^n$ and choosing the action $u[t] \in \mathcal{U}^n$, the n-th MDP receives a penalty $y^n(u[t], s[t])$ and L types of resource costs $z_1^n(u[t], s[t]), \cdots, z_L^n(u[t], s[t])$, where these functions are all bounded mappings from $\mathcal{S}^n \times \mathcal{U}^n$ to $\mathbb{R}$. For simplicity we write $y^n[t] = y^n(u[t], s[t])$ and $z_l^n[t] = z_l^n(u[t], s[t])$. The goal is to minimize the time average overall penalty with constraints on time average overall costs, where these MDPs are weakly coupled through the time average constraints. This problem can be written in the form (1)-(2).

In order to define the renewal frame, we need one more assumption on the MDPs. We assume each of the MDPs is *ergodic*, i.e. there exists a state which is recurrent and the corresponding Markov chain is aperiodic under any randomized stationary policy[2], with bounded expected recurrence time. Under this assumption, the renewals for each MDP can be defined as successive revisitations to the recurrent state, and the action set $\mathcal{A}^n$ in such scenario is defined as the set of all randomized stationary policies that can be implemented in one renewal frame. Thus, our formulation includes coupled ergodic MDPs. We refer to [1], [3], and [17] for more details on MDP theory and related topics.

As a side remark, this multi-MDP problem can be viewed as a single MDP on an enlarged state space. Constrained MDPs are discussed previously in [1]. One can show that under the previous ergodic assumption, the minimum of (1)-(2) is achieved by a randomized stationary policy, and furthermore, such a policy can be obtained via solving a linear program reformulated from (1)-(2) offline. However, formulating such LP requires the knowledge of all the parameters in the problem, including the statistics of the external process $\{\mathbf{d}[t]\}_{t=0}^{\infty}$, and the resulting LP is often computationally intractable when the number of MDPs is very large. On the contrary, our newly developed algorithm is carried out in an online manner, does't require the statistics of the external process and enjoys a natural "decoupled" structure, effectively reducing the computational load.

1.2. *Challenges and previous approaches.*
As mentioned above, for the special case of coupled ergodic MDPs, this problem can be solved via a linear program (see [1] and also [8] for detailed discussions on formulating MDPs as linear programs). However, this approach becomes intractable as the number of MDPs gets very large. On the other hand, existing asynchronous algorithms and analysis (e.g. [4][6][15][19]) treat only the case where system delays (frames) are of fixed distribution independent of the actions or even deterministic, which are not readily extendable to our problem.

The main technical challenge is the dilemma on how to pick a correct time scale to carry out an algorithm and corresponding analysis. On one hand, since time is slotted, one would naturally think of synchronizing all systems on the slot scale and designing a slot-based algorithm. However, since each renewal spans multiple slots, any such algorithm would essentially cut some renewals in the middle and it would be difficult to analyze any particular system. On the other hand, if we analyze each system over its own renewal frames, it is not clear how to piece together these individual analyses.

Prior approaches treat this challenge only in special cases. The works [12] and [13] consider a special case where all quantities introduced above are deterministic functions of the actions. The work [20] considers another special case of the current formulation in server scheduling, where there is only one queue stability constraint and it can be easily met via controlling the arrival rate to the system. These two methods circumvent the aforementioned dilemma by making extra assumptions on the system and thus can not be generalized to the current setting.

---

[2]A *randomized stationary policy* π is an algorithm which chooses actions at state $s \in \mathcal{S}^n$ according to a fixed conditional distribution $\pi(u|s)$, $u \in \mathcal{U}^n$ and is independent of all other past information, i.e. $Pr(u[t]|\mathcal{F}_t) = \pi(u[t]|s[t])$, $u[t] \in \mathcal{U}^n$, $s[t] \in \mathcal{S}^n$ and $\mathcal{F}_t$ is the past information up to time t.

1.3. *Our contributions.* The current paper develops a new algorithm where each system operates on its own renewal frame. It is fully analyzed with convergence as well as convergence time results. As a first technical contribution, we fully characterize the fundamental performance region of the problem (1)-(2) (Lemma 3.2). We then resolve the aforementioned dilemma by constructing a supermartingale along with a stopping-time to "synchronize" all systems on a slot basis, by which we could piece together analysis of each individual system to prove the convergence of the proposed algorithm. Furthermore, encapsulating this new idea into convex analysis tools, we prove the $\mathcal{O}(1/\varepsilon^2)$ convergence time of the proposed algorithm to reach $\mathcal{O}(\varepsilon)$ near optimality under a mild assumption on the existence of a Lagrange multiplier (Section 4). Specifically, we show that for any accuracy $\epsilon > 0$ and any time $T \geq 1/\varepsilon^2$, the sequence $\{y^n[t]\}$ and $\{\mathbf{z}^n[t]\}$ produced by our algorithm satisfies,

$$\frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} \mathbb{E}(y^n[t]) \leq f_* + \mathcal{O}(\varepsilon),$$

$$\frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} \mathbb{E}(z_l^n[t]) \leq d_l + \mathcal{O}(\varepsilon), l \in \{1, 2, \cdots, L\},$$

where f_* denotes the optimal objective value of (1)-(2). Simulation experiments on the aforementioned multi-server energy-aware scheduling problem also demonstrate the effectiveness of the proposed algorithm.

1.4. *Other related works.* The problem considered in the current paper is a generalization of optimization over a single renewal system. It is shown in [14] that for the single renewal system with finite action set, the problem can be solved (offline) via a linear fractional program. Methods for solving linear fractional programs can be found in [5] and [18]. The *drift-plus-penalty ratio* approach is also developed in [11] and [14] for the single renewal system.

On the other hand, our problem is also related to the multi-server scheduling as is shown in one of the example applications. When assuming proper statistics of the arrivals and/or services, energy optimization problems in multi-server systems can also be treated via queueing theory. Specifically, by assuming both arrivals and services are Poisson distributed, [9] treats the multi-server system as an M/M/k/setup queue and explicitly computes several performance metrics via the renewal reward theorem. By assuming arrivals are Poisson and only one server, [10] and [21] treat the system as a multi-class M/G/1 queue and optimize the average energy consumption via polymatroid optimization.

1.5. *Notation and organization of the paper.* Throughout the paper, we use superscript $n \in \{1, 2, \cdots, N\}$ to index different systems, use the subscript $l \in \{1, 2, \cdots, L\}$ to index different constraints and use the subscript $k \in \mathbb{N}$ to index the frames. For any vector $\mathbf{x} \in \mathbb{R}^d$, the considered norms are $\|\mathbf{x}\| := \sqrt{\sum_{i=1}^{d} x_i^2}$, $\|\mathbf{x}\|_1 := \sum_{i=1}^{d} |x_i|$ and $\|\mathbf{x}\|_\infty := \max_i |x_i|$.

The rest of the paper is organized as follows: Section 2 introduces the proposed algorithm along with technical assumptions. Section 3 introduces our main technical argument proving the convergence of the proposed algorithm via supermartingale and stopping time constructions. Building upon these technical tools, Section 4 takes one step further and proves the convergence time of the proposed algorithm. Finally, a simulation study regarding multi-server energy-aware scheduling is given in Section 5.

2. Algorithm.

2.1. *Technical preliminaries.* Throughout the paper, we make the following assumptions.

ASSUMPTION 2.1. *The problem* (1)-(2) *is feasible, i.e. there are action sequences* $\{\alpha_k^n\}_{k=0}^{\infty}$ *for all* $n \in \{1, 2, \cdots, N\}$ *so that the corresponding process* $\{\mathbf{z}^n[t]\}_{t=0}^{\infty}$ *satisfies the constraints* (2).

Following this assumption, we define f_* as the infimum objective value for (1)-(2) over all decision sequences that satisfy the constraints.

ASSUMPTION 2.2 (Boundedness). *For any* $k \in \mathbb{N}$ *and any* $n \in \{1, 2, \cdots, N\}$, *there exist absolute constants* $y_{\max}$, $z_{\max}$ *and* $d_{\max}$ *such that*

$$|y^n[t]| \le y_{\max}, \quad |z_l^n[t]| \le z_{\max}, \quad |d_l[t]| \le d_{\max}, \quad \forall t \in \mathcal{T}_k^n, \ \forall l \in \{1, 2, \cdots, L\}.$$

Furthermore, there exists an absolute constant $B \ge 1$ *such that for every fixed* $\alpha^n \in \mathcal{A}^n$ *and every* $s \in \mathbb{N}$ *for which* $Pr(T_k^n \ge s | \alpha_k^n = \alpha^n) > 0$,

$$(5) \qquad\qquad\qquad \mathbb{E}\big((T_k^n - s)^2 \big| \ \alpha_k^n = \alpha^n, T_k^n \ge s\big) \le B.$$

REMARK 2.1. *The quantity* $T_k^n - s$ *is usually referred to as the residual lifetime. In the special case where* $s = 0$, (5) *gives the uniform second moment bound of the renewal frames as*

$$\mathbb{E}\big((T_k^n)^2 \big| \ \alpha_k^n = \alpha^n\big) \le B.$$

Note that (5) *is satisfied for a large class of problems. In particular, it can be shown to hold in the following three cases:*

  1. *If the inter-renewal* T_k^n *is deterministically bounded.*
  2. *If the inter-renewal* T_k^n *is geometrically distributed.*
  3. *If each system is a finite state ergodic MDP with a finite action set.*

DEFINITION 2.1. *For any* $\alpha^n \in \mathcal{A}^n$, *let*

$$\widehat{y}^n(\alpha^n) := \mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} y^n[t] \middle| \alpha_k^n = \alpha^n\right), \quad \widehat{z}_l^n(\alpha^n) := \mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} z_l^n[t] \middle| \alpha_k^n = \alpha^n\right),$$

and $\widehat{T}^n(\alpha^n) := \mathbb{E}(T_k^n | \alpha_k^n = \alpha^n)$. *Define*

$$\widehat{f}^n(\alpha^n) := \widehat{y}^n(\alpha^n)/\widehat{T}^n(\alpha^n),$$
$$\widehat{g}_l^n(\alpha^n) := \widehat{z}_l^n(\alpha^n)/\widehat{T}^n(\alpha^n), \ \forall l \in \{1, 2, \cdots, L\},$$

and let $\left(\widehat{f}^n(\alpha^n), \ \widehat{\mathbf{g}}^n(\alpha^n)\right)$ *be a performance vector under the action* α^n.

Note that by Assumption 2.2, $\widehat{y}^n(\alpha^n)$ and $\widehat{\mathbf{z}}^n(\alpha^n)$ in Definition 2.1 are both bounded, and $T_k^n \ge 1$, $\forall k \in \mathbb{N}$, thus, the set $\left\{\left(\widehat{f}^n(\alpha^n), \ \widehat{\mathbf{g}}^n(\alpha^n)\right), \ \alpha^n \in \mathcal{A}^n\right\}$ is also bounded. The following mild assumption states that this set is also closed.

ASSUMPTION 2.3. *The set* $\left\{\left(\widehat{f}^n(\alpha^n), \ \widehat{\mathbf{g}}^n(\alpha^n)\right), \ \alpha^n \in \mathcal{A}^n\right\}$ *is compact.*

Finally, we define the performance region of each individual system as follows.

DEFINITION 2.2. *Let $\mathcal{S}^n$ be the convex hull of $\left\{ \left(\widehat{y}^n(\alpha^n), \ \widehat{\mathbf{z}}^n(\alpha^n), \ \widehat{T}^n(\alpha^n) \right) : \ \alpha^n \in \mathcal{A}^n \right\} \subseteq$ $\mathbb{R}^{L+2}$. Define*

$$\mathcal{P}^n := \{(y/T, \ \mathbf{z}/T) : \ (y, \mathbf{z}, T) \in \mathcal{S}^n\} \subseteq \mathbb{R}^{L+1}$$

as the performance region of system n.

2.2. *Proposed algorithm.* In this section, we propose an algorithm where each system can make its own decision after observing a global vector of multipliers which is updated using the global information from all systems. We start by defining a vector of virtual queues $\mathbf{Q}[t] := (Q_1[t], \ Q_2[t], \ \cdots, \ Q_L[t])$, which are 0 at $t = 0$ and updated as follows,

$$(6) \qquad Q_l[t+1] = \max \left\{ Q_l[t] + \sum_{n=1}^{N} z_l^n[t] - d_l[t], \ 0 \right\}, \ l \in \{1, 2, \cdots, L\}.$$

These virtual queues will serve as global multipliers to control the growth of corresponding resource consumptions.

Then, the proposed algorithm runs as follows via a fixed trade-off parameter $V > 0$:

- At the beginning of k-th frame of system n, the system observes the vector of virtual queues $\mathbf{Q}[t_k^n]$ and makes a decision $\alpha_k^n \in \mathcal{A}^n$ so as to solve the following subproblem:

$$(7) \qquad D_k^n := \min_{\alpha^n \in \mathcal{A}^n} \frac{\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} \left(V y^n[t] + \langle \mathbf{Q}[t_k^n], \mathbf{z}^n[t] \rangle \right) \Big| \alpha_k^n = \alpha^n, \mathbf{Q}[t_k^n] \right)}{\mathbb{E}\left(T_k^n \Big| \alpha_k^n = \alpha^n, \mathbf{Q}[t_k^n] \right)}.$$

- Update the virtual queue after each slot:

$$Q_l[t+1] = \max \left\{ Q_l[t] + \sum_{n=1}^{N} z_l^n[t] - d_l[t], \ 0 \right\}, \ l \in \{1, 2, \cdots, L\}.$$

Note that using the notation specified in Definition 2.1, we can rewrite (7) in a more concise way as follows:

$$(8) \qquad \min_{\alpha^n \in \mathcal{A}^n} \left\{ V \widehat{f}^n(\alpha^n) + \langle \mathbf{Q}[t_k^n], \widehat{\mathbf{g}}^n(\alpha^n) \rangle \right\},$$

which is a deterministic optimization problem. Then, by the compactness assumption (Assumption 2.3), there always exists a solution to this subproblem.

