arXiv:2506.18278v3 [math.OC] 27 Nov 2025

Finite-Time Minimax Bounds and an Optimal Lyapunov Policy in Queueing Control

Yujie Liu Vincent Y. F. Tan Yunbei Xu
National University of Singapore
{yj-liu, vtan, yunbei}@nus.edu.sg

Abstract

We introduce an original minimax framework for finite-time performance analysis in queueing control and propose a surprisingly simple Lyapunov-based scheduling policy with superior finite-time performance. The framework quantitatively characterizes how the expected total queue length scales with key system parameters, including the capacity of the scheduling set and the variability of arrivals and departures across queues. To our knowledge, this provides the first firm foundation for evaluating and comparing scheduling policies in the finite-time regime, including nonstationary settings, and shows that the proposed policy can provably and empirically outperform classical MaxWeight in finite time. Within this framework, we establish three main sets of results. First, we derive minimax lower bounds on the expected total queue length for parallel-queue scheduling via a novel Brownian coupling argument. Second, we propose a new policy, LyapOpt, which minimizes the full quadratic Lyapunov drift—capturing both first- and second-order terms—and achieves optimal finite-time performance in heavy traffic while retaining classical stability guarantees. Third, we identify a key limitation of the classical MaxWeight policy, which optimizes only the first-order drift: its finite-time performance depends suboptimally on system parameters, leading to substantially larger backlogs in explicitly characterized settings. Together, these results delineate the scope and limitations of classical drift-based scheduling and motivate new queueing-control methods with rigorous finite-time guarantees.

1 Introduction

In this paper, we study the fundamental scheduling problem in parallel queueing systems. We extend the celebrated "generalized switch" model of [43] to nonstationary settings and characterize its finite-time behavior. The system comprises n job flows, each with its own queue, competing for a limited pool of potentially shared, time-varying many-server resources, modeled by general scheduling sets in $\mathbb{R}_+^n$. At each discrete time step, a central controller chooses how to allocate the available service capacity across the n queues. The objective is to optimize system-wide performance measures such as total queue length, delay, and throughput.

Scheduling plays a central role in modern infrastructures. It underpins cloud and communication platforms such as data centers, wireless networks, and service systems, while also serving as a

cornerstone of AI infrastructure by enabling throughput-optimized schedulers for GPU clusters that support large-scale model training and inference [3, 26]. Whereas classical analyses focus on long-run performance under stationary environment, contemporary platforms routinely encounter bursty, nonstationary workloads and stringent service-level objectives, as exemplified by production systems including NVIDIA's Run:AI and Google's Borg [13, 47, 39]. In these regimes, system performance is determined by transient congestion and workload variability, highlighting the need for rigorous finite-time guarantees. Developing scheduling policies with provable delay bounds in these nonstationary, time-critical regimes is therefore essential to ensure responsiveness and resource efficiency that next-generation systems require.

The celebrated MaxWeight scheduling policy [45, 43] has been widely studied in the queueing literature. It is known to be throughput-optimal [45] and to achieve optimality in the diffusion-scaled, heavy-traffic regime, as established through elegant Lyapunov drift analysis and diffusion approximations [43, 33, 9, 10]. However, in finite-time regime, its performance remains poorly understood, leaving a crucial gap in both theory and practice.

Precisely analyzing the finite-time queueing performance of scheduling policies for individual instances is analytically intractable. To address this challenge, we introduce a minimax framework, akin to the classical minimax-risk formulation in statistical decision theory pioneered by Wald [49] and developed in modern texts [25, 48, 24, 7], which evaluates non-asymptotic performance over model classes rather than a single instance, enabling a unified finite-time analysis. This perspective is distinct from the instance- or problem-dependent approaches of prior works [45, 43]. Our framework provides finite-time upper and lower bounds that hold uniformly and optimally over well-defined classes of instances, up to universal constants. These results offering a principled basis for queueing control with explicit, parameter-dependent optimality guarantees at finite horizons.

In this work, we ask the following central questions:

- How can we formulate a minimax framework for the scheduling problem in queueing systems? Within this framework, what is the minimum achievable queue length by time T?
- Can MaxWeight attain this minimum? If not, what alternative scheduling policies can possibly achieve it, and under what conditions?

1.1 Contributions

We provide rigorous and definitive answers to our central questions:

1. *Proposal of a Minimax Framework For Queuing.* We develop what is, to our knowledge, the first minimax framework for scheduling problems in queueing systems (Section 2.2). We focus on finite-time performance metrics such as the expected total queue length and the expected sum of squared queue lengths. The model class comprises capacity-constrained scheduling sets that may vary over time, nonstationary arrival processes subject to mean and variance constraints, and departure processes with bounded variance, thereby capturing a broad range of uncertainties in queueing systems. The minimax objective is to find a scheduling policy that minimizes

the worst-case expected total queue length over all instances in this class. This viewpoint sidesteps classical instance-specific assumptions, enables explicit finite-time analysis (including nonstationary settings). It also complements classical asymptotic tools such as fluid and diffusion approximations by explicitly revealing how the queue length scales with the underlying system parameters in the finite-time regime, a level of detail that asymptotic methods inherently obscure.

2. *Derivation of Finite-Time Lower Bounds.* Within the minimax framework, we derive the first finite-time minimax lower bound for any scheduling policy in parallel queueing systems (Section 3). In the heavy-traffic regime, the bound grows on the order of $\Omega(\sqrt{T})$ over a horizon T, while in the interior regime it scales as $\Omega((1-\rho)^{-1})$, where ρ denotes the traffic intensity; other system parameters characterized in the bounds include the number of queues, the arrival variance, the departure variance, and the capacity parameters of the scheduling sets. This lower bound thus provides a fundamental benchmark for all scheduling policies. Its proof leverages classical minimax lower-bound theory to construct hard yet analytically tractable instances within the model class, and then analyzes the resulting stochastic queueing process through an innovative Brownian-coupling argument from applied probability. This novel, blended proof strategy may have broader applicability for establishing lower bounds in other dynamic-programming–type control problems.