This algorithm requires knowledge of the conditional expectations associated with the performance vectors $\left(\widehat{f}^n(\alpha^n), \ \widehat{\mathbf{g}}^n(\alpha^n) \right)$, $\alpha^n \in \mathcal{A}^n$, but only requires individual systems n to know their own $\left(\widehat{f}^n(\alpha^n), \ \widehat{\mathbf{g}}^n(\alpha^n) \right)$, $\alpha^n \in \mathcal{A}^n$, and therefore decouples these systems. Furthermore, the virtual queue update uses observed $d_l[t]$ and does not require knowledge of distribution or mean of $d_l[t]$.

In addition, we introduce $\mathbf{Q}[t]$ as "virtual queues" for the following two reasons: First, it can be mapped to real queues in applications (such as the server scheduling problem mentioned in Section 1.1.1), where $\mathbf{d}[t]$ stands for the arrival process and $\mathbf{z}[t]$ is the service process. Second, stabilizing these virtual queues implies the constraints (2) are satisfied, as is illustrated in the following lemma, whose proof is given in the appendix.

LEMMA 2.1. *If $Q_l[0] = 0$ and $\lim_{T \to \infty} \frac{1}{T} \mathbb{E}(Q_l[T]) = 0$, then, $\limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} \mathbb{E}(z_l^n[t]) \leq d_l$.*

2.3. *Computing subproblems.* Since a key step in the algorithm is to solve the optimization problem (8), we make several comments on the computation of the ratio minimization (8). In general, one can solve the ratio optimization problem (7) (therefore (8)) via a bisection search algorithm. For more details, see section 7 of [11]. However, more often than not, bisection search is not the most efficient one. We will discuss two special cases arising from applications where we can find a simpler way of solving the subproblem.

First of all, when there are only a finite number of actions in the set $\mathcal{A}^n$, one can solve (8) simply via enumerating. This is a typical scenario in energy-aware scheduling where a finite action set consists of different processing modes that can be chosen by servers.

Second, when the set $\left\{\left(\widehat{y}^n(\alpha^n),\ \widehat{\mathbf{z}}^n(\alpha^n),\ \widehat{T}^n(\alpha^n)\right):\ \alpha^n \in \mathcal{A}^n\right\}$ specified in Definition 2.2 is itself a convex hull of a finite sequence $\{(y_j, \mathbf{z}_j, T_j)\}_{j=1}^m$, then, (8) can be rewritten as a simple enumeration:

$$\min_{i \in \{1,2,\cdots,m\}}\ \left\{V \frac{y_i}{T_i} + \left\langle \mathbf{Q}[t_k^n], \frac{\mathbf{z}_i}{T_i} \right\rangle\right\}.$$

To see this, note that by definition of convex hull, for any $\alpha^n \in \mathcal{A}^n$, $\left(\widehat{y}^n(\alpha^n),\ \widehat{\mathbf{z}}^n(\alpha^n),\ \widehat{T}^n(\alpha^n)\right) = \sum_{j=1}^m p_j \cdot (y_j, z_j, T_j)$ for some $\{p_j\}_{j=1}^m$, $p_j \geq 0$ and $\sum_{j=1}^m p_j = 1$. Thus,

$$\begin{aligned}
V\widehat{f}^n(\alpha^n) + \langle \mathbf{Q}[t_k^n], \widehat{\mathbf{g}}^n(\alpha^n)\rangle =& V \frac{\sum_{j=1}^m p_j y_j}{\sum_{j=1}^m p_j T_j} + \left\langle \mathbf{Q}[t_k^n], \frac{\sum_{j=1}^m p_j \mathbf{z}_j}{\sum_{j=1}^m p_j T_j} \right\rangle \\
=& \sum_{i=1}^m \frac{p_i T_i}{\sum_{j=1}^m p_j T_j} \left(V \frac{y_i}{T_i} + \left\langle \mathbf{Q}[t_k^n], \frac{\mathbf{z}_i}{T_i} \right\rangle\right) \\
=:& \sum_{i=1}^m q_i \left(V \frac{y_i}{T_i} + \left\langle \mathbf{Q}[t_k^n], \frac{\mathbf{z}_i}{T_i} \right\rangle\right),
\end{aligned}$$

where we let $q_i = \frac{p_i T_i}{\sum_{j=1}^m p_j T_j}$. Note that $q_i \geq 0$ and $\sum_{i=1}^m q_i = 1$ because $T_i \geq 1$. Hence, solving (8) is equivalent to choosing $\{q_i\}_{i=1}^m$ to minimize the above expression, which boils down to choosing a single $(y_i, \mathbf{z}_i, T_i)$ among $\{(y_j, \mathbf{z}_j, T_j)\}_{j=1}^m$ which achieves the minimum.

Note that such a convex hull case stands out not only because it yields a simple solution, but also because of the fact that ergodic coupled MDPs discussed in Section 1.1.2 have the region $\left\{\left(\widehat{y}^n(\alpha^n),\ \widehat{\mathbf{z}}^n(\alpha^n),\ \widehat{T}^n(\alpha^n)\right):\ \alpha^n \in \mathcal{A}^n\right\}$ being the convex hull of a finite sequence of points $\{(y_j, \mathbf{z}_j, T_j)\}_{j=1}^m$, where each point $(y_j, \mathbf{z}_j, T_j)$ results from a pure stationary policy ([1]). [3] Thus, solving (8) for the ergodic coupled MDPs reduces to choosing a pure policy among a finite number of pure policies.

3. Limiting Performance. For the rest of the paper, the underlying probability space is denoted as the tuple $(\Omega,\ \mathcal{F},\ P)$. Let $\mathcal{F}[t]$ be the system history up until time slot t. Formally, $\{\mathcal{F}[t]\}_{t=0}^\infty$ is a filtration with $\mathcal{F}[0] = \{\emptyset, \Omega\}$ and each $\mathcal{F}[t]$, $t \geq 1$ is the σ-algebra generated by all random variables from slot 0 to $t-1$.

For the rest of the paper, we always assume Assumptions 2.1-2.3 hold without explicitly mentioning them.

3.1. *Convexity.* The following lemma demonstrates the convexity of $\mathcal{P}^n$ in Definition 2.2.

---

[3]A pure stationary policy is an algorithm where the decision to be taken at any time t is a deterministic function of the state at time t, and independent of all other past information.

LEMMA 3.1. *The performance region $\mathcal{P}^n$ specified in Definition 2.2 is convex for any $n \in \{1, 2, \cdots, N\}$. Furthermore, it is the convex hull of the set $\left\{ \left(\widehat{f}^n(\alpha^n), \ \widehat{\mathbf{g}}^n(\alpha^n) \right) : \alpha^n \in \mathcal{A}^n \right\}$ and thus compact, where $\left(\widehat{f}^n(\alpha^n), \ \widehat{\mathbf{g}}^n(\alpha^n) \right)$ is specified Definition 2.1.*

PROOF. We first prove the convexity of $\mathcal{P}^n$. Consider any two points $(f_1, \mathbf{g}_1), \ (f_2, \mathbf{g}_2) \in \mathcal{P}^n$. We aim to show that for any $q \in (0, 1)$, $(qf_1 + (1-q)f_2, q\mathbf{g}_1 + (1-q)\mathbf{g}_2) \in \mathcal{P}^n$. Notice that by definition of $\mathcal{P}^n$, there exists $(y_1, \mathbf{z}_1, T_1), \ (y_2, \mathbf{z}_2, T_2) \in \mathcal{S}^n$ such that $f_1 = y_1/T_1$, $\mathbf{g}_1 = \mathbf{z}_1/T_1$, $f_2 = y_2/T_2$, and $\mathbf{g}_2 = \mathbf{z}_2/T_2$. Thus, it is enough to show

$$(9) \qquad \left(q\frac{y_1}{T_1} + (1-q)\frac{y_2}{T_2}, q\frac{\mathbf{z}_1}{T_1} + (1-q)\frac{\mathbf{z}_2}{T_2} \right) \in \mathcal{P}^n.$$

To show this, we make a change of variable by letting $p = \frac{qT_2}{(1-q)T_1 + qT_2}$. It is obvious that $p \in (0, 1)$. Furthermore, $q = \frac{pT_1}{pT_1 + (1-p)T_2}$ and

$$q\frac{y_1}{T_1} + (1-q)\frac{y_2}{T_2} = \frac{py_1 + (1-p)y_2}{pT_1 + (1-p)T_2},$$
$$q\frac{\mathbf{z}_1}{T_1} + (1-q)\frac{\mathbf{z}_2}{T_2} = \frac{p\mathbf{z}_1 + (1-p)\mathbf{z}_2}{pT_1 + (1-p)T_2}.$$

Since $\mathcal{S}^n$ is convex,

$$(py_1 + (1-p)y_2, \ p\mathbf{z}_1 + (1-p)\mathbf{z}_2, \ pT_1 + (1-p)T_2) \in \mathcal{S}^n.$$

Thus, by definition of $\mathcal{P}^n$ again, (9) holds and the first part of the proof is finished.

To show the second part of the claim, let

$$\mathcal{Q}^n := \left\{ \left(\widehat{f}^n(\alpha^n), \ \widehat{\mathbf{g}}^n(\alpha^n) \right) : \alpha^n \in \mathcal{A}^n \right\} = \left\{ \left(\widehat{y}^n(\alpha^n) \Big/ \widehat{T}^n(\alpha^n) \, , \ \widehat{\mathbf{z}}^n(\alpha^n) \Big/ \widehat{T}^n(\alpha^n) \right) : \alpha^n \in \mathcal{A}^n \right\}$$

and let $\text{conv}(\mathcal{Q}^n)$ be the convex hull of $\mathcal{Q}^n$. First of all, By Definition 2.2,

$$\mathcal{P}^n = \{(y/T, \ \mathbf{z}/T) : \ (y, \mathbf{z}, T) \in \mathcal{S}^n\} \subseteq \mathbb{R}^{L+1},$$

for $\mathcal{S}^n$ being the convex hull of $\left\{ \left(\widehat{y}^n(\alpha^n), \ \widehat{\mathbf{z}}^n(\alpha^n), \ \widehat{T}^n(\alpha^n) \right) : \ \alpha^n \in \mathcal{A}^n \right\}$, thus, in view of the definition of $\mathcal{Q}^n$, we have $\mathcal{Q}^n \subseteq \mathcal{P}^n$. Since both $\mathcal{P}^n$ and $\text{conv}(\mathcal{Q}^n)$ are convex, by definition of convex hull ([16]) that $\text{conv}(\mathcal{Q}^n)$ is the smallest convex set containing $\mathcal{Q}^n$, we have $\text{conv}(\mathcal{Q}^n) \subseteq \mathcal{P}^n$.

To show the reverse inclusion $\mathcal{P}^n \subseteq \text{conv}(\mathcal{Q}^n)$, note that any point in $\mathcal{P}^n$ can be written in the form $\left(\frac{y}{T}, \frac{\mathbf{z}}{T} \right)$, where $(y, \mathbf{z}, T) \in \mathcal{S}^n$. Since $\mathcal{S}^n$ by definition is the convex hull of

$$\left\{ \left(\widehat{y}^n(\alpha^n), \ \widehat{\mathbf{z}}^n(\alpha^n), \ \widehat{T}^n(\alpha^n) \right) : \ \alpha^n \in \mathcal{A}^n \right\} \subseteq \mathbb{R}^{L+2},$$

by the definition of convex hull, $(y, \mathbf{z}, T)$ can be written as a convex combination of m points in the above set. Let $\left\{ \left(\widehat{y}^n(\alpha_i^n), \ \widehat{\mathbf{z}}^n(\alpha_i^n), \ \widehat{T}^n(\alpha_i^n) \right) \right\}_{i=1}^m$ be these points, so that

$$(y, \mathbf{z}, T) = \sum_{i=1}^m p_i \cdot \left(\widehat{y}^n(\alpha_i^n), \ \widehat{\mathbf{z}}^n(\alpha_i^n), \ \widehat{T}^n(\alpha_i^n) \right),$$
$$p_i \geq 0, \quad \sum_{i=1}^m p_i = 1.$$

As a result, we have

$$\left(\frac{y}{T}, \frac{\mathbf{z}}{T}\right) = \left(\frac{\sum_{i=1}^{m} p_i y^n(\alpha_i^n)}{\sum_{i=1}^{m} p_i T^n(\alpha_i^n)}, \frac{\sum_{i=1}^{m} p_i \mathbf{z}^n(\alpha_i^n)}{\sum_{i=1}^{m} p_i T^n(\alpha_i^n)}\right).$$

We make a change of variable by letting $q_j = \frac{p_j T^n(\alpha_j^n)}{\sum_{i=1}^{m} p_i T^n(\alpha_i^n)}$, $\forall j = 1, 2, \cdots, m$, then,

$$p_j = \frac{q_j}{T^n(\alpha_j^n)} \cdot \sum_{i=1}^{m} p_i T^n(\alpha_i^n),$$

it follows,

$$\left(\frac{y}{T}, \frac{\mathbf{z}}{T}\right) = \sum_{i=1}^{m} q_i \cdot \left(\frac{y^n(\alpha_i^n)}{T^n(\alpha_i^n)}, \frac{\mathbf{z}^n(\alpha_i^n)}{T^n(\alpha_i^n)}\right) = \sum_{i=1}^{m} q_i \cdot \left(\widehat{f}^n(\alpha_i^n), \widehat{\mathbf{g}}^n(\alpha_i^n)\right).$$

Since $\sum_{i=1}^{m} q_i = 1$ and $q_i \geq 0$, it follows any point in $\mathcal{P}^n$ can be written as a convex combination of finite number of points in $\mathcal{Q}^n$, which implies $\mathcal{P}^n \subseteq \text{conv}(\mathcal{Q}^n)$. Overall, we have $\mathcal{P}^n = \text{conv}(\mathcal{Q}^n)$.

Finally, by Assumption 2.3, we have $\mathcal{Q}^n = \left\{\left(\widehat{f}^n(\alpha^n), \widehat{\mathbf{g}}^n(\alpha^n)\right) : \alpha^n \in \mathcal{A}^n\right\}$ is compact. Thus, $\mathcal{P}^n$, being a convex hull of a compact set, is also compact. $\qquad\square$

3.2. *Key-feature inequality and supermartingale construction.* First of all, we have the following fundamental performance lemma which states that the optimality of (1)-(2) is achievable within $\mathcal{P}^n$ specified in Definition 2.2.

LEMMA 3.2.   *For each $n \in \{1, 2, \cdots, N\}$, there exists a pair $\left(\overline{f}_*^n, \overline{\mathbf{g}}_*^n\right) \in \mathcal{P}^n$ such that the following hold:*

$$\sum_{n=1}^{N} \overline{f}_*^n = f_*$$

$$\sum_{n=1}^{N} \overline{g}_{l,*}^n \leq d_l, \ l \in \{1, 2, \cdots, L\},$$

where f^* is the optimal objective value for problem (1)-(2), i.e. the optimality is achievable within $\otimes_{n=1}^{N} \mathcal{P}^n$, the Cartesian product of $\mathcal{P}^n$.

Furthermore, for any $\left(\overline{f}^n, \overline{\mathbf{g}}^n\right) \in \mathcal{P}^n$, $n \in \{1, 2, \cdots, N\}$, satisfying $\sum_{n=1}^{N} \overline{g}_l^n \leq d_l$, $l \in \{1, 2, \cdots, L\}$, we have $\sum_{n=1}^{N} \overline{f}^n \geq f_*$, i.e. one cannot achieve better performance than (1)-(2) in $\otimes_{n=1}^{N} \mathcal{P}^n$.

The proof of this Lemma is delayed to Appendix A. In particular, the proof uses the following lemma, which also plays an important role in several lemmas later.

LEMMA 3.3.   *Suppose $\{y^n[t]\}_{t=0}^{\infty}$, $\{\mathbf{z}^n[t]\}_{t=0}^{\infty}$ and $\{T_k^n\}_{k=0}^{\infty}$ are processes resulting from any algorithm,[4] then, $\forall T \in \mathbb{N}$,*

$$(10) \qquad\qquad \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}(f^n[t] - y^n[t]) \leq \frac{B_1}{T},$$

$$(11) \qquad\qquad \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}(g_l^n[t] - z_l^n[t]) \leq \frac{B_2}{T}, \ l \in \{1, 2, \cdots, L\},$$

---

[4]Note that this algorithm might make decisions using the past information.

where $B_1 = 2y_{\max}\sqrt{B}$, $B_2 = 2z_{\max}\sqrt{B}$ and $f^n[t]$, $\mathbf{g}^n[t]$ are constant over each renewal frame for system n defined by

$$f^n[t] = \widehat{f}^n(\alpha^n), \quad if\ t \in \mathcal{T}_k^n, \alpha_k^n = \alpha^n$$
$$\mathbf{g}^n[t] = \widehat{\mathbf{g}}^n(\alpha^n), \quad if\ t \in \mathcal{T}_k^n, \alpha_k^n = \alpha^n,$$

and $\left(\widehat{f}^n(\alpha^n), \widehat{\mathbf{g}}^n(\alpha^n)\right)$ are defined in Definition 2.1.

The proof of this lemma is delayed to Appendix A.

REMARK 3.1. *Note that directly computing $\overline{f}_*^n$ and $\overline{g}_{l,*}^n$ indicated by Lemma 3.2 would be difficult because of the fractional nature of $\mathcal{P}^n$, the coupling between different systems through time average constraints and the fact that $d_l = \mathbb{E}(d_l[t])$ might be unknown. However, Lemma 3.2 can be used to prove important performance theorems regarding our proposed algorithm as is indicated by the following lemma.*

The following key-feature inequality connects our proposed algorithm with the performance vectors inside $\mathcal{P}^n$.

LEMMA 3.4. *Consider the stochastic processes $\{y^n[t]\}_{t=0}^\infty$, $\{\mathbf{z}^n[t]\}_{t=0}^\infty$, and $\{T_k^n\}_{k=0}^\infty$ resulting from the proposed algorithm. For any system n, the following holds for any $k \in \mathbb{N}$ and any $(\overline{f}^n, \overline{\mathbf{g}}^n) \in \mathcal{P}^n$,*

$$(12) \qquad \frac{\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} (Vy^n[t] + \langle \mathbf{Q}[t_k^n], \mathbf{z}^n[t]\rangle) \middle| \mathbf{Q}[t_k^n]\right)}{\mathbb{E}\left(T_k^n | \mathbf{Q}[t_k^n]\right)} \le V\overline{f}^n + \langle \mathbf{Q}[t_k^n], \overline{\mathbf{g}}^n\rangle,$$

PROOF. First of all, since the proposed algorithm solves (7) over all possible decisions in $\mathcal{A}^n$, it must achieve value less than or equal to that of any action $\alpha^n \in \mathcal{A}^n$ at the same frame. This gives,

$$D_k^n \le \frac{\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} (Vy^n[t] + \langle \mathbf{Q}[t_k^n], \mathbf{z}^n[t]\rangle) \middle| \mathbf{Q}[t_k^n], \alpha_k^n = \alpha^n\right)}{\mathbb{E}\left(T_k^n \middle| \mathbf{Q}[t_k^n], \alpha_k^n = \alpha^n\right)} = \frac{V\widehat{y}^n(\alpha^n) + \langle \mathbf{Q}[t_k^n], \widehat{\mathbf{z}}^n(\alpha^n)\rangle}{\widehat{T}^n(\alpha^n)},$$

where D_k^n is defined in (7) and the equality follows from the renewal property of the system that T_k^n, $\sum_{t \in \mathcal{T}_k^n} y^n[t]$ and $\sum_{t \in \mathcal{T}_k^n} \mathbf{z}^n[t]$ are conditionally independent of $\mathbf{Q}[t_k^n]$ given $\alpha_k^n = \alpha^n$.

Since $\widehat{T}_k^n \ge 1$, this implies

$$\widehat{T}^n(\alpha^n) \cdot D_k^n \le V\widehat{y}^n(\alpha^n) + \langle \mathbf{Q}[t_k^n], \widehat{\mathbf{z}}^n(\alpha^n)\rangle,$$

thus, for any $\alpha^n \in \mathcal{A}^n$,

$$V\widehat{y}^n(\alpha^n) + \langle \mathbf{Q}[t_k^n], \widehat{\mathbf{z}}^n(\alpha^n)\rangle - D_k^n \cdot \widehat{T}^n(\alpha^n) \ge 0.$$

Since $\mathcal{S}^n$ specified in Definition 2.2 is the convex hull of $\left\{(\widehat{y}^n(\alpha^n), \widehat{\mathbf{z}}^n(\alpha^n), \widehat{T}^n(\alpha^n)), \alpha^n \in \mathcal{A}^n\right\}$, it follows for any vector $(y, \mathbf{z}, T) \in \mathcal{S}^n$, we have

$$Vy + \langle \mathbf{Q}[t_k^n], \mathbf{z}\rangle - D_k^n \cdot T \ge 0.$$

Dividing both sides by T and using the definition of $\mathcal{P}^n$ in Definition 2.2 give

$$D_k^n \le V\overline{f}^n + \langle \mathbf{Q}[t_k^n], \overline{\mathbf{g}}^n\rangle, \ \forall (\overline{f}^n, \overline{\mathbf{g}}^n) \in \mathcal{P}^n.$$

Finally, since $\{y^n[t]\}_{t=0}^\infty$, $\{\mathbf{z}^n[t]\}_{t=0}^\infty$, and $\{T_k^n\}_{k=0}^\infty$ result from the proposed algorithm and the action chosen is determined by $\mathbf{Q}[t_k^n]$ as in (7),

$$D_k^n = \frac{\mathbb{E}\Big(\sum_{t\in\mathcal{T}_k^n} \big(V y^n[t] + \langle \mathbf{Q}[t_k^n], \mathbf{z}^n[t]\rangle \big) \Big| \mathbf{Q}[t_k^n]\Big)}{\mathbb{E}\big(T_k^n | \mathbf{Q}[t_k^n]\big)}.$$

This finishes the proof.   □

Our next step is to give a frame-based analysis for each system by constructing a supermartingale on the per-frame timescale. Recall that $\{\mathcal{F}[t]\}_{t=0}^\infty$ is a filtration (with $\mathcal{F}[t]$ representing system history during slots $\{0, \cdots, t-1\}$). Fix a system n and recall that t_k^n is the time slot where the k-th renewal occurs for system n. We would like to define a filtration corresponding to the random times t_k^n. To this end, define the collection of sets $\{\mathcal{F}_k^n\}_{k=0}^\infty$ such that for each k,

$$\mathcal{F}_k^n := \{A \in \mathcal{F} : A \cap \{t_k^n \le t\} \in \mathcal{F}[t], \forall t \in \{0, 1, 2, \cdots\}\}$$

For example, the following set A is an element of $\mathcal{F}_3^n$:

$$A = \{t_3^n = 5\} \cap \{y[0] = y_0, y[1] = y_1, y[2] = y_2, y[3] = y_3, y[4] = y_4\}$$

where $y_0, \cdots, y_4$ are specific values. Then $A \in \mathcal{F}_3^n$ because for $i \in \{0, 1, 2, 3, 4\}$ we have $A \cap \{t_3^n \le i\} = \emptyset \in \mathcal{F}[i]$, and for $i \in \{5, 6, 7, \cdots\}$ we have $A \cap \{t \le i\} = A \in \mathcal{F}[i]$. The following technical lemma is proved in the appendix.

LEMMA 3.5.   *The sequence $\{\mathcal{F}_k^n\}_{k=0}^\infty$ is a valid filtration, i.e. $\mathcal{F}_k^n \subseteq \mathcal{F}_{k+1}^n$, $\forall k \ge 0$. Furthermore, for any real-valued adapted process $\{Z^n[t-1]\}_{t=1}^\infty$ with respect to $\{\mathcal{F}[t]\}_{t=1}^\infty$,* [5]

$$\big\{G_{t_k^n}(Z^n[0],\ Z^n[1],\ \cdots, Z^n[t_k^n-1])\big\}_{k=1}^\infty$$

is also adapted to $\{\mathcal{F}_k^n\}_{k=1}^\infty$, where for any $t \in \mathbb{N}$, $G_t(\cdot)$ is a fixed real-valued measurable mappings. That is, for any k, it holds that any measurable function of $(Z^n[0], \cdots, Z[t_k^n-1])$ is determined by events in $\mathcal{F}_k^n$.

With Lemma 3.4 and Lemma 3.5, we can construct a supermartingale as follows,

LEMMA 3.6.   *Consider the stochastic processes $\{y^n[t]\}_{t=0}^\infty$, $\{\mathbf{z}^n[t]\}_{t=0}^\infty$, and $\{T_k^n\}_{k=0}^\infty$ resulting from the proposed algorithm. For any $(\overline{f}^n, \overline{\mathbf{g}}^n) \in \mathcal{P}^n$, let*

(13) $$X^n[t] := V\left(y^n[t] - \overline{f}^n\right) + \langle \mathbf{Q}[t], \mathbf{z}^n[t] - \overline{\mathbf{g}}^n\rangle,$$

then,

$$\mathbb{E}\left(\sum_{t\in\mathcal{T}_k^n} X^n[t] \bigg| \mathcal{F}_k^n\right) \le L z_{\max}(N z_{\max} + d_{\max})B := C_0,$$

where B, $z_{\max}$ and $d_{\max}$ are as defined in Assumption 2.2. Furthermore, define a real-valued process $\{Y_K^n\}_{K=0}^\infty$ on the frame such that $Y_0^n = 0$ and

$$Y_K^n = \sum_{k=0}^{K-1}\left(\sum_{t\in\mathcal{T}_k^n} X^n[t] - C_0\right),\ K \ge 1.$$

Then, $\{Y_K^n\}_{K=0}^\infty$ is a supermartingale adapted to the aforementioned filtration $\{\mathcal{F}_k^n\}_{K=0}^\infty$.

---

[5]Meaning that for each t in $\{1, 2, 3, \cdots\}$, the random variable $Z^n[t-1]$ is determined by events in $\mathcal{F}[t]$.

PROOF. Consider any $t \in \mathcal{T}_k^n$, then, we can decompose $X^n[t]$ as follows

$$(14) \qquad X^n[t] = V(y^n[t] - \overline{f}^n) + \langle \mathbf{Q}[t_k^n], \mathbf{z}^n[t] - \overline{\mathbf{g}}^n \rangle + \langle \mathbf{Q}[t] - \mathbf{Q}[t_k^n], \mathbf{z}^n[t] - \overline{\mathbf{g}}^n \rangle.$$

By the queue updating rule (6), we have for any $l \in \{1, 2, \cdots, L\}$ and any $t > t_k^n$,

$$(15) \qquad |Q_l[t] - Q_l[t_k^n]| \le \sum_{s=t_k^n}^{t-1} \left| \sum_{m=1}^{N} z_l^m[s] - d_l[t] \right| \le (t - t_k^n)(N z_{\max} + d_{\max})$$

Thus, for the last term in (14), by Hölder's inequality,

$$\begin{aligned}
\langle \mathbf{Q}[t] - \mathbf{Q}[t_k^n], \mathbf{z}^n[t] - \overline{\mathbf{g}}^n \rangle &\le \|\mathbf{Q}[t] - \mathbf{Q}[t_k^n]\|_1 \cdot \|\mathbf{z}^n[t] - \overline{\mathbf{g}}^n\|_\infty \\
&\le \sum_{s=t_k^n}^{t-1} \left\| \sum_{m=1}^{N} \mathbf{z}^n[s] - \mathbf{d}[t] \right\|_1 \cdot \|\mathbf{z}^n[t] - \overline{\mathbf{g}}^n\|_\infty \\
&\le (t - t_k^n) L (N z_{\max} + d_{\max}) \cdot 2 z_{\max},
\end{aligned}$$

where the second inequality follows from (15) and the last inequality follows from the boundedness assumption (Assumption 2.2) of corresponding quantities. Substituting the above bound into (14) gives a bound on $\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} X^n[t] \Big| \mathcal{F}_k^n \right)$ as

$$\begin{aligned}
\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} X^n[t] \Big| \mathcal{F}_k^n \right) &\le \mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} \left(V\left(y^n[t] - \overline{f}^n\right) + \langle \mathbf{Q}[t_k^n], \mathbf{z}^n[t] - \overline{\mathbf{g}}^n \rangle \right) \Big| \mathcal{F}_k^n \right) \\
&\quad + \mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} (t - t_k^n) \Big| \mathcal{F}_k^n \right) \cdot 2L(N z_{\max} + d_{\max}) z_{\max} \\
&\le \mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} \left(V\left(y^n[t] - \overline{f}^n\right) + \langle \mathbf{Q}[t_k^n], \mathbf{z}^n[t] - \overline{\mathbf{g}}^n \rangle \right) \Big| \mathcal{F}_k^n \right) \\
(16) &\quad + \mathbb{E}\left((T_k^n)^2 \big| \mathcal{F}_k^n \right) \cdot L(N z_{\max} + d_{\max}) z_{\max},
\end{aligned}$$

where we use the fact that $0 + 1 + \cdots + T_k^n - 1 = (T_k^n - 1)T_k^n / 2 \le (T_k^n)^2$ in the last inequality.