3. *Design and Analysis of an Optimal Lyapunov Policy.* Building on the finite-time minimax framework, we design and analyze LyapOpt, a novel scheduling policy that minimizes the *full* Lyapunov drift, explicitly incorporating both the first- and second-order terms of the quadratic Lyapunov function (Section 4.1). We show that LyapOpt attains the minimax lower bound up to universal constants under certain structured conditions, thereby establishing its minimax optimality in finite time. By including the second-order term, LyapOpt *adapts* the scheduling decision to the *geometry* of the queue length vector, rather than defaulting to extreme points as in traditional drift-based methods like MaxWeight. This enhancement improves the finite-time performance, particularly in scenarios with asymmetric scheduling sets—for example, when servers have significantly different service rates. Moreover, LyapOpt retains the classical stability guarantees proven for MaxWeight whenever the traffic intensity is strictly less than 1 (Section 5), since it directly minimizes the one-step Lyapunov drift and the standard stability arguments for MaxWeight carry over verbatim.

4. *Identification of a Key Limitation of MaxWeight.* Within the finite-time minimax framework, we show that MaxWeight is not minimax optimal (Section 4.2). In particular, its expected total queue length exceeds the minimax lower bound by a factor that depends explicitly on the capacity of the scheduling set. In highly asymmetric scheduling sets, even under deterministic arrivals and departures, MaxWeight exhibits a $\sqrt{T}$ growth for some range of the finite time T, whereas LyapOpt maintains a constant total queue length independent of T (Section 4.2). This finding does not contradict MaxWeight's classical optimality in the diffusion-scaled, heavy-traffic regime; rather, it underscores that traditional drift-based methods inevitably overlook finer geometric

structures which become especially important when considering the more practically relevant finite-time regime.

Bridging Queueing Control and Non-Asymptotic Learning. Dynamic programming provides a principled framework for optimal control, but exact solutions are often computationally infeasible in high-dimensional systems. As a result, performance is typically studied using asymptotic methods such as fluid and diffusion approximations [8] or approximate optimality bounds based on bias and performance-difference analyses [4], which offer only limited insight into behavior over finite time horizons. As queueing theory is one of the most widely studied domains in dynamic programming, our work provides a representative minimax framework for finite-time analysis of structured control problems. In doing so, it strengthens the theoretical foundations of both learning theory (with its non-asymptotic, lower-bound viewpoint) and dynamic programming (by bringing queueing-based tools into finite-time analysis), and opens new directions for short-horizon decision making.

Not just regret: Unlike conventional analyses of reinforcement learning algorithms (e.g., [6]), which benchmark the performance of proposed methods against an *oracle* dynamic programming policy, our study examines the intrinsic performance of the policy itself. In this context, our findings complement existing learning frameworks centered on parameter estimation [25] by emphasizing the fundamental limitations of the control policies that these algorithms seek to approximate.

1.2 Related Works

Stability and Diffusion Optimality of MaxWeight. The MaxWeight policy (also referred to as the backpressure policy in multi-hop queueing networks) was introduced by Tassiulas and Ephremides [45] and has since been extensively studied (see, e.g., [2, 44, 43, 33, 9, 10]). Much of this literature has established the policy's throughput optimality across a wide range of queueing models, demonstrating that MaxWeight stabilizes the queue length process whenever stability is achievable [45, 44, 2, 9]. Beyond throughput guarantees, another stream of research has shown that the MaxWeight policy minimizes the diffusion-scaled workload process among all scheduling policies in the heavy-traffic regime, both in parallel queueing systems [43, 33] and in fairly general stochastic processing networks [10]. These diffusion-scaled analyses are a major step toward understanding the asymptotic efficiency of MaxWeight and its connection to reflected Brownian motion models. However, because they focus on diffusion-scaled dynamics, the resulting characterizations capture only coarse parameter dependence in finite time—typically through first- and second-order moments—while other key quantities (such as the capacity parameters of the scheduling sets) do not appear explicitly. As a result, such asymptotic approximations cannot reveal how performance depends on these system parameters over a fixed, finite time horizon.

Existing Finite-Time Analysis in Queueing. Understanding queueing behavior in the finite-time regime remains challenging, even for basic models such as the $M/M/1$ queue [1]. Much of the existing literature focuses on uncontrolled systems, examining the convergence rate to steady

state—typically via coupling arguments for the $M/M/1$ queue [41] or spectral methods for the $GI/M/n$ queue in the heavy traffic regime [16, 46]. Finite-time analyses for controlled systems is far less developed. There are a few results for routing policies such as Join-the-Shortest-Queue (see, e.g., [28, 29]). Luczak and McDiarmid [28] use mean-field and coupling techniques to characterize the convergence rate of the maximum queue length to equilibrium in large-scale systems, while Ma and Maguluri [29] establish the convergence rate for the two-server symmetric case based on the original system rather than its diffusion-scaled counterpart. In contrast, our framework directly analyzes finite-time performance in the original system, without relying on diffusion-scaled regimes or making any reference to steady-state behavior.

Parameter Learning in Queueing. Some strands of online learning research in queueing systems study scheduling problems in which either service statistics [22, 23, 52, 42, 15, 27, 11] or the utility function [11, 21] are initially unknown and must be learned online. A significant body of work establishes upper bounds on the time-averaged queue length of learning algorithms within the finite-time regime [52, 11], with aims such as ensuring system stability or evaluating utility maximization relative to benchmark policies [11]. Another common performance metric in the literature is the queue length regret, which quantifies the excess backlog incurred by a learning algorithm relative to an oracle policy such as MaxWeight or $c\mu$ (see, e.g., [22, 23, 42, 15, 27]). Despite these advances, little is known about the optimal policies in the finite-time regime when system parameters are already known—beyond the classical result that the $c\mu$ rule is optimal in the single-server case [20]. To address this gap in the literature, we go beyond the parameter learning paradigm and propose a minimax formulation with the expected total queue length as the performance metric, thereby elucidating the fundamental finite-time limits of scheduling policies and identifying a minimax-optimal strategy.

Finite-Time Performance of Dynamic Programming Policies. Queueing control problems can also be modeled within the framework of structured dynamic programming. However, the exponential growth of the state and action spaces makes obtaining exact solutions computationally infeasible—an illustration of the well-known curse of dimensionality [38, 36]. These challenges highlight the crucial need to establish fundamental performance limits for any admissible policy. Despite their importance, few results have addressed this issue; a notable exception is Gupta and Shroff [17], who derived lower bounds on delay in multi-hop networks by reducing the problem to the delay of a $G/D/1$ queue. Nonetheless, such reductions remain difficult to analyze in the finite-time regime. In contrast, our work provides finite-time lower bounds that explicitly reveal the dependence on underlying system parameters, offering key benchmarks for evaluating scheduling policies in general and challenging settings.

Lyapunov Drift Methods. Lyapunov drift methods are widely used to establish stability and characterize steady-state behavior in queueing networks [14, 30]. Many control policies can, in fact, be interpreted as solutions that optimize the one-step Lyapunov drift. The MaxWeight policy directly

minimizes the first-order term of the one-step quadratic Lyapunov drift, whereas the drift-plus-penalty policy [37] minimizes the same term augmented with auxiliary objectives such as delay or energy consumption, thereby ensuring stability while simultaneously optimizing a desired performance criterion. Although such policies guarantee stability, their finite-time performance remains largely unexplored. The proposed LyapOpt policy minimizes the full Lyapunov drift, incorporating both the first- and second-order terms, and thereby directly targets performance in the finite-time regime. This refinement highlights that optimizing only the first-order term, as done in traditional Lyapunov drift-based methods, is too coarse to achieve finite-time optimality.

The preceding discussion reviewed related works from several perspectives. After introducing our model and the minimax framework in Section 2, we present a comprehensive and detailed comparison of our approach with the most closely related studies in Table 2 in Section 2.2.

1.3 Outline of Paper

The remainder of the paper is organized as follows. Section 2 presents the model setup and establishes the minimax framework. Section 3 develops the minimax finite-time lower bound. In Section 4, we propose the LyapOpt policy, establish its minimax optimality in the finite-time regime, and characterize a limitation of the MaxWeight policy. Section 5 further proves the stability of LyapOpt. Section 6 presents numerical experiments comparing LyapOpt with MaxWeight. Section 7 extends the results by deriving the minimax finite-time lower bound under specific conditions and considering additional model classes. Finally, Section 8 concludes the paper and discusses directions for future research.

2 Problem Setup and Minimax Framework

2.1 Problem Setup of Queueing Control

Notation. Let $\mathbb{N} = \{1, 2, \ldots\}$ be the set of natural numbers, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x_i \geq 0,\ 1 \leq i \leq n\}$. Let $\mathbb{Z}^n_+ := \{x \in \mathbb{Z}^n : x_i \geq 0,\ 1 \leq i \leq n\}$ be the set of integer-valued vectors. For any set $U \subseteq \mathbb{R}^n_+$, let $\mathrm{conv}(U)$ denote its convex hull, and define $\Pi(U) = \big\{\gamma \in \mathbb{R}^n_+ : \exists\, u \in \mathrm{conv}(U) \text{ with } \gamma \leq u\big\}$. That is, $\Pi(U)$ consists of all nonnegative vectors that are componentwise bounded above by some convex combination of elements of U. We use $\mathbf{0}$ to denote the zero vector, with the dimension understood from context. For $x, y \in \mathbb{R}^n$, the componentwise maximum and minimum are defined by $(\max\{x, y\})_i = \max\{x_i, y_i\}$ and $(\min\{x, y\})_i = \min\{x_i, y_i\}$ for $1 \leq i \leq n$. $\langle x, y\rangle$ denotes the inner product of x and y. Independence of random elements X and Y is written as $X \perp\!\!\!\perp Y$, while conditional independence of X and Y given a σ-algebra $\mathcal{G}$ is written $X \perp\!\!\!\perp Y \mid \mathcal{G}$. We adopt the convention $\frac{c}{0} = +\infty$ for any constant $c > 0$.

Probability space. We work in a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Expectation and variance under $\mathbb{P}$ are denoted as $\mathbb{E}[\cdot]$ and $\mathrm{Var}(\cdot)$, respectively, and "a.s." is a shorthand for almost surely. A

stochastic process $X : \mathbb{N}_0 \times \Omega \to E$ is denoted as $\{X(t)\}_{t\in\mathbb{N}_0}$, where $X(t) = X(t,\omega)$ refers to the random variable at time $t \in \mathbb{N}_0$.

Queueing Model and Dynamics. We consider a discrete-time queueing system consisting of n parallel, infinite-capacity queues, where jobs arrive exogenously and depart after a single service. The queue length process is $Q : \mathbb{N}_0 \times \Omega \to \mathbb{R}^n_+$ with $Q(0) = \mathbf{0}$ almost surely. The arrival process is $A : \mathbb{N} \times \Omega \to \mathbb{R}^n_+$, where $A_i(t)$ is the number of jobs arriving to queue i for $1 \le i \le n$. We consider the scheduling sequence (also referred to as a *switch*) $\mathcal{D} := \{\mathcal{D}_t\}_{t\in\mathbb{N}_0}$, where each scheduling set $\mathcal{D}_t \subset \mathbb{R}^n_+$ is compact. An element of $\mathcal{D}_t$, termed a *schedule*, specifies the expected number of jobs that can depart from each queue at time t. At each time t, the controller chooses $d(t) \in \mathcal{D}_t$ according to a (possibly randomized) policy, and the realized departure vector $D(t)$ is a random realization consistent with $d(t)$, as formalized in (2) in the following. The queue lengths then evolve almost surely as

$$\begin{aligned} Q(0) &= \mathbf{0}; \\ Q(t+1) &= \max\{Q(t) - D(t), \mathbf{0}\} + A(t+1), \quad \text{for } t \in \mathbb{N}. \end{aligned} \tag{1}$$

Note that instead of specifying servers explicitly as in [45], we represent service through time-indexed scheduling sets $\{\mathcal{D}_t\}_{t\in\mathbb{N}_0}$, where each $\mathcal{D}_t \subset \mathbb{R}^n_+$ denotes the set of all feasible instantaneous expected departure vectors at time t. This abstraction generalizes classical multi-server models. Two illustrative examples are shown in Table 1. The number in the row for queue i and the column for server j specifies the service rate of server j for queue i (a zero indicates that server j cannot serve queue i). Each server may serve at most one queue per slot. For the single-server case (Table 1a), the instantaneous scheduling set is $\mathcal{D}_1 = \{(1,0), (0,2)\}$. For the two-server case (Table 1b), the instantaneous scheduling set is $\mathcal{D}_2 = \{(1,0), (0,1.5), (2,0), (1,1.5)\}$. Figure 1 depicts the sets $\mathcal{D}_1$ and $\mathcal{D}_2$ together with their associated regions $\Pi(\mathcal{D}_1)$ and $\Pi(\mathcal{D}_2)$.