Next, by the queue updating rule (6), $Q_l[t_k^n]$ is determined by $z_l^n[0], \cdots, z_l^n[t_k^n - 1]$ ($n = 1, 2, \cdots, N$) and $d_l[0], \cdots, d_l[t_k^n - 1]$ for any $l \in \{1, 2, \cdots, L\}$. Thus, by Lemma 3.5, $\mathbf{Q}[t_k^n]$ is determined by $\mathcal{F}_k^n$. For the proposed algorithm, each system makes decisions purely based on the virtual queue state $\mathbf{Q}[t_k^n]$, and by the renewal property of each system, given the decision at the k-th renewal, the random quantities T_k^n, $\mathbf{z}^n[t]$ and $y^n[t]$, $t \in \mathcal{T}_k^n$ are independent of the

outcomes from the slots before t_k^n. This implies the following display,

$$\mathbb{E}\left(\sum_{t\in\mathcal{T}_k^n}\left(V\left(y^n[t]-\overline{f}^n\right)+\langle\mathbf{Q}[t_k^n],\mathbf{z}^n[t]-\overline{\mathbf{g}}^n\rangle\right)\middle|\mathcal{F}_k^n\right)$$

$$=\mathbb{E}\left(\sum_{t\in\mathcal{T}_k^n}V\left(y^n[t]-\overline{f}^n\right)\middle|\mathcal{F}_k^n\right)+\left\langle\mathbf{Q}[t_k^n],\mathbb{E}\left(\sum_{t\in\mathcal{T}_k^n}(\mathbf{z}^n[t]-\overline{\mathbf{g}}^n)\middle|\mathcal{F}_k^n\right)\right\rangle$$

$$=\mathbb{E}\left(\sum_{t\in\mathcal{T}_k^n}V\left(y^n[t]-\overline{f}^n\right)\middle|\mathbf{Q}[t_k^n]\right)+\left\langle\mathbf{Q}[t_k^n],\mathbb{E}\left(\sum_{t\in\mathcal{T}_k^n}(\mathbf{z}^n[t]-\overline{\mathbf{g}}^n)\middle|\mathbf{Q}[t_k^n]\right)\right\rangle$$

$$(17)\qquad=\mathbb{E}\left(\sum_{t\in\mathcal{T}_k^n}\left(V\left(y^n[t]-\overline{f}^n\right)+\langle\mathbf{Q}[t_k^n],\mathbf{z}^n[t]-\overline{\mathbf{g}}^n\rangle\right)\middle|\mathbf{Q}[t_k^n]\right),$$

By Lemma 3.4, we have the following:

$$\mathbb{E}\left(\sum_{t\in\mathcal{T}_k^n}(Vy^n[t]+\langle\mathbf{Q}[t_k^n],\mathbf{z}^n[t]\rangle)\middle|\mathbf{Q}[t_k^n]\right)\leq\left(V\overline{f}^n+\langle\mathbf{Q}[t_k^n],\overline{\mathbf{g}}^n\rangle\right)\cdot\mathbb{E}(T_k^n|\mathbf{Q}[t_k^n]).$$

Thus, rearranging terms in above inequality gives the expectation on the right hand side of (17) is no greater than 0 and hence the first expectation on the right hand side of (16) is also no greater than 0. For the second expectation in (16), using (5) in Assumption 2.2 gives $\mathbb{E}\left((T_k^n)^2\middle|\mathcal{F}_k^n\right)\leq B$ and the first part of the lemma is proved.

For the second part of the lemma, by Lemma 3.5 and the definition of Y_K^n, the process $\{Y_K^n\}_{K=0}^\infty$ is adapted to $\{\mathcal{F}_k^n\}_{K=0}^\infty$. Moreover, by Assumption 2.2,

$$\mathbb{E}\left(\left|\sum_{t\in\mathcal{T}_k^n}X^n[t]\right|\right)\leq\mathbb{E}\left(\sum_{t\in\mathcal{T}_k^n}|X^n[t]|\right)<\infty,\ \forall k.$$

Thus, $\mathbb{E}(|Y_K^n|)<\infty$, $\forall K\in\mathbb{N}$, i.e. it is absolutely integrable. Furthermore, by the first part of the lemma,

$$\mathbb{E}\left(Y_{K+1}^n\mid\mathcal{F}_k^n\right)=Y_K^n+\mathbb{E}\left(\left(\sum_{t\in\mathcal{T}_K^n}X^n[t]-C_0\right)\middle|\mathcal{F}_k^n\right)\leq Y_K^n,$$

finishing the proof.   □

3.3. *Synchronization lemma.* So far, we have analyzed the processes related to each individual system over its renewal frames. However, due the asynchronous behavior of different systems, the supermartingales of each system cannot be immediately summed.

In order to get a global performance bound, we have to get rid of any index related to individual renewal frames only. In other words, we need to look at the system property at any time slot T as opposed to any renewal t_k^n. We start with the following standard definition of stopping time:

DEFINITION 3.1.   *Given a probability space $(\Omega,\mathcal{F},P)$ and a filtration $\{\varnothing,\Omega\}=\mathcal{F}_0\subseteq\mathcal{F}_1\subseteq\mathcal{F}_2\cdots$ in $\mathcal{F}$. A stopping time τ with respect to the filtration $\{\mathcal{F}_i\}_{i=0}^\infty$ is a random variable such that for any $i\in\mathbb{N}$,*

$$\{\tau=i\}\in\mathcal{F}_i,$$

i.e. the stopping time occurring at time i is contained in the information during slots 0, 1, 2, $\cdots$, $i-1$.

Next, for any fixed slot $T > 0$, let $S^n[T]$ be the number of renewals up to (and including) time slot T, with with the convention that the first renewal occurs at time $t = 0$, so $t_0^n = 0$ and $S^n[0] = 1$, i.e. $t_0^n = 0$. The next lemma shows $S^n[T]$ is a valid stopping time, whose proof is in the appendix.

LEMMA 3.7. *For each $n \in \{1, 2, \cdots, N\}$, the random variable $S^n[T]$ is a stopping time with respect to the filtration $\{\mathcal{F}_k^n\}_{k=0}^\infty$, i.e. $\{S^n[T] = k\} \in \mathcal{F}_k^n$, $\forall k \in \mathbb{N}$.*

The following theorem tells us a stopping-time truncated supermartingale is still a supermartingale.

THEOREM 3.1 (Theorem 5.2.6 in [7]). *If τ is a stopping time and $Z[i]$ is a supermartingale with respect to $\{\mathcal{F}_i\}_{i=0}^\infty$, then $Z[i \wedge \tau]$ is also a supermartingale, where $a \wedge b \triangleq \min\{a, b\}$.*

With this theorem and the above stopping time construction, we have the following lemma:

LEMMA 3.8. *For each $n \in \{1, 2, \cdots, N\}$ and any fixed $T \in \mathbb{N}$, we have*

$$\frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}(X^n[t]) \le C_1 + \frac{C_2 V}{T},$$

where

$$C_1 := 6L z_{\max}(N z_{\max} + d_{\max})B, \quad C_2 := 2y_{\max}\sqrt{B}.$$

PROOF. First, note that the renewal index k starts from 0. Thus, for any fixed $T \in \mathbb{N}$, $t_{S^n[T]-1}^n \le T < t_{S^n[T]}^n$, and

$$
\begin{aligned}
\mathbb{E}\left(\sum_{t=0}^{T-1} X^n[t]\right) &= \mathbb{E}\left(\sum_{t=0}^{t_{S^n[T]}^n-1} X^n[t] - \sum_{t=T}^{t_{S^n[T]}^n-1} X^n[t]\right) \\
&= \mathbb{E}\left(\sum_{t=1}^{t_{S^n[T]}^n-1} X^n[t]\right) - \mathbb{E}\left(\sum_{t=T}^{t_{S^n[T]}^n-1} X^n[t]\right) \\
&= \mathbb{E}\left(Y_{S^n[T]}^n\right) + C_0 \mathbb{E}(S^n[T]) - \mathbb{E}\left(\sum_{t=T}^{t_{S^n[T]}^n-1} X^n[t]\right) \\
&\le \mathbb{E}\left(Y_{S^n[T]}^n\right) + C_0(T+1) - \mathbb{E}\left(\sum_{t=T}^{t_{S^n[T]}^n-1} X^n[t]\right),
\end{aligned}
$$

(18)

where the third equality follows from the definition of Y_K^n in Lemma 3.6 and the last inequality follows from the fact that the number of renewals up to time slot T is no more than the total number of slots, i.e. $S^n[T] \le T + 1$. For the term $\mathbb{E}\left(Y_{S^n[T]}^n\right)$, we apply Theorem 3.1 with $\tau = S^n[T]$ and index K to obtain $\{Y_{K \wedge S^n[T]}^n\}_{K=0}^\infty$ is a supermartingale. This implies

$$\mathbb{E}\left(Y_{K \wedge S^n[T]}^n\right) \le \mathbb{E}\left(Y_{0 \wedge S^n[T]}^n\right) = \mathbb{E}(Y_0^n) = 0, \; \forall K \in \mathbb{N}.$$

Since $S^n[T] \leq T + 1$, it follows by substituting $K = T + 1$,

$$\mathbb{E}\left(Y_{S^n[T]}^n\right) = \mathbb{E}\left(Y_{(T+1)\wedge S^n[T]}^n\right) \leq 0.$$

For the last term in (18), by queue updating rule (6), for any $l \in \{1, 2, \cdots, L\}$,

$$|Q_l[t]| \leq \sum_{s=0}^{t-1} \left|\sum_{m=1}^{N} z_l^m[s] - d_l[t]\right| \leq t(Nz_{\max} + d_{\max}),$$

it then follows from Hölder's inequality again that

$$
\begin{aligned}
\mathbb{E}\left(\left|\sum_{t=T}^{t_{S^n[T]}^n - 1} X^n[t]\right|\right) &= \mathbb{E}\left(\left|\sum_{t=T}^{t_{S^n[T]}^n - 1} \left(V(y^n[t] - \overline{f}^n) + \langle \mathbf{Q}[t], \mathbf{z}^n[t] - \overline{\mathbf{g}}^n\rangle\right)\right|\right) \\
&\leq \mathbb{E}\left(\sum_{t=T}^{t_{S^n[T]}^n - 1} \left(V\left|y^n[t] - \overline{f}^n\right| + \|\mathbf{Q}[t]\|_1 \cdot \|\mathbf{z}^n[t] - \overline{\mathbf{g}}^n\|_\infty\right)\right) \\
&\leq \mathbb{E}\left(\sum_{t=T}^{t_{S^n[T]}^n - 1} \left(2Vy_{\max} + L(Nz_{\max} + d_{\max})t \cdot 2z_{\max}\right)\right) \\
&= 2Vy_{\max} \cdot \mathbb{E}\left(t_{S^n[T]}^n - T\right) + Lz_{\max}(Nz_{\max} + d_{\max}) \\
&\quad \cdot \left((2T - 1) \cdot \mathbb{E}\left(t_{S^n[T]}^n - T\right) + \mathbb{E}\left(t_{S^n[T]}^n - T\right)^2\right) \\
&\leq 2Vy_{\max}\sqrt{B} + 2Lz_{\max}(Nz_{\max} + d_{\max})\sqrt{B}T + Lz_{\max}(Nz_{\max} + d_{\max})B \\
&\leq 2Vy_{\max}\sqrt{B} + 2Lz_{\max}(Nz_{\max} + d_{\max})B(T + 1),
\end{aligned}
$$

where in the second from last inequality we use (5) of Assumption 2.2 that the residual life $t_{S^n[T]}^n - T$ satisfies

$$\mathbb{E}\left((t_{S^n[T]}^n - T)^2\right) = \mathbb{E}\left(\mathbb{E}\left((t_{S^n[T]}^n - T)^2\Big|\ t_{S^n[T]}^n - t_{S^n[T]-1}^n \geq T - t_{S^n[T]-1}^n\right)\right) \leq B$$

and $\mathbb{E}\left(t_{S^n[T]}^n - T\right) \leq \sqrt{B}$, and in the last inequality we use the fact that $B \geq 1$, thus, $\sqrt{B} \leq B$. Substitute the above bound into (18) gives

$$
\begin{aligned}
\mathbb{E}\left(\sum_{t=0}^{T-1} X^n[t]\right) &\leq C_0(T + 1) + 2Vy_{\max}B + 2Lz_{\max}(Nz_{\max} + d_{\max})B(T + 1) \\
&= 2Vy_{\max}\sqrt{B} + 3Lz_{\max}(Nz_{\max} + d_{\max})B(T + 1) \\
&\leq 2Vy_{\max}\sqrt{B} + 6Lz_{\max}(z_{\max} + d_{\max})BT
\end{aligned}
$$

where we use the definition $C_0 = Lz_{\max}(z_{\max} + d_{\max})B$ from Lemma 3.6 in the equality and use $T + 1 \leq 2T$ in the final equality. Dividing both sides by T finishes the proof. $\square$

3.4. *Achieving near optimality.* The following theorem gives the performance bound of our proposed algorithm.

THEOREM 3.2. *The sequences $\{y^n[t]\}_{t=0}^{\infty}$ and $\{\mathbf{z}^n[t]\}_{t=0}^{\infty}$ produced by the proposed algorithm satisfy all the constraints in* (2) *and achieves* $\mathcal{O}(1/V)$ *near optimality, i.e.*

$$\limsup_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(y^n[t]) \leq f_* + \frac{NC_1 + C_3}{V},$$

where f_* is the optimal objective of (1)-(2), C_1 *is defined in Lemma* 3.8 *and* $C_3 := (Nz_{\max} + d_{\max})^2 L/2$.