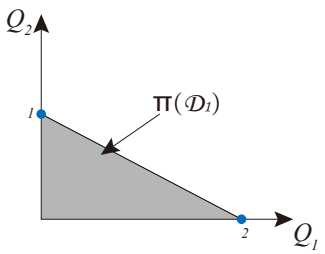
(a) Single server

	Server 1
Queue 1	1
Queue 2	2

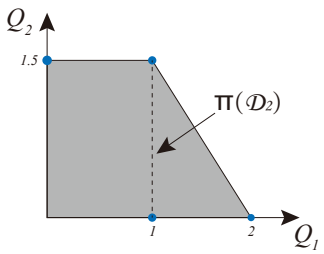
(b) Two servers

	Server 1	Server 2
Queue 1	1	1
Queue 2	0	1.5

Table 1: Service tables



(a) Scheduling set $\mathcal{D}_1$ and $\Pi(\mathcal{D}_1)$ induced by Table 1a



(b) Scheduling set $\mathcal{D}_2$ and $\Pi(\mathcal{D}_2)$ induced by Table 1b

Figure 1: Scheduling sets induced by the service matrices in Table 1.

Arrival Process. In our model, we allow the arrivals to be nonstationary. That is, the distribution of $A(t)$ may vary with time t, while the sequence $\{A(t)\}_{t\in\mathbb{N}}$ remains independent across t. Moreover, for each fixed t, the components $\{A_i(t)\}_{i=1}^n$ may exhibit arbitrary dependence across queues indexed by $1 \le i \le n$. Let $\lambda(t) := \mathbb{E}[A(t)] \in \mathbb{R}_+^n$ be the arrival rate vector at time t.

Policy. Let the history (i.e., the information available) prior to choosing $d(t)$ at time t be

$$\mathcal{F}_t := \sigma\{\mathcal{D}_0, D(0), A(1), \cdots, \mathcal{D}_{t-1}, D(t-1), A(t), \mathcal{D}_t\}.$$

A *policy* Φ is defined as a sequence of scheduling rules, $\Phi = \{\phi_t\}_{t\in\mathbb{N}_0}$. Each ϕ_t specifies, based on the history $\mathcal{F}_t$, a probability distribution over the scheduling set $\mathcal{D}_t$. Formally, given $\mathcal{F}_t$, the schedule $d(t) \in \mathcal{D}_t$ is sampled according to

$$\mathbb{P}\big(d(t) \in \cdot \,\big|\, \mathcal{F}_t\big) = \phi_t(\cdot).$$

Departure Mechanism and Capacity Region. The departure variability is modeled by a *random departure field* $S = \{S_t\}_{t\in\mathbb{N}_0}$, where

$$S_t : \mathcal{D}_t \times \Omega \to \mathbb{R}_+^n, \qquad (d, \omega) \mapsto S(d, \omega),$$

is jointly measurable with the natural product σ-algebra. Write $S_d(t) := S_t(d, \omega)$. For each $t \in \mathbb{N}_0$, the σ-algebra generated by $\{S_d(t) : d \in \mathcal{D}_t\}$ is independent of $\mathcal{F}_t$ conditioned on $\mathcal{D}_t$. Moreover, the random departure field S is assumed to be independent of the arrival process A. The field satisfies *mean consistency*:

$$\mathbb{E}[S_d(t)] = d \quad \text{a.s.} \quad \text{for all } t \in \mathbb{N}_0,\ d \in \mathcal{D}_t.$$

At time t, once the policy selects $d(t) \in \mathcal{D}_t$, the departure vector is realized by

$$D(t) := S_{d(t)}(t). \tag{2}$$

Consequently,

$$D(t) \perp\!\!\!\perp \mathcal{F}_t \mid d(t), \qquad \mathbb{E}[D(t) \mid \mathcal{F}_t,\, d(t)] = d(t) \quad \text{a.s.}$$

Our goal is to minimize the queue length in the finite-time regime. To achieve this, we develop a minimax framework that captures the fundamental finite-time performance limits of the scheduling problem and provides a benchmark against which all scheduling policies can be evaluated.

2.2 Minimax Criteria

The minimax criterion is a standard approach to studying the intrinsic difficulty of problems in statistics and machine learning [49]. In this work, we extend this criterion to the domain of queueing control, aiming to bridge queueing theory with learning-theoretic methodologies and to motivate the use of statistical decision-theoretic tools in a broader class of dynamic programming problems.

Performance Metrics. We consider two standard performance metrics: the total queue length (see Neely [37]), which captures the overall system backlog, and the sum of squares of queue lengths (see Stolyar [43]), which reflects the degree of imbalance across queues. These are formally defined as

$$\mathbb{E}\left[\sum_{i=1}^{n} Q_i(T)\right] \quad \text{and} \quad \mathbb{E}\left[\sum_{i=1}^{n} Q_i(T)^2\right],$$

respectively. In this work, we primarily present results for the total queue length, while the corresponding matching bounds for the sum of squares objective are provided in Appendix A.1. Note that this is a dynamic programming problem, where the objective at time t is a function of all past decisions $d(0), d(1), \ldots, d(t-1)$. A central challenge in dynamic programming and control is to develop solutions that are both tractable and practically effective.