PROOF. Define the drift-plus-penalty expression at time t as

$$(19) \qquad P[t] := \mathbb{E}\left(\sum_{n=1}^{N} Vy^n[t] + \frac{1}{2}\left(\|\mathbf{Q}[t+1]\|^2 - \|\mathbf{Q}[t]\|^2\right)\right).$$

By the queue updating rule (6), we have

$$P[t] \leq \mathbb{E}\left(\sum_{n=1}^{N} Vy^n[t] + \frac{1}{2}\sum_{l=1}^{L}\left(\sum_{n=1}^{N} z_l^n[t] - d_l[t]\right)^2 + \sum_{l=1}^{L} Q_l[t]\left(\sum_{n=1}^{N} z_l^n[t] - d_l[t]\right)\right)$$

$$\leq \frac{1}{2}(Nz_{\max} + d_{\max})^2 L + \mathbb{E}\left(\sum_{n=1}^{N} Vy^n[t] + \sum_{l=1}^{L} Q_l[t]\left(\sum_{n=1}^{N} z_l^n[t] - d_l[t]\right)\right)$$

$$= \frac{1}{2}(Nz_{\max} + d_{\max})^2 L + \mathbb{E}\left(\sum_{n=1}^{N} Vy^n[t] + \sum_{l=1}^{L} Q_l[t]\left(\sum_{n=1}^{N} z_l^n[t] - d_l\right)\right)$$

where the second inequality follows from the boundedness assumption (Assumption 2.2) that $\sum_{l=1}^{L}\left(\sum_{n=1}^{N} z_l^n[t] - d_l[t]\right)^2 \leq (Nz_{\max} + d_{\max})^2 L$, and the equality follows from the fact that $d_l[t]$ is i.i.d. and independent of $Q_l[t]$, thus,

$$\mathbb{E}(Q_l[t]d_l[t]) = \mathbb{E}(Q_l[t] \cdot \mathbb{E}(d_l[t]|Q_l[t])) = \mathbb{E}(Q_l[t]d_l).$$

For simplicity, define $C_3 = \frac{1}{2}(Nz_{\max} + d_{\max})^2 L$. Now, by the achievability of optimality in $\otimes_{n=1}^{N}\mathcal{P}^n$ (Lemma 3.2), we have $\sum_{n=1}^{N} \overline{g}_{l,*}^n \leq d_l$, thus, substituting this inequality into the above bound for $P[t]$ gives

$$P[t] \leq C_3 + \mathbb{E}\left(\sum_{n=1}^{N} Vy^n[t] + \sum_{n=1}^{N}\sum_{l=1}^{L} Q_l[t]\left(z_l^n[t] - \overline{g}_{l,*}^n\right)\right)$$

$$= C_3 + \sum_{n=1}^{N}\mathbb{E}(Vy^n[t] + \langle \mathbf{Q}[t], \mathbf{z}^n[t] - \overline{\mathbf{g}}_*^n\rangle)$$

$$= C_3 + \sum_{n=1}^{N}\mathbb{E}(X^n[t]) + V\sum_{n=1}^{N} \overline{f}_*^n$$

$$= C_3 + \sum_{n=1}^{N}\mathbb{E}(X^n[t]) + Vf_*,$$

where we use the definition of $X^n[t]$ in (13) by substituting $(\overline{f}^n, \overline{\mathbf{g}}^n)$ with $(\overline{f}_*^n, \overline{\mathbf{g}}_*^n)$, i.e. $X^n[t] = V(y^n[t] - \overline{f}_*^n) + \langle\mathbf{Q}[t], \mathbf{z}^n[t] - \overline{\mathbf{g}}_*^n\rangle$, in the second from last equality and use the optimality condition

(Lemma 3.2) in the final equality. Now, by Lemma 3.8, we have for any $T \in \mathbb{N}$,

$$
\begin{aligned}
\frac{1}{T} \sum_{t=0}^{T-1} P[t] \leq & C_3 + V f_* + \frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} \mathbb{E}(X^n[t]) \\
= & C_3 + V f_* + \sum_{n=1}^{N} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}(X^n[t]) \\
\leq & N C_1 + C_3 + V f_* + \frac{N C_2 V}{T}.
\end{aligned}
$$

(20)

On the other hand, by the definition of $P[t]$ in (19) and then telescoping sums with $\mathbf{Q}[0] = 0$, we have

$$
\begin{aligned}
\frac{1}{T} \sum_{t=0}^{T-1} P[t] = & \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\left(\sum_{n=1}^{N} V y^n[t] + \frac{1}{2} \left(\|\mathbf{Q}[t+1]\|^2 - \|\mathbf{Q}[t]\|^2 \right) \right) \\
= & \frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} V \mathbb{E}(y^n[t]) + \frac{1}{2T} \mathbb{E}(\|\mathbf{Q}[T]\|^2).
\end{aligned}
$$

Combining this with inequality (20) gives

(21)
$$
\frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} V \mathbb{E}(y^n[t]) + \frac{1}{2T} \mathbb{E}(\|\mathbf{Q}[T]\|^2) \leq N C_1 + C_3 + V f_* + \frac{N C_2 V}{T}.
$$

Since $\frac{1}{2T} \mathbb{E}(\|\mathbf{Q}[T]\|_2^2) \geq 0$, we can throw away the term and the inequality still holds, i.e.

(22)
$$
\frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} \mathbb{E}(y^n[t]) \leq f_* + \frac{N C_1 + C_3}{V} + \frac{N C_2}{T}.
$$

Taking $\limsup_{T \to \infty}$ from both sides gives the near optimality in the theorem.

To get the constraint violation bound, we use Assumption 2.2 that $|y^n[t]| \leq y_{\max}$, then, by (21) again, we have

$$
\frac{1}{T} \mathbb{E}(\|\mathbf{Q}[T]\|^2) \leq 2(N C_1 + C_3) + 4 V y_{\max} + \frac{2 N C_2 V}{T}.
$$

By Jensen's inequality $\mathbb{E}(\|\mathbf{Q}[T]\|^2) \geq \mathbb{E}(\|\mathbf{Q}[T]\|)^2$. This implies that

$$
\mathbb{E}(\|\mathbf{Q}[T]\|) \leq \sqrt{(2(N C_1 + C_3) + 4 V y_{\max})T + 2 N C_2 V},
$$

which implies

(23)
$$
\frac{1}{T} \mathbb{E}(\|\mathbf{Q}[T]\|) \leq \sqrt{\frac{2(N C_1 + C_3) + 4 V y_{\max}}{T} + \frac{2 N C_2 V}{T^2}}.
$$

Sending $T \to \infty$ gives

$$
\lim_{T \to \infty} \frac{1}{T} \mathbb{E}(Q_l[T]) = 0, \quad \forall l \in \{1, 2, \cdots, L\}.
$$

Finally, by Lemma 2.1, all constraints are satisfied. $\square$

Note that the above proof implies a more refined result that illustrates the convergence time. Fix an $\varepsilon > 0$, let $V = 1/\varepsilon$, then, for all $T \geq 1/\varepsilon$, (22) implies that

$$\frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} \mathbb{E}(y^n[t]) \leq f_* + \mathcal{O}(\varepsilon).$$

However, (23) suggests a larger convergence time is required for constraint satisfaction! For $V = 1/\varepsilon$, it can be shown that (23) implies that

$$\frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} \mathbb{E}(z_l^n[t]) \leq d_l + \mathcal{O}(\varepsilon),$$

whenever $T \geq 1/\varepsilon^3$. The next section shows a tighter $1/\varepsilon^2$ convergence time with a mild Lagrange multiplier assumption.

4. Convergence Time Analysis.

4.1. *Lagrange Multipliers.* Consider the following optimization problem:

$$\text{(24)} \qquad \min \ \sum_{n=1}^{N} \overline{f}^n$$

$$\text{(25)} \qquad s.t. \ \sum_{n=1}^{N} \overline{g}_l^n \leq d_l, \ \forall l \in \{1, 2, \cdots, L\},$$

$$\text{(26)} \qquad (\overline{f}^n, \overline{\mathbf{g}}^n) \in \mathcal{P}^n, \ \forall n \in \{1, 2, \cdots, N\}.$$

Since $\mathcal{P}^n$ is convex, it follows $\mathcal{P}^n$ is convex and $\otimes_{n=1}^{N} \mathcal{P}^n$ is also convex. Thus, (24)-(26) is a convex program. Furthermore, by Lemma 3.2, we have (24)-(26) is feasible if and only if (1)-(2) is feasible, and when assuming feasibility, they have the same optimality f_* as is specified in Lemma 3.2.

Since $\mathcal{P}^n$ is convex, one can show (see Proposition 5.1.1 of [2]) that there *always* exists a sequence $(\gamma_0, \gamma_1, \cdots, \gamma_L)$ so that $\gamma_i \geq 0, \ i = 0, 1, \cdots, L$ and

$$\sum_{n=1}^{N} \gamma_0 \overline{f}^n + \sum_{l=1}^{L} \gamma_l \sum_{n=1}^{N} \overline{g}_l^n \geq \gamma_0 f_* + \sum_{l=1}^{L} \gamma_l d_l, \ \forall (\overline{f}^n, \overline{\mathbf{g}}^n) \in \mathcal{P}^n,$$

i.e. there always exists a hyperplane parametrized by $(\gamma_0, \gamma_1, \cdots, \gamma_L)$, supported at $(f_*, d_1, \cdots, d_L)$ and containing the set $\left\{ \left(\sum_{n=1}^{N} \overline{f}^n, \ \sum_{n=1}^{N} \overline{\mathbf{g}}^n \right) : (\overline{f}^n, \overline{\mathbf{g}}^n) \in \mathcal{P}^n, \ \forall n \in \{1, 2, \cdots, N\} \right\}$ on one side. This hyperplane is called "separating hyperplane". The following assumption stems from this property and simply assumes this separating hyperplane to be non-vertical (i.e. $\gamma_0 > 0$):

ASSUMPTION 4.1. *There exists non-negative finite constants* $\gamma_1, \ \gamma_2, \ \cdots, \ \gamma_L$ *such that the following holds,*

$$\sum_{n=1}^{N} \overline{f}^n + \sum_{l=1}^{L} \gamma_l \sum_{n=1}^{N} \overline{g}_l^n \geq f_* + \sum_{l=1}^{L} \gamma_l d_l, \ \forall (\overline{f}^n, \overline{\mathbf{g}}^n) \in \mathcal{P}^n,$$

i.e. there exists a separating hyperplane parametrized by $(1, \gamma_1, \cdots, \gamma_L)$.

REMARK 4.1.   *The parameters $\gamma_1, \cdots, \gamma_L$ are called Lagrange multipliers and this assumption is equivalent to the existence of Lagrange multipliers for constrained convex program (24)-(26). It is known that Lagrange multipliers exist if the Slater's condition holds ([2]), which states that there exists a nonempty interior of the feasible region for the convex program. Slater's condition is very common in convex optimization theory and plays an important role in convergence rate analysis, such as the analysis of the interior point algorithm ([5]). In the current context, this condition is satisfied, for example, in energy aware server scheduling problems, if the highest possible sum of service rates from all servers is strictly higher than the arrival rate.*

LEMMA 4.1.   *Suppose $\{y^n[t]\}_{t=0}^{\infty}$, $\{\mathbf{z}^n[t]\}_{t=0}^{\infty}$ and $\{T_k^n\}_{k=0}^{\infty}$ are processes resulting from the proposed algorithm. Under the Assumption 4.1,*

$$\frac{1}{T}\sum_{t=0}^{T-1}\left(f_* - \sum_{n=1}^{N}\mathbb{E}(y^n[t])\right) \leq \frac{1}{T}\sum_{t=0}^{T-1}\sum_{l=1}^{L}\gamma_l\left(\sum_{n=1}^{N}\mathbb{E}(z_l^n[t]) - d_l\right) + \frac{C_4}{T},$$

where $C_4 = B_1 N + B_2 N \sum_{l=1}^{L}\gamma_l$, and B_1, B_2 are defined in Lemma 3.3.