Model Class. The queueing system under consideration is specified by a scheduling sequence $\mathcal{D}$, an arrival process A, and a random departure field S. Adopting a minimax perspective, we define the following model class that includes capacity-constrained scheduling sequence, arrival processes subject to mean and variance constraints, and departure processes with bounded variance.

$$\mathcal{M}^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}}) = \left\{ (\mathcal{D}, A, S) \;:\; \begin{array}{c} \frac{1}{n}\sum_{i=1}^{n} d_i^2 \le B^2, \;\; \forall d \in \mathcal{D}_t, t \in \mathbb{N}_0; \\ \lambda(t) \in \rho\Pi(\mathcal{D}_t),\; \frac{1}{n}\sum_{i=1}^{n} \mathrm{Var}(A_i(t)) \le C_{\mathrm{a}}^2, \;\; \forall\, t \in \mathbb{N}; \\ \frac{1}{n}\sum_{i=1}^{n} \mathrm{Var}((S_d(t))_i) \le C_{\mathrm{d}}^2, \;\; \forall\, d \in \mathcal{D}_t, t \in \mathbb{N}_0 \end{array} \right\}$$

where the scheduling sequence $\mathcal{D}$, the arrival process A, and the random departure field S are those defined in Section 2.1. The parameter $\rho \in (0, 1]$ denotes the *traffic intensity*; $B > 0$ constrains the capacity of the scheduling set; and $C_{\mathrm{a}} \ge 0$ and $C_{\mathrm{d}} \ge 0$ respectively control the variability of the arrivals and departures across queues. We refer to $\rho \uparrow 1$ as the *heavy-traffic regime* and $\rho \in (0, 1)$ as the *interior* (of the capacity region) regime.

Similar constraints of the form $\lambda(t) \in \rho\Pi(\mathcal{D}_t)$ have been studied in prior works (e.g., [45, 43, 37]). Tassiulas and Ephremides [45] considered time-invariant scheduling sets and showed that stability requires arrival rates to be dominated by the convex hull of feasible schedules. Stolyar [43] and Neely [37] modeled time-varying scheduling sets as finite-state Markov chains and required arrival rates to be dominated by schedules averaged under the stationary distribution. In contrast, we allow the scheduling sequence to be time-varying and impose this domination constraint at each time step.

We aim to characterize the fundamental minimax expected total queue length at time T:

$$\Gamma^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}}, T) := \inf_{\Phi} \sup_{(\mathcal{D},A,S)\in\mathcal{M}^{\rho}(B,C_{\mathrm{a}},C_{\mathrm{d}})} \mathbb{E}^{\Phi}_{(A,S)} \left[\sum_{i=1}^{n} Q_i(T) \right]. \tag{3}$$

This work introduces, for the first time, a minimax formulation for characterizing the finite-time fundamental limits of queueing control, offering a principled approach to quantifying the hardness of structured dynamic scheduling problems.

Table 2 provides a structured comparison of most related works on scheduling in queueing systems. "Resource Type" specifies the fundamental scheduling unit, such as individual servers or the more abstract scheduling set $\mathcal{D}_t$. "Arrival Pattern" characterizes the arrival process, represented by $A = \{A(t)\}_{t\in\mathbb{N}_0}$, distinguishing between stationary arrivals, which are i.i.d. over time, and nonstationary arrivals, which may vary arbitrarily over time. "Unknown Parameters" lists the system characteristics that are not assumed to be known a priori and must be inferred. Finally, "Performance Metric" specifies the criterion used to evaluate policies; for our setting, this is defined precisely in (3).

	Resource Type	Arrival Pattern	Unknown Parameters	Performance Metric
Krishnasamy et al. [23]	Server	Stationary	Arrival rate Service rate	Queue length regret
Yang et al. [52]	Server	Nonstationary	Arrival rate Service rate	Time-averaged queue length
Stolyar [43]	Scheduling set	Stationary	Arrival rate	Workload in diffusion-scaled, heavy-traffic regime
Ours	Scheduling set	Nonstationary	Arrival rate	Finite-time queue length

Table 2: Comparison of most related works on scheduling in queueing systems.

As shown in Table 2, our framework represents a departure from traditional online learning approaches which typically focus on estimating unknown service rates to ensure stability. Additionally, our framework extends beyond classical diffusion-scaled optimality results. It addresses the *finite-time* scheduling problem and aims to identify the minimax-optimal policy in a nonstationary environment.

3 Minimax Lower Bound

In this section, we derive a finite-time minimax lower bound on the expected total queue length for the system introduced in Section 2. Formally, we seek a lower bound for the quantity (3). To this end, we employ a minimax approach together with Brownian coupling techniques to construct hard yet analytically tractable instances within the model class and analyze their behavior. The resulting bound explicitly quantifies how queue lengths scale with the time horizon T, the traffic intensity ρ, the capacity parameter B imposed by the scheduling sequences, and the variance parameters C_{a} and C_{d} on the arrival processes and random departure fields.

Intuitively, this lower bound reveals the challenging scenarios an adversary can construct under the prescribed capacity and variance constraints, thereby establishing fundamental performance limits for any scheduling policy in parallel queueing systems.

3.1 The Bound and Its Interpretation

The following theorem formalizes this lower bound for arbitrary scheduling policies within the model class $\mathcal{M}^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}})$.

Theorem 1 (Minimax Lower Bound)**.** *For any scheduling policy, and for scheduling sequences, arrival processes and random departure fields within the model class $\mathcal{M}^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}})$, there exist an absolute constant $c_1 > 0$ and a finite constant $\psi_1(n, B, C_{\mathrm{a}}, C_{\mathrm{d}}) > 0$ such that, for*

$$\rho \in \begin{cases} \left[\left(1 - \frac{\sqrt{C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2}}{B\psi_1(n,B,C_{\mathrm{a}},C_{\mathrm{d}})}\right) \vee 1/2, 1 \right], & \text{if } C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2 > 0, \\ (0,1], & \text{if } C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2 = 0, \end{cases}$$

the following lower bound holds for all $T \geq \psi_1(n, B, C_{\mathrm{a}}, C_{\mathrm{d}})^2 \mathbb{1}_{\{C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2 > 0\}} + 1$,

$$\Gamma^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}}, T) \geq c_1 \min\left\{ n\sqrt{(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)(T-1)}, \frac{n(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)}{B(1-\rho)} \right\} + n\rho B. \tag{4}$$

Theorem 1 demonstrates that, for any scheduling policy and any system within the model class $\mathcal{M}^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}})$, the expected total queue length scales at least on the order of $\Omega\big(\min\{\sqrt{T}, (1-\rho)^{-1}\}\big)$, up to problem-dependent constants.

The first term inside the min in (4) is active in the heavy traffic regime where $\rho \uparrow 1$. It reflects the intrinsic stochastic fluctuations over a finite horizon. When ρ assumes its extremal value of 1, the lower bound scales as $n\sqrt{(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)T}$, and this scaling persists for all T.

The second term in the min becomes dominant in the interior regime, when ρ is strictly less than 1 and T is sufficiently large. It captures the long-term congestion effect and depends explicitly on the traffic intensity parameter. As ρ approaches 1, the $\sqrt{T}$ term dominates for sufficiently large but finite horizons, indicating that the lower bound grows with a rate of at least $n\sqrt{(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)T}$. Beyond this point, the lower bound effectively plateaus, with a value determined by the system

parameters $B, C_{\mathrm{a}}, C_{\mathrm{d}}$, and the traffic intensity ρ. This dependence of the queue length on the reciprocal of $1-\rho$ in the interior regime has also been observed in previous work [14].

Overall, the two terms in the lower bound together establish a fundamental performance limit that no scheduling policy can surpass.

3.2 Proof Sketch of Theorem 1

As the proof of Theorem 1 contains some interesting and novel elements, we sketch its main arguments in the following.

Proof Sketch of Theorem 1. The proof proceeds in several conceptual steps, reducing the queueing dynamics to a drifted random walk and then approximating it via Brownian motion. The complete details are provided in Appendix A.3.2.

(1) Lower bound: reducing vector processes to total sums. From the lower bound on the total queue length (as stated in Lemma 4), we have

$$\sum_{i=1}^{n} Q_i(T) \geq \max_{1\leq k\leq T-1} \sum_{t=k}^{T-1} \left(\sum_{i=1}^{n} A_i(t) - \sum_{i=1}^{n} D_i(t) \right) + \sum_{i=1}^{n} A_i(T). \tag{5}$$

In other words, the total queue length is at least the maximum of cumulative net input (total arrivals minus total departures) plus the final arrivals. Hence, bounding $\mathbb{E}[\sum_{i=1}^{n} Q_i(T)]$ from below reduces to bounding the expected maximum of the total net-input process.

(2) Hard-instance construction: lower-bounding the net input by a drifted random walk. Consider a time-invariant scheduling set $\mathcal{D}_\star$ with total capacity M, i.e., $\max_{d\in\mathcal{D}_\star} \sum_{i=1}^{n} d_i = M$, where M depends on the parameter B. We construct i.i.d. arrivals $\{A(t)\}_{t\in\mathbb{N}}$ over time, where all components $A_i(t)$ are correlated across i, such that the aggregate mean satisfies $\mathbb{E}[X(t)] = \rho M$ and the variance satisfies $\mathrm{Var}(X(t)) = n^2 C_{\mathrm{a}}^2$, with $X(t) = \sum_{i=1}^{n} A_i(t)$. We then introduce an i.i.d. total-departure proxy $\{G(t)\}_{t\in\mathbb{N}}$ that dominates any policy's total departures, with mean M and variance $n^2 C_{\mathrm{d}}^2$. Replacing true total departures by $G(t)$ yields a valid lower bound in (5). The resulting net-input sequence $\{X(t) - G(t)\}$ is i.i.d., and the problem reduces to bound the expected maximum of a drifted random walk with negative drift $-\theta$, where $\theta = (1-\rho)M$ and variance $\sigma^2 = n^2(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)$.

(3) Probabilistic lower bound via Brownian coupling. For a negatively drifted Brownian motion $Y_t = \sigma W_t - \theta t$, Lemma 1 gives

$$\mathbb{E}\left[\sup_{0\leq s\leq t} Y_s\right] \gtrsim \min\left\{\sigma\sqrt{t},\ \sigma^2/\theta\right\}. \tag{6}$$

By the Brownian coupling in Lemma 2, the partial sum of k i.i.d. zero-mean increments with variance σ^2 can be coupled with σW_k so that their maximum difference in the interval $[0, t]$ is $O(\log t)$ in expectation. Consequently, as formalized in Lemma 3, the drifted random walk $S(k) = \sum_{t=1}^{k}(X(t) - G(t))$ inherits the same lower bound scaling as in (6) (up to constants).

Specifically, for all $t \geq \psi^2$ (a finite constant determined by the characteristics of $X(t) - G(t)$) and for all θ satisfying $0 \leq \theta \leq \sigma/\psi$, the expected maximum satisfies

$$\mathbb{E}\left[\max_{0\leq k\leq t} S(k)\right] \gtrsim \min\left\{\sigma\sqrt{t}, \sigma^2/\theta\right\}.$$

(4) Application to the queueing instance. For the rectangular scheduling set $\mathcal{D}_\star = \{x \in \mathbb{R}^n : 0 \leq x_i \leq B \text{ for } 1 \leq i \leq n\}$, we have $M = nB$. Applying the above bound to the constructed instance yields the desired lower bound in (4). □

In summary, the proof leverages a classical minimax approach by constructing hard instances with arrival rates close to their maximum admissible limits and with significant fluctuations in both arrivals and departures. Unlike traditional minimax arguments that focus on static parameter estimation, our challenge is to control the evolution of a stochastic queueing process. We demonstrate that the expected total queue length is lower bounded by the expected maximum of a random walk with i.i.d. negative-mean increments. To analyze this, we introduce innovative Brownian coupling techniques, coupling the random walk with a drifted Brownian motion. This coupling introduces a small logarithmic error, which is negligible compared to the scaling of the Brownian motion. We surmise that this proof strategy may be broadly applicable to establishing lower bounds for other dynamic programming-type problems.

Theorem 1 reveals that no policy can guarantee a better scaling than $n\sqrt{(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)T}$ in the finite-time setting. This establishes a fundamental lower bound, serving as a benchmark for evaluating the performance of any scheduling algorithm. Building on this perspective, the next section introduces a policy framework that attains this bound up to universal constants by jointly optimizing the first- and second-order Lyapunov drift, thereby illustrating how the proposed approach bridges the gap in traditional scheduling methods.