PROOF. First of all, from the statement of Lemma 3.3, for the proposed algorithm, we can define the corresponding processes $(f^n[t], \mathbf{g}^n[t])$ for all n as

$$f^n[t] = \widehat{f}^n(\alpha^n) = \widehat{y}^n(\alpha^n)/\widehat{T}^n(\alpha^n), \quad \text{if } t \in \mathcal{T}_k^n, \alpha_k^n = \alpha^n$$
$$\mathbf{g}^n[t] = \widehat{\mathbf{g}}^n(\alpha^n) = \widehat{\mathbf{z}}^n(\alpha^n)/\widehat{T}^n(\alpha^n), \quad \text{if } t \in \mathcal{T}_k^n, \alpha_k^n = \alpha^n,$$

where the last equality follows from the definition of $\widehat{f}^n(\alpha^n)$ and $\widehat{\mathbf{g}}^n(\alpha^n)$ in Definition 2.1. Since $\left(\widehat{y}^n(\alpha^n), \widehat{\mathbf{z}}^n(\alpha^n), \widehat{T}^n(\alpha^n)\right) \in \mathcal{S}^n$, by definition of $\mathcal{P}^n$ in Definition 2.2, $(f^n[t], \mathbf{g}^n[t]) \in \mathcal{P}^n \subseteq \mathcal{P}^n$, $\forall n$, $\forall t$. Since $\mathcal{P}^n$ is a convex set by Lemma 3.1, it follows

$$(\mathbb{E}(f^n[t]), \mathbb{E}(\mathbf{g}^n[t])) \in \mathcal{P}^n, \quad \forall t, \forall n.$$

By Assumption 4.1, we have

$$\sum_{n=1}^{N}\mathbb{E}(f^n[t]) + \sum_{l=1}^{L}\gamma_l\sum_{n=1}^{N}\mathbb{E}(g_l^n[t]) \geq f_* + \sum_{l=1}^{L}\gamma_l d_l, \quad \forall t.$$

Rearranging terms gives

$$f_* - \sum_{n=1}^{N}\mathbb{E}(f^n[t]) \leq \sum_{l=1}^{L}\gamma_l\left(\sum_{n=1}^{N}\mathbb{E}(g_l^n[t]) - d_l\right), \quad \forall t.$$

Taking the time average from 0 to $T-1$ gives

$$(27) \qquad \frac{1}{T}\sum_{t=0}^{T-1}\left(f_* - \sum_{n=1}^{N}\mathbb{E}(f^n[t])\right) \leq \frac{1}{T}\sum_{t=0}^{T-1}\sum_{l=1}^{L}\gamma_l\left(\sum_{n=1}^{N}\mathbb{E}(g_l^n[t]) - d_l\right).$$

For the left hand side of (27), we have

$$l.h.s. = \frac{1}{T}\sum_{t=0}^{T-1}\left(f_* - \sum_{n=1}^{N}\mathbb{E}(y^n[t])\right) + \frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(y^n[t] - f^n[t])$$

$$(28) \qquad \geq \frac{1}{T}\sum_{t=0}^{T-1}\left(f_* - \sum_{n=1}^{N}\mathbb{E}(y^n[t])\right) - \frac{B_1 N}{T}.$$

where the inequality follows from (10) in Lemma 3.3. For the right hand side of (27), we have

$$r.h.s. = \frac{1}{T}\sum_{t=0}^{T-1}\sum_{l=1}^{L}\gamma_l\left(\sum_{n=1}^{N}\mathbb{E}(z_l^n[t]) - d_l\right) + \frac{1}{T}\sum_{t=0}^{T-1}\sum_{l=1}^{L}\gamma_l\sum_{n=1}^{N}\mathbb{E}(g_l^n[t] - z_l^n[t])$$

$$(29) \qquad \leq \frac{1}{T}\sum_{t=0}^{T-1}\sum_{l=1}^{L}\gamma_l\left(\sum_{n=1}^{N}\mathbb{E}(z_l^n[t]) - d_l\right) + \frac{B_2 N\sum_{l=1}^{L}\gamma_l}{T},$$

where the inequality follows from the fact that $\gamma_l \geq 0$, $\forall l$ and (11) in Lemma 3.3. Substituting (28) and (29) into (27) finishes the proof. $\qquad\square$

4.2. Convergence time theorem.

THEOREM 4.1. *Fix $\varepsilon \in (0,1)$ and define $V = 1/\varepsilon$. If the problem (1)-(2) is feasible and the Assumption 4.1 holds, then, for all $T \geq 1/\varepsilon^2$,*

$$(30) \qquad \frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(y^n[t]) \leq f_* + \mathcal{O}(\varepsilon),$$

$$(31) \qquad \frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(z_l^n[t]) \leq d_l + \mathcal{O}(\varepsilon), l \in \{1,2,\cdots,L\}.$$

Thus, the algorithm provides $\mathcal{O}(\varepsilon)$ approximation with the convergence time $\mathcal{O}(1/\varepsilon^2)$.

PROOF. First of all, by queue updating rule (6),

$$(32) \qquad \sum_{t=0}^{T-1}\left(\sum_{n=1}^{N}\mathbb{E}(z_l^n[t]) - d_l\right) \leq \mathbb{E}(Q_l[T]).$$

By Lemma 4.1, we have

$$\frac{1}{T}\sum_{t=0}^{T-1}\left(f_* - \sum_{n=1}^{N}\mathbb{E}(y^n[t])\right) \leq \frac{1}{T}\sum_{t=0}^{T-1}\sum_{l=1}^{L}\gamma_l\left(\sum_{n=1}^{N}\mathbb{E}(z_l^n[t]) - d_l\right) + \frac{C_4}{T},$$

$$(33) \qquad \leq \sum_{l=1}^{L}\frac{\gamma_l}{T}\mathbb{E}(Q_l[T]) + \frac{C_4}{T}.$$

Combining this with (21) gives

$$\frac{1}{2T}\mathbb{E}\big(\|\mathbf{Q}[T]\|^2\big) \leq NC_1 + C_3 + \frac{V}{T}\sum_{t=0}^{T-1}\left(f_* - \sum_{n=1}^{N}\mathbb{E}(y^n[t])\right) + \frac{NC_2 V}{T}$$

$$\leq NC_1 + C_3 + \frac{(NC_2 + C_4)V}{T} + V\sum_{l=1}^{L}\frac{\gamma_l}{T}\mathbb{E}(Q_l[T])$$

$$(34) \qquad \leq NC_1 + C_3 + \frac{(NC_2 + C_4)V}{T} + \frac{V}{T}\|\gamma\| \cdot \|\mathbb{E}(\mathbf{Q}[T])\|,$$

where $\gamma := (\gamma_1, \cdots, \gamma_L)$, the second inequality follows from (33) and the final inequality follows from Cauchy-Schwarz. Then, by Jensen's inequality, we have

$$\|\mathbb{E}(\mathbf{Q}[T])\|^2 \leq \mathbb{E}\big(\|\mathbf{Q}[T]\|^2\big).$$

Thus, it follows by (34) that

$$\|\mathbb{E}(\mathbf{Q}[T])\|^2 - 2V\|\gamma\| \cdot \|\mathbb{E}(\mathbf{Q}[T])\| - 2(NC_1 + C_3)T - 2(NC_2 + C_4)V \leq 0.$$

The left hand side is a quadratic form on $\|\mathbb{E}(\mathbf{Q}[T])\|$, and the inequality implies that $\|\mathbb{E}(\mathbf{Q}[T])\|$ is deterministically upper bounded by the largest root of the equation $x^2 - bx - c = 0$ with $b = 2V\|\gamma\|$ and $c = 2(NC_1 + C_3)T + 2(NC_2 + C_4)V$. Thus,

$$
\begin{aligned}
\|\mathbb{E}(\mathbf{Q}[T])\| \leq & \frac{b + \sqrt{b^2 + 4c}}{2} \\
= & V\|\gamma\| + \sqrt{V^2\|\gamma\|^2 + 2(NC_1 + C_3)T + 2(NC_2 + C_4)V} \\
\leq & 2V\|\gamma\| + \sqrt{2(NC_1 + C_3)T} + \sqrt{2(NC_2 + C_4)V}.
\end{aligned}
$$

Thus, for any $l \in \{1, 2, \cdots, L\}$,

$$\frac{1}{T}\mathbb{E}(Q_l[T]) \leq \frac{2V\|\gamma\|}{T} + \sqrt{\frac{2(NC_1 + C_3)}{T}} + \frac{\sqrt{2(NC_2 + C_4)V}}{T}.$$

By (32) again,

$$\frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(z_l^n[t]) \leq d_l + \frac{2V\|\gamma\|}{T} + \sqrt{\frac{2(NC_1 + C_3)}{T}} + \frac{\sqrt{2(NC_2 + C_4)V}}{T}.$$

Substituting $V = 1/\varepsilon$ and $T \geq 1/\varepsilon^2$ into the above inequality gives $\forall l \in \{1, 2, \cdots, L\}$,

$$
\begin{aligned}
\frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(z_l^n[t]) \leq & d_l + \left(2\|\gamma\| + \sqrt{2(NC_1 + C_3)}\right)\varepsilon + \sqrt{2(NC_2 + C4)}\varepsilon^{3/2} \\
= & d_l + \mathcal{O}(\varepsilon).
\end{aligned}
$$

Finally, substituting $V = 1/\varepsilon$ and $T \geq 1/\varepsilon^2$ into (22) gives

$$\frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(y^n[t]) \leq f_* + \mathcal{O}(\varepsilon),$$

finishing the proof. □

5. Simulation Study in Energy-aware Scheduling. Here, we apply the algorithm introduced in Section 2 to deal with the energy-aware scheduling problem described in Section 1.1. To be specific, we consider a scenario with 5 homogeneous servers and 3 different classes of jobs, i.e. $N = 5$ and $L = 3$. We assume that each server can only choose one class of jobs to serve during each frame. So the mode set $\mathcal{M}^n$ contains three actions $\{1, 2, 3\}$ and the action i stands for serving the i-th class of jobs and we count the number of serviced jobs at the end of each service duration. The action m_k^n determines the following quantities:

- The uniformly distributed total number of class l jobs that can be served with expectation $\mathbb{E}\left(\sum_{t\in\mathcal{T}_k^n}\mu_l^n[t]\Big|\ m_k^n\right) := \widehat{\mu}_l^n(m_k^n)$.
- The geometrically distributed service duration H_k^n slots with expectation $\mathbb{E}(H_k^n|\ m_k^n) := \widehat{H}^n(m_k^n)$.
- The energy consumption $\widehat{e}^n(m_k^n)$ for serving all these jobs.

TABLE 1
Problem parameters

	λ_i	$\widehat{H}^n(i)$	$\widehat{\mu}^n(i)$	$\widehat{e}^n(i)$	$\widehat{I}^n(i)$
Class 1	2	5.5	15 (Uniform $[9, 21] \cap \mathbb{N}$)	16	2.5
Class 2	3	4.6	21 (Uniform $[15, 27] \cap \mathbb{N}$)	20	4.3
Class 3	4	3.8	17 (Uniform $[11, 23] \cap \mathbb{N}$)	13	3.7

- The geometrically distributed idle/setup time I_k^n slots with constant energy consumption p^n per slot and zero job service. The expectation $\mathbb{E}(I_k^n|\ m_k^n) := \widehat{I}^n(m_k^n)$.

The idle/setup cost is $p^n = 3$ units per slot and the rest of the parameters are listed in Table 1.

Following the algorithm description in Section 2, the proposed algorithm has the queue updating rule

$$Q_l[t+1] = \max\left\{Q_l[t] + \lambda_l[t] - \sum_{n=1}^{N} \mu_l^n[t],\ 0\right\},$$

and each system minimizes (7) each frame, which can be written as

$$\min_{m_k^n \in \mathcal{M}^n} \frac{V\left(\widehat{e}_l^n(m_k^n) + p^n\widehat{I}^n(m_k^n)\right) - \langle \mathbf{Q}[t_k^n], \widehat{\mu}^n(m_k^n)\rangle}{\widehat{H}^n(m_k^n) + \widehat{I}^n(m_k^n)}.$$

Each plot for the proposed algorithm is the result of running 1 million slots and taking the time average as the performance of the proposed algorithm. The benchmark is the optimal stationary performance obtained by performing a change of variable and solving a linear program, knowing the arrival rates (see also [12] for details).

Fig. 3 shows as the trade-off parameter V gets larger, the time average energy consumptions under the proposed algorithm approaches the optimal energy consumption. Fig. 4 shows as V gets large, the time average number of services also approaches the optimal service rate for each class of jobs. In Fig. 5, we plot the time average queue backlog for each class of jobs verses V parameter. We see that the queue backlog for the first class is always low whereas the rest queue backlogs scale up linearly with V. This is because the service rate for the first class is always strictly larger than the arrival rate whereas for the rest classes, as V gets larger, the service rates approach the arrival rates. This plot, together with Fig. 3, also demonstrate that V is indeed a trade-off parameter which trades queue backlog for near optimality.

APPENDIX A: ADDITIONAL LEMMAS AND PROOFS.

PROOF OF LEMMA 2.1. Fix $l \in \{1, 2, \cdots, L\}$. For any fixed T, $Q_l[T] = \sum_{t=0}^{T-1}(Q_l[t+1] - Q_l[t])$. For each summand, by queue updating rule (6),

$$Q_l[t+1] - Q_l[t] = \max\left\{Q_l[t] + \sum_{n=1}^{N} z_l^n[t] - d_l[t],\ 0\right\} - Q_l[t]$$

$$\geq Q_l[t] + \sum_{n=1}^{N} z_l^n[t] - d_l[t] - Q_l[t] = \sum_{n=1}^{N} z_l^n[t] - d_l[t].$$

Thus, by the assumption $Q_l[0] = 0$,

$$Q_l[T] \geq \sum_{t=0}^{T-1}\left(\sum_{n=1}^{N} z_l^n[t] - d_l[t]\right).$$
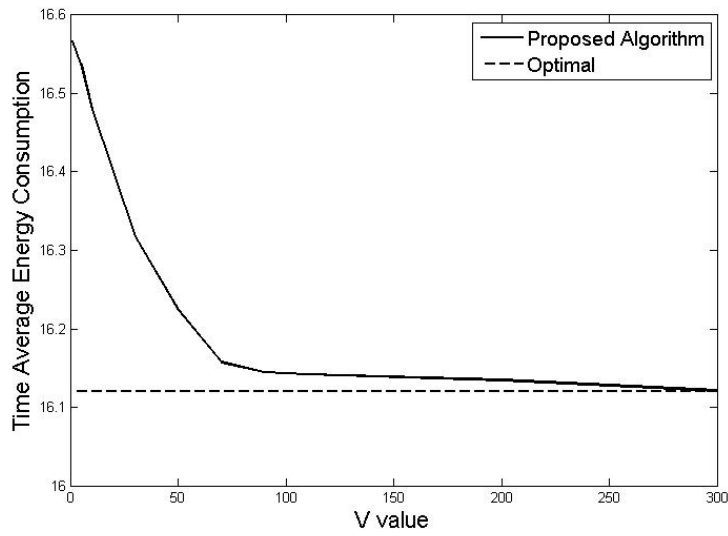
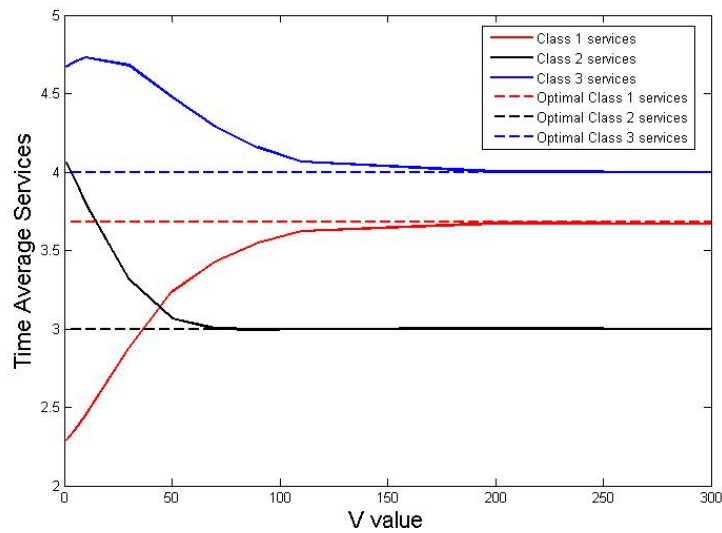Fig 3. *Time average energy consumption verses V parameter over 1 millon slots.*

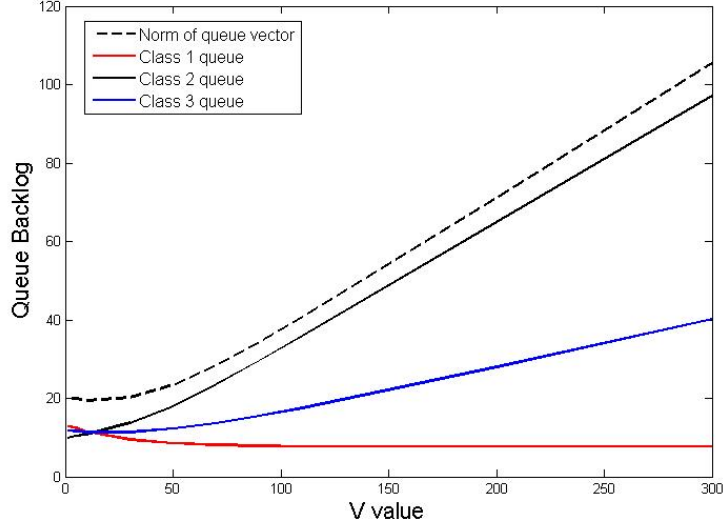Fig 4. *Time average services verses V parameter over 1 millon slots.*

FIG 5. *Time average queue size verses V parameter over 1 million slots.*

Taking expectations of both sides with $\mathbb{E}(d_l[t]) = d_l$, $\forall l$, gives

$$\mathbb{E}(Q_l[T]) \geq \sum_{t=0}^{T-1} \left(\sum_{n=1}^{N} \mathbb{E}(z_l^n[t]) - d_l \right).$$

Dividing both sides by T and passing to the limit gives

$$\limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \left(\sum_{n=1}^{N} \mathbb{E}(z_l^n[t]) - d_l \right) \leq \lim_{T \to \infty} \frac{1}{T} \mathbb{E}(Q_l[T]) = 0,$$

finishing the proof. $\qquad\square$

PROOF OF LEMMA 3.3. We prove bound (10) ((11) is proved similarly). By definition of $\widehat{f}^n(\alpha^n)$ in Definition 2.1, we have for any $\alpha^n \in \mathcal{A}^n$,

$$\widehat{f}^n(\alpha^n) = \frac{\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} y^n[t] \,\middle|\, \alpha_k^n = \alpha^n \right)}{\mathbb{E}\left(T_k^n \,\middle|\, \alpha_k^n = \alpha^n \right)},$$

thus,

$$\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} \left(\widehat{f}^n(\alpha_k^n) - y^n[t] \right) \,\middle|\, \alpha_k^n = \alpha^n \right) = 0.$$

By the renewal property of the system, given $\alpha_k^n = \alpha^n$, T_k^n and $\sum_{t \in \mathcal{T}_k^n} y^n[t]$ are independent of the past information before t_k^n. Thus, the same equality holds if conditioning also on $\mathcal{F}_k^n$, i.e.

$$\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} \left(\widehat{f}^n(\alpha_k^n) - y^n[t] \right) \,\middle|\, \alpha_k^n = \alpha^n, \ \mathcal{F}_k^n \right) = 0.$$

Hence,

$$\mathbb{E}\left(\left. \sum_{t\in\mathcal{T}_k^n} \left(\widehat{f}^n(\alpha_k^n) - y^n[t] \right) \right| \mathcal{F}_k^n \right) = 0.$$

By the definition of $f^n[t]$, this further implies that

$$\mathbb{E}\left(\left. \sum_{t\in\mathcal{T}_k^n} \left(f^n[t] - y^n[t] \right) \right| \mathcal{F}_k^n \right) = 0.$$

Since $|y^n[t]| \leq y_{\max}$ and $\mathbb{E}(T_k^n) \leq \sqrt{B}$, it follows $\mathbb{E}\left(\left| \sum_{t\in\mathcal{T}_k^n} \left(f^n[t] - y^n[t] \right) \right| \right) < \infty$ and the process $\{F_K^n\}_{K=0}^\infty$ defined as

$$F_K^n = \sum_{k=0}^{K-1} \sum_{t\in\mathcal{T}_k^n} \left(f^n[t] - y^n[t] \right), \ \ K \geq 1,$$

$F_0^n = 0$ is a *martingale*.

Consider any fixed $T \in \mathbb{N}$ and define $S^n[T]$ as the number of renewals up to T. Lemma 3.7 shows $S^n[T]$ is a valid stopping time with respect to the filtration $\{\mathcal{F}_k^n\}_{k=0}^\infty$. Furthermore, $\{F_{K\wedge S^n[T]}^n\}_{K=0}^\infty$ is a supermartingale by Theorem 3.1, where $a \wedge b := \min\{a, b\}$.

For this fixed T, we have

$$\mathbb{E}\left(\sum_{t=0}^{T-1} \left(f^n[t] - y^n[t] \right) \right) = \mathbb{E}\left(\sum_{t=0}^{t_{S^n[T]}^n - 1} \left(f^n[t] - y^n[t] \right) \right) - \mathbb{E}\left(\sum_{t=T}^{t_{S^n[T]}^n - 1} \left(f^n[t] - y^n[t] \right) \right)$$

$$= \mathbb{E}\left(F_{S^n[T]}^n \right) - \mathbb{E}\left(\sum_{t=T}^{t_{S^n[T]}^n - 1} \left(f^n[t] - y^n[t] \right) \right).$$

Since the number of renewals is always bounded by the number of slots at any time, i.e. $S^n[T] \leq T + 1$, it follows

$$\mathbb{E}\left(F_{S^n[T]}^n \right) = \mathbb{E}\left(F_{(T+1)\wedge S^n[T]}^n \right) \leq 0.$$

On the other hand,

$$\left| \mathbb{E}\left(\sum_{t=T}^{t_{S^n[T]}^n - 1} \left(f^n[t] - y^n[t] \right) \right) \right| \leq \mathbb{E}\left(t_{S^n[T]}^n - T \right) \cdot 2y_{\max} \leq 2y_{\max}\sqrt{B}.$$

where the last inequality follows from Assumption 2.2 for the residual life time. Thus,

$$\mathbb{E}\left(\sum_{t=0}^{T-1} \left(f^n[t] - y^n[t] \right) \right) \leq 2y_{\max}\sqrt{B}.$$

Dividing both sides by T finishes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

PROOF OF LEMMA 3.5. Recall that t_k^n is the time slot where the k-th renewal occurs ($k = 0, 1, 2, \cdots$), then, it follows from the definition of stopping time ([7]) that $\{t_k^n\}_{k=0}^\infty$ is a sequence of stopping times with respect to $\{\mathcal{F}[t]\}_{t=0}^\infty$ satisfying $t_k^n < t_{k+1}^n$, $\forall k$. Thus, by definition of $\mathcal{F}_k^n$, for any set $A \in \mathcal{F}_k^n$,

$$A \cap \{t_{k+1}^n \leq t\} = A \cap \{t_k^n \leq t\} \cap \{t_{k+1}^n \leq t\} \in \mathcal{F}[t].$$

Thus, $A \in \mathcal{F}_{k+1}^n$, which implies $\mathcal{F}_k^n \subseteq \mathcal{F}_{k+1}^n$, $\forall k$, and $\{\mathcal{F}_k^n\}_{k=0}^\infty$ is indeed a filtration. This finishes the first part of the proof.

Next, we would like to show that $G_{t_k^n}(Z_0^n, \cdots, Z^n[t_k^n - 1])$ is measurable with respect to $\mathcal{F}_k^n$, $\forall k \geq 1$, i.e. $\{G_{t_k^n}(Z_0^n, \cdots, Z^n[t_k^n - 1]) \in B\} \in \mathcal{F}_k^n$, for any Borel set $B \subseteq \mathbb{R}$. By definition of $\mathcal{F}_k^n$, this is equivalent to showing $\{G_{t_k^n}(Z_0^n, \cdots, Z^n[t_k^n - 1]) \in B\} \cap \{t_k^n \leq s\} \in \mathcal{F}[s]$ for any slot $s \geq 0$. For $s = 0$, this is obvious because $\{t_k^n \leq 0\} = \emptyset$, $\forall k \geq 1$. Consider any $s \geq 1$,

$$
\begin{aligned}
& \left\{ G_{t_k^n}(Z_0^n, \cdots, Z^n[t_k^n - 1]) \in B \right\} \cap \{t_k^n \leq s\} \\
&= \bigcup_{i=1}^s \left(\{G_i(Z_0^n, \cdots, Z^n[i-1]) \in B\} \bigcap \{t_k^n = i\} \right) \\
&= \bigcup_{i=1}^s \left(\{(Z_0^n, \cdots, Z^n[i-1]) \in G_i^{-1}(B)\} \bigcap \{t_k^n = i\} \right) \in \mathcal{F}[s], \ \forall k \geq 1,
\end{aligned}
$$

where the last step follows from the assumption that the random variable $Z^n[t-1]$ is measurable with respect to $\mathcal{F}[t]$ for any $t > 0$ and t_k^n is a stopping time with respect to $\{\mathcal{F}[t]\}_{t=0}^\infty$ for all $k \geq 1$. This gives the second part of the claim. □

PROOF OF LEMMA 3.7. We aim to prove $\{S^n[T] = k\} \in \mathcal{F}_k^n$, $\forall k \in \mathbb{N}$. First of all, recall that the index of the renewal starts from $k = 0$ and $t_0^n = 0$, thus, for any $k \in \mathbb{N}$, $\{S^n[T] = k\} = \{t_k^n > T\} \cap \{t_{k-1}^n \leq T\}$, and any $t \in \mathbb{N}$,

(35) $$\{S^n[T] = k\} \cap \{t_k^n \leq t\} = \{t_k^n > T\} \cap \{t_{k-1}^n \leq T\} \cap \{t_k^n \leq t\}.$$

Consider two cases as follows:

1. $t \leq T$. In this case, the set (35) is empty and obviously belongs to $\mathcal{F}[t]$.
2. $t > T$. In this case, we have $\{t_k^n > T\} \cap \{t_k^n \leq t\} = \{T < t_k^n \leq t\} \in \mathcal{F}[t]$ as well as $\{t_{k-1}^n \leq T\} \in \mathcal{F}[T] \subseteq \mathcal{F}[t]$. Thus, the set (35) belongs to $\mathcal{F}[t]$.

Overall, we have $\{S^n[T] = k\} \cap \{t_k^n \leq t\} \in \mathcal{F}[t]$, $\forall t \in \mathbb{N}$. Thus, $\{S^n[T] = k\} \in \mathcal{F}_k^n$ and $S^n[T]$ is indeed a valid stopping time with respect to the filtration $\{\mathcal{F}_k^n\}_{k=0}^\infty$. □

The rest of the section is devoted to the proof of Lemma 3.2.

PROOF OF LEMMA 3.2. To prove the first part of the claim, we define the following notation:

$$
\bigoplus_{n=1}^N \mathcal{P}^n := \left\{ \sum_{n=1}^N \mathbf{p}_n, \ \mathbf{p}_n \in \mathcal{P}^n, \ \forall n \right\}
$$

is the Minkowski sum of sets $\mathcal{P}_n$, $n \in \{1, 2, \cdots, N\}$, and for any sequence $\{\mathbf{x}[t]\}_{t=0}^\infty$ taking values in $\mathbb{R}^d$, define

$$
\limsup_{T \to \infty} \mathbf{x}[T] := \left(\limsup_{T \to \infty} x_1[T], \ \cdots, \limsup_{T \to \infty} x_d[T] \right)
$$

is a vector of lim sups. By definition, any vector in $\oplus_{n=1}^N \mathcal{P}^n$ can be constructed from $\otimes_{n=1}^N \mathcal{P}^n$, thus, it is enough to show that there exists a vector $\mathbf{r}^* \in \oplus_{n=1}^N \mathcal{P}^n$ such that $r_0^* = f^*$ and the rest of the entries $r_l^* \leq d_l$, $l = 1, 2, \cdots, L$.

By the feasibility assumption for (1)-(2), we can consider *any algorithm that achieves the optimality* of (1)-(2) and the corresponding process $\{(f^n[t], \mathbf{g}^n[t])\}_{t=0}^\infty$ defined in Lemma 3.3 for

any system n. Notice that $(f^n[t], \mathbf{g}^n[t]) \in \mathcal{P}^n$, $\forall n$, $\forall t$. This follows from the definition of $\widehat{f}^n(\alpha^n)$ and $\widehat{\mathbf{g}}^n(\alpha^n)$ in Definition 2.1 that

$$f^n[t] = \widehat{f}^n(\alpha^n) = \widehat{y}^n(\alpha^n)/\widehat{T}^n(\alpha^n), \quad \text{if } t \in \mathcal{T}_k^n, \alpha_k^n = \alpha^n$$
$$\mathbf{g}^n[t] = \widehat{\mathbf{g}}^n(\alpha^n) = \widehat{\mathbf{z}}^n(\alpha^n)/\widehat{T}^n(\alpha^n), \quad \text{if } t \in \mathcal{T}_k^n, \alpha_k^n = \alpha^n,$$

and $\left(\widehat{y}^n(\alpha^n),\ \widehat{\mathbf{z}}^n(\alpha^n),\ \widehat{T}^n(\alpha^n)\right) \in \mathcal{S}^n$. By definition of $\mathcal{P}^n$ in Definition 2.2, $(f^n[t], \mathbf{g}^n[t]) \in \mathcal{P}^n$, $\forall n$, $\forall t$.