4 Finite-Time Performance Guarantees

4.1 Optimal Lyapunov Policy

Building on the finite-time minimax framework introduced earlier, we develop a novel scheduling policy that optimizes the full one-step derandomized Lyapunov drift at each decision time. By "Lyapunov drift", we refer to the one-step difference of the Lyapunov function. Unlike traditional drift-based methods that optimize only the first-order term of this drift, our policy anticipates future queue dynamics and directly controls them through an additional second-order objective, yielding more balanced and stable queue evolution.

Recall that $Q(t) \in \mathbb{R}_+^n$ denotes the queue-length vector at time t. The LyapOpt policy selects

the schedule $d(t)$ as the solution to the following optimization problem:

$$d^{\text{LyapOpt}}(t) \in \operatorname*{argmin}_{d \in \mathcal{D}_t} \sum_{i=1}^{n} \left(\max\{Q_i(t) - d_i, 0\}\right)^2, \tag{7}$$

with ties broken arbitrarily. This expression represents a one-step derandomized lookahead of the quadratic Lyapunov function $V(x) = \|x\|_2^2$, capturing the anticipated impact of the chosen schedule $d \in \mathcal{D}_t$. Specifically, based on the queue dynamics in (1), the one-step Lyapunov drift is given by

$$\begin{aligned}\Delta V(t) &= \mathbb{E}\left[V\big(Q(t+1) - A(t+1)\big) - V\big(Q(t) - A(t)\big) \mid \mathcal{F}_t\right] \\ &= \mathbb{E}\left[\sum_{i=1}^{n} (\max\{Q_i(t) - D_i(t), 0\})^2 \middle| \mathcal{F}_t\right] - \sum_{i=1}^{n} (Q_i(t) - A_i(t))^2. \end{aligned} \tag{8}$$

Replacing the random departure $D(t)$ with its conditional mean $d(t)$ yields a deterministic surrogate objective in (7). As shown in Appendix A.4.1, this replacement only incurs an additive variance correction term:

$$\mathbb{E}\left[\sum_{i=1}^{n} (\max\{Q_i(t) - D_i(t), 0\})^2 \middle| \mathcal{F}_t\right] \le \mathbb{E}\left[\sum_{i=1}^{n} (\max\{Q_i(t) - d_i(t), 0\})^2 \middle| \mathcal{F}_t\right] + \sum_{i=1}^{n} \operatorname{Var}(D_i(t)). \tag{9}$$

Before presenting the main performance guarantee of this policy, we first develop a general Lyapunov drift analysis applicable to any scheduling policy.

In the following, for notational simplicity and without risk of ambiguity, we write $\mathbb{E}^{\Phi}_{(A,S)}$ as $\mathbb{E}$. Then, for any scheduling sequence, arrival processes and random departure fields in $\mathcal{M}^{\rho}(B, C_{\text{a}}, C_{\text{d}})$, and any scheduling policy Φ, consider the one-step Lyapunov drift as follows:

$$\begin{aligned}\Delta V(t) &= \mathbb{E}\left[\sum_{i=1}^{n} (\max\{Q_i(t) - D_i(t), 0\})^2 \middle| \mathcal{F}_t\right] - \sum_{i=1}^{n} (Q_i(t) - A_i(t))^2 \\ &\le \mathbb{E}\left[\sum_{i=1}^{n} (\max\{Q_i(t) - d_i(t), 0\})^2 \middle| \mathcal{F}_t\right] + \sum_{i=1}^{n} \operatorname{Var}(D_i(t)) - \sum_{i=1}^{n} (Q_i(t) - A_i(t))^2 \\ &\le \mathbb{E}\left[\sum_{i=1}^{n} (Q_i(t) - d_i(t))^2 \middle| \mathcal{F}_t\right] - \sum_{i=1}^{n} (Q_i(t) - A_i(t))^2 + \sum_{i=1}^{n} \operatorname{Var}(D_i(t)) \\ &= f(Q(t), d(t)) + r(Q(t), A(t)) + \sum_{i=1}^{n} \operatorname{Var}(D_i(t)), \end{aligned} \tag{10}$$

where

$$f(Q(t), d) := \underbrace{\mathbb{E}\left[2 \sum_{i=1}^{n} Q_i(t)\big(\lambda_i(t) - d_i\big) \middle| \mathcal{F}_t\right]}_{\text{first-order term}} + \underbrace{\sum_{i=1}^{n} \big(d_i^2 - \lambda_i(t)^2\big)}_{\text{second-order term}}, \tag{11}$$

and

$$\mathbb{E}[r(Q(t), A(t))] = \sum_{i=1}^{n} \mathrm{Var}(A_i(t)). \tag{12}$$

The complete arguments for both results (10) and (12) can be found in Appendix A.4.1. Taking expectations, summing from $t = 1$ to $T - 1$ and applying Jensen's and Cauchy–Schwarz inequalities (see Appendix A.4.1 for the detailed derivation), we obtain:

$$\mathbb{E}\left[\sum_{i=1}^{n} Q_i(T)\right] \le n\sqrt{\sum_{t=1}^{T-1} \frac{1}{n}\, \mathbb{E}[f(Q(t), d(t))] + (T-1)(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)} + \sum_{i=1}^{n} \mathbb{E}[A_i(T)]. \tag{13}$$

We now present the performance guarantee of the LyapOpt policy.

Theorem 2 (Finite-Time Performance of the LyapOpt Policy)**.** *Within the model class* $\mathcal{M}^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}})$ *for* $\rho \in (0, 1]$*, the* LyapOpt *policy achieves the following bound on the expected total queue length:*

$$\mathbb{E}^{\mathrm{LyapOpt}}_{(A,S)}\left[\sum_{i=1}^{n} Q_i(T)\right] \le n\sqrt{\sum_{t=1}^{T-1} \frac{1}{n}\, \mathbb{E}^{\mathrm{LyapOpt}}_{(A,S)}\Big[\min_{d\in\mathcal{D}_t} f(Q(t), d)\Big] + (C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)(T-1)} + \sum_{i=1}^{n} \mathbb{E}_{(A,S)}[A_i(T)] \tag{14}$$

with $f(Q(t), d)$ *given by* (11)*. Moreover, for any fixed* $\rho \in (0, 1]$*, suppose that* $\lambda(t) \in \mathrm{star}(\mathcal{D}_t) := \{\alpha x : x \in \mathcal{D}_t,\, \alpha \in [0, 1]\}$ *for all* $t \in \mathbb{N}$*, then the expected total queue length is bounded by*

$$\mathbb{E}^{\mathrm{LyapOpt}}_{(A,S)}\left[\sum_{i=1}^{n} Q_i(T)\right] \le n\sqrt{(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)(T-1)} + \sum_{i=1}^{n} \mathbb{E}_{(A,S)}[A_i(T)]. \tag{15}$$