Since $\mathcal{P}^n$ is convex by Lemma 3.1, it follows that $(\mathbb{E}(f^n[t]), \mathbb{E}(\mathbf{g}^n[t])) \in \mathcal{P}^n$, $\forall n$, $\forall t$. Hence,

$$\left(\frac{1}{T}\sum_{t=1}^{T-1}\mathbb{E}(f^n[t]),\ \frac{1}{T}\sum_{t=1}^{T-1}\mathbb{E}(\mathbf{g}^n[t])\right) \in \mathcal{P}^n, \ \forall T, \forall n.$$

This further implies that

$$\mathbf{r}(T) := \left(\frac{1}{T}\sum_{t=1}^{T-1}\sum_{n=1}^{N}\mathbb{E}(f^n[t]),\ \frac{1}{T}\sum_{t=1}^{T-1}\sum_{n=1}^{N}\mathbb{E}(\mathbf{g}^n[t])\right) \in \bigoplus_{n=1}^{N}\mathcal{P}^n.$$

By Lemma 3.1, $\mathcal{P}^n$ is compact in $\mathbb{R}^{L+1}$. Thus, $\oplus_{n=1}^{N}\mathcal{P}^n$ is also compact. This implies that the sequence $\{\mathbf{r}(T)\}_{T=1}^{\infty}$ has at least one limit point, and any such limit point is contained in $\oplus_{n=1}^{N}\mathcal{P}^n$.

We consider a specific limit point of $\{\mathbf{r}(T)\}_{T=1}^{\infty}$ denoted as $\mathbf{r}^* \in \oplus_{n=1}^{N}\mathcal{P}^n$, with the first entry denoted as r_0^* satisfying

$$r_0^* = \limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(f^n[t]).$$

Then, we have the rest of the entries of $\mathbf{r}^*$ must satisfy

$$r_l^* \leq \limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(\mathbf{g}^n[t]), \ \forall l \in \{1, 2, \cdots, L\}.$$

Now, by Lemma 3.3, we can connect the $\limsup$ with respect to $f^n[t]$ and $\mathbf{g}^n[t]$ to that of $y^n[t]$ and $\mathbf{z}^n[t]$ as follows:

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(y^n[t])$$
$$= \limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\left(\mathbb{E}(y^n[t] - f^n[t]) + \mathbb{E}(f^n[t])\right)$$
$$= \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(y^n[t] - f^n[t]) + \limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(f^n[t])$$
$$= \limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(f^n[t]).$$

Similarly, we can show that

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(\mathbf{z}^n[t]) = \limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=1}^{N}\mathbb{E}(\mathbf{g}^n[t]).$$

Thus, by our preceeding assumption that the algorithm under consideration achieves the optimality of (1)-(2), we have

$$r_0^* = \limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} \mathbb{E}(y^n[t]) = f^*$$

$$r_l^* \leq \limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{n=1}^{N} \mathbb{E}(z_l^n[t]) \leq d_l, \ \forall i \in \{1, 2, \cdots, L\}.$$

Overall, we have shown that $\mathbf{r}^* \in \oplus_{n=1}^{N} \mathcal{P}^n$ achieves the optimality of (1)-(2), and the first part of the lemma is proved.

To prove the second part of the lemma, we show that any point in $\otimes_{n=1}^{N} \mathcal{P}^n$ is achievable by the corresponding time averages of some algorithm. Specifically, consider the following class of *randomized stationary algorithms*: For each system n, at the beginning of k-th frame, the controller independently chooses an action α_k^n from the set $\mathcal{A}^n$ with a fixed probability distribution.

Thus, the actions $\{\alpha_k^n\}_{k=0}^{\infty}$ result from any randomized stationary algorithm is i.i.d.. By the renewal property of each system, we have

$$\left\{ \left(\sum_{t \in \mathcal{T}_k^n} y^n[t], \ \sum_{t \in \mathcal{T}_k^n} \mathbf{z}^n[t], \ T_k^n \right) \right\}_{k=0}^{\infty},$$

is also an i.i.d. process for each system n.

Next, we would like to show that any point in $\mathcal{S}^n$ can be achieved by the corresponding expectations of some randomized stationary algorithm. Recall that $\mathcal{S}^n$ defined in Definition 2.2 is the convex hull of

$$\mathcal{G}^n := \left\{ \left(\widehat{y}^n(\alpha^n), \ \widehat{\mathbf{z}}^n(\alpha^n), \ \widehat{T}^n(\alpha^n) \right), \ \alpha^n \in \mathcal{A}^n \right\} \subseteq \mathbb{R}^{L+2},$$

By definition of convex hull, any point $(y, \mathbf{z}, T) \in \mathcal{S}^n$, can be written as a convex combination of a finite number of points from the set $\mathcal{G}^n$. Let $\left\{ \left(\widehat{y}^n(\alpha_i^n), \ \widehat{\mathbf{z}}^n(\alpha_i^n), \ \widehat{T}^n(\alpha_i^n) \right) \right\}_{i=1}^{m}$ be these points, then, we have there exists a finite sequence $\{p_i\}_{i=1}^{m}$, such that

$$(y, \mathbf{z}, T) = \sum_{i=1}^{m} p_i \cdot \left(\widehat{y}^n(\alpha_i^n), \ \widehat{\mathbf{z}}^n(\alpha_i^n), \ \widehat{T}^n(\alpha_i^n) \right),$$

$$p_i \geq 0, \ \sum_{i=1}^{m} p_i = 1.$$

We can then use $\{p_i\}_{i=1}^{m}$ to construct the following randomized stationary algorithm: At the start of each frame k, the controller independently chooses action $\alpha_i \in \mathcal{A}^n$ with probability p_i defined above for $i = 1, 2, \cdots, m$. Then, the one-shot expectation of this particular randomized stationary algorithm on system n satisfies

$$\left(\mathbb{E} \left(\sum_{t \in \mathcal{T}_k^n} y^n[t] \right), \ \mathbb{E} \left(\sum_{t \in \mathcal{T}_k^n} \mathbf{z}^n[t] \right), \ \mathbb{E}(T_k^n) \right) = \sum_{i=1}^{m} p_i \cdot \left(\widehat{y}^n(\alpha_i^n), \ \widehat{\mathbf{z}}^n(\alpha_i^n), \ \widehat{T}^n(\alpha_i^n) \right) = (y, \mathbf{z}, T),$$

which implies any point in $\mathcal{S}^n$ can be achieved by the corresponding expectations of a randomized stationary algorithm.

Next, by definition of $\mathcal{P}^n$ in Definition 2.2, any $(\overline{f}^n, \overline{\mathbf{g}}^n) \in \mathcal{P}^n$ can be written as $(\overline{f}^n, \overline{\mathbf{g}}^n) = (y/T, \mathbf{z}/T)$, where $(y, \mathbf{z}, T) \in \mathcal{S}^n$. Thus, it is achievable by the ratio of one-shot expectations from a randomized stationary algorithm, i.e.

$$\frac{\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} y^n[t]\right)}{\mathbb{E}(T_k^n)} = \frac{y}{T} = \overline{f}^n, \quad \frac{\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} \mathbf{z}^n[t]\right)}{\mathbb{E}(T_k^n)} = \frac{\mathbf{z}}{T} = \overline{\mathbf{g}}^n.$$

Now we claim that for $y^n[t]$, $\mathbf{z}^n[t]$ and T_k^n result from the randomized stationary algorithm,

$$(36) \qquad \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}(y^n[t]) = \frac{\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} y^n[t]\right)}{\mathbb{E}(T_k^n)},$$

$$(37) \qquad \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}(\mathbf{z}^n[t]) = \frac{\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} \mathbf{z}^n[t]\right)}{\mathbb{E}(T_k^n)}.$$

We prove (36) and (37) is shown in a similar way. Consider any fixed T, and let $S^n[T]$ be the number of renewals up to (and including) time T. Then, from Lemma 3.7 in Section 3, $S^n[T]$ is a valid stopping time with respect to the filtration $\{\mathcal{F}_k^n\}_{k=0}^\infty$. We write

$$(38) \qquad \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}(y^n[t]) = \frac{1}{T}\mathbb{E}\left(\sum_{k=0}^{S^n[T]} \sum_{t \in \mathcal{T}_k^n} y^n[t]\right) - \frac{1}{T}\mathbb{E}\left(\sum_{t=T}^{t_{S^n[T]}^n - 1} y^n[t]\right).$$

For the first part on the right hand side of (38), since $\left\{\sum_{t \in \mathcal{T}_k^n} y^n[t]\right\}_{k=0}^\infty$ is an i.i.d. process, by Wald's equality (Theorem 4.1.5 of [7]),

$$\frac{1}{T}\mathbb{E}\left(\sum_{k=0}^{S^n[T]} \sum_{t \in \mathcal{T}_k^n} y^n[t]\right) = \mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} y^n[t]\right) \cdot \frac{\mathbb{E}(S^n[T])}{T}.$$

By renewal reward theorem (Theorem 4.4.2 of [7]),

$$\lim_{T \to \infty} \frac{\mathbb{E}(S^n[T])}{T} = \frac{1}{\mathbb{E}(T_k^n)}.$$

Thus,

$$\lim_{T \to \infty} \frac{1}{T}\mathbb{E}\left(\sum_{k=0}^{S^n[T]} \sum_{t \in \mathcal{T}_k^n} y^n[t]\right) = \frac{\mathbb{E}\left(\sum_{t \in \mathcal{T}_k^n} y^n[t]\right)}{\mathbb{E}(T_k^n)}.$$

For the second part on the right hand side of (38), by Assumption 2.2,

$$\left|\mathbb{E}\left(\sum_{t=T}^{t_{S^n[T]}^n - 1} y^n[t]\right)\right| \le y_{\max} \cdot \mathbb{E}\left(t_{S^n[T]}^n - T\right) \le \sqrt{B} y_{\max},$$

which implies $\lim_{T \to \infty} \frac{1}{T}\mathbb{E}\left(\sum_{t=T}^{t_{S^n[T]}^n - 1} y^n[t]\right) = 0$. Overall, we have (36) holds.

To this point, we have shown that for any $(\overline{f}^n, \overline{\mathbf{g}}^n) \in \mathcal{P}^n$, $n \in \{1, 2, \cdots, N\}$, there exists a randomized stationary algorithm so that

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}(y^n[t]) = \overline{f}^n, \quad \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}(\mathbf{z}^n[t]) = \overline{\mathbf{g}}^n,$$

for any $n \in \{1, 2, \cdots, N\}$. Since f^* is the optimal solution to (1)-(2) over all algorithms, it follows for any $(\overline{f}^n, \overline{\mathbf{g}}^n) \in \mathcal{P}^n$, $n \in \{1, 2, \cdots, N\}$ satisfying $\sum_{n=1}^{N} \overline{g}_l^n \leq d_l$, $\forall l \in \{1, 2, \cdots, L\}$, we have $\sum_{n=1}^{N} \overline{f}^n \geq f^*$, and the second part of the lemma is proved. $\qquad \square$

REFERENCES

[1] ALTMAN, E. (1999). *Constrained Markov decision processes.* Chapman and Hall/CRC Press.

[2] BERTSEKAS, D. (2009). *Convex Optimization Theory.* Athena Scientific.

[3] BERTSEKAS, D. P. (2001). *Dynamic Programming and Optimal Control, 2nd edition, Vol. I.* Athena Scientific, Nashua, NH.

[4] BERTSEKAS, D. P. and TSITSIKLIS, J. N. (1997). *Parallel and Distributed Computation: Numerical Methods.* Athena Scientific, Nashua, NH.

[5] BOYD, S. and VANDENBERGHE, L. (2004). *Convex Optimization.* Cambridge University Press.

[6] BYOD, S., GHOSH, A., PRABHAKAR, B. and SHAH, D. (2006). Randomized gossip algorithms. *IEEE/ACM Transactions on Networking* **14,** 2508-2530.

[7] DURRETT, R. (2013). *Probability: Theory and Examples, 4th edition.* Cambridge University Press.

[8] FOX, B. (1966). Markov renewal programming by linear fractional programming. *SIAM Journal on Applied Mathematics* **14,** 1418-1432.

[9] GANDHI, A., DOROUDI, S., HARCHOL-BALTER, M. and SCHELLER-WOLF, A. (2013). Exact analysis of the M/M/k/setup class of Markov chains via recursive renewal reward. *Proc. ACM Sigmetrics* 153-166.

[10] LI, C. and NEELY, M. J. (2014). Solving convex optimization with side constraints in a multi-class queue by adaptive $c\mu$ rule. *Queueing System* **77,** 331-372.

[11] NEELY, M. J. (2010). *Stochastic Network Optimization with Application to Communication and Queueing Systems.* Morgan & Claypool.

[12] NEELY, M. J. (2012). Asynchronous Scheduling for Energy Optimality in Systems with Multiple Servers. *Proceedings of 46th Annual Conference on Information Sciences and Systems (CISS).*

[13] NEELY, M. J. (2012). Asynchronous Control for Coupled Markov Decision Systems. *Information Theory Workshop (ITW).*

[14] NEELY, M. J. (2013). Dynamic Optimization and Learning for Renewal Systems. *IEEE Transactions on Automatic Control* **58,** 32-46.

[15] PENG, Z., XU, Y., YAN, M. and YIN, W. (2016). ARock: an Algorithmic Framework for Asynchronous Parallel Coordinate Updates. *To appear in SIAM Journal on Scientific Computing.*

[16] ROCKAFELLAR, R. T. (2015). *Convex analysis.* Princeton university press.

[17] ROSS, S. (2002). *Introduction to Probability Models, 8th edition.* Academic Press.

[18] SCHAIBLE, S. (1983). Fractional programming. *Zeitschrift fur Operations Research* **27,** 39-54.

[19] SRIVASTAVA, K. and NEDIC, A. (2011). Distributed Asynchronous Constrained Stochastic Optimization. *IEEE Journal of Selected Topics in Signal Processing* **5,** 772-790.

[20] WEI, X. and NEELY, M. J. (2017). Data Center Server Provision: Distributed Asynchronous Control for Coupled Renewal Systems. *IEEE/ACM Transactions on Networking* **25,** 2180 - 2194.

[21] YAO, D. D. (2002). Dynamic Scheduling via Polymatroid Optimization. *Proceeding Performance Evaluation of Complex Systems: Techniques and Tools* 89-113.