Proof. By inequality (9) and the definition of LyapOpt in (7), the one-step Lyapunov drift satisfies

$$\begin{aligned}
&\Delta V^{\mathrm{LyapOpt}}(t) \\
&\le \mathbb{E}\left[\sum_{i=1}^{n} (\max\{Q_i(t) - d_i^{\mathrm{LyapOpt}}(t), 0\})^2 \,\middle|\, \mathcal{F}_t\right] + \sum_{i=1}^{n} \mathrm{Var}(D_i^{\mathrm{LyapOpt}}(t)) - \sum_{i=1}^{n} (Q_i(t) - A_i(t))^2 \\
&= \min_{d\in\mathcal{D}_t} \left(\sum_{i=1}^{n} (\max\{Q_i(t) - d_i, 0\})^2\right) + \sum_{i=1}^{n} \mathrm{Var}(D_i^{\mathrm{LyapOpt}}(t)) - \sum_{i=1}^{n} (Q_i(t) - A_i(t))^2.
\end{aligned} \tag{16}$$

Then, by (10), we obtain

$$\Delta V^{\mathrm{LyapOpt}}(t) \le \min_{d\in\mathcal{D}_t} f(Q(t), d) + r(Q(t), A(t)) + nC_{\mathrm{d}}^2. \tag{17}$$

Applying the same argument as in the proof of (13) yields (14).

Fix any $\rho \in (0, 1]$. If $\lambda(t) \in \mathrm{star}(\mathcal{D}_t)$ for all $t \in \mathbb{N}$, then define $d^{(\rho)}(t) := \lambda(t)/\alpha^{(\rho)} \in \mathcal{D}_t$ for some

$\alpha^{(\rho)} < 1$. Substituting $d = d^{(\rho)}(t)$ into (16) gives,

$$\begin{aligned}
\Delta V^{\text{LyapOpt}}(t) &\leq \sum_{i=1}^{n}(\max\{Q_i(t) - d_i^{(\rho)}(t), 0\})^2 + nC_{\text{d}}^2 - \sum_{i=1}^{n}(Q_i(t) - A_i(t))^2 \\
&\leq \sum_{i=1}^{n}(\max\{Q_i(t) - \alpha^{(\rho)} d_i^{(\rho)}(t), 0\})^2 + nC_{\text{d}}^2 - \sum_{i=1}^{n}(Q_i(t) - A_i(t))^2 \\
&= f(Q(t), \alpha^{(\rho)} d^{(\rho)}(t)) + r(Q(t), A(t)) + nC_{\text{d}}^2 \\
&= r(Q(t), A(t)) + nC_{\text{d}}^2,
\end{aligned}$$

where the last equality follows from $f(Q(t), \alpha^{(\rho)} d^{(\rho)}(t)) = 0$. Taking expectations and applying the same telescoping argument as in the proof of (13) yields the bound in (15). □

From (14), it is evident that the key advantage of the LyapOpt policy lies in its simultaneous optimization of the first- and second-order terms by minimizing the full Lyapunov drift. This joint optimization enables the departure vector to align more precisely with the arrival rate vector, particularly when the scheduling set exhibits geometric irregularities.

As shown in (15), when the arrival rate lies within the star hull of the scheduling set, the LyapOpt policy perfectly aligns with the arrival rate vector (see Eqn. (11)), regardless of the underlying geometry of the scheduling set. Under this setting, LyapOpt attains, up to constant factors, the first term in the min in the fundamental minimax lower bound in (4). This demonstrates that in the finite-time, heavy-traffic regime, LyapOpt is optimal up to constants. In contrast, in the same heavy-traffic regime, MaxWeight is not optimal in finite-time (see Theorem 4 in the next section); rather it is optimal in the diffusion-scaled regime as demonstrated in Stolyar [43].

4.2 A Limitation of MaxWeight

The MaxWeight policy is widely recognized for its throughput optimality and asymptotic guarantees. Nevertheless, its performance in the finite-time regime is much less understood. In this subsection, we revisit established upper-bound results from the literature [45, 43, 37, 50] and present a complementary lower bound that clarifies the necessity of the factors appearing in this bound, thereby providing a more complete understanding of the policy's fundamental limitations.

4.2.1 Upper Bound of MaxWeight

Consider the MaxWeight policy, which selects the schedule according to

$$d^{\text{MaxWeight}}(t) \in \operatorname*{argmax}_{d \in \mathcal{D}_t} \langle Q(t), d \rangle.$$

By optimizing the first-order Lyapunov drift specified in (11), MaxWeight gives priority to queues with larger backlogs. The finite-time upper bound on the total queue length under MaxWeight, within the model class $\mathcal{M}^{\rho}(B, C_{\text{a}}, C_{\text{d}})$, is given below.

Theorem 3 (Upper Bound of MaxWeight Policy). *Within the model class $\mathcal{M}^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}})$ for $\rho \in (0, 1]$, the MaxWeight policy satisfies the following upper bound on the expected total queue length:*

$$\mathbb{E}_{(A,S)}^{\text{MaxWeight}}\left[\sum_{i=1}^{n} Q_i(T)\right] \leq n\sqrt{(B^2 + C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)(T-1)} + \sum_{i=1}^{n} \mathbb{E}_{(A,S)}[A_i(T)]. \tag{18}$$

Proof. Recall (11). The MaxWeight policy minimizes the first-order term of $f(Q(t), d)$ and ensures that this term remains non-positive at each step, since $\lambda(t) \in \rho\Pi(\mathcal{D}_t)$ and, by definition of $\Pi(\mathcal{D}_t)$, $d^{\text{MaxWeight}}(t)$ maximizes $\langle Q(t), d\rangle$ over all $d \in \Pi(\mathcal{D}_t)$. Besides, the second-order term of $f(Q(t), d)$, namely, $\sum_{i=1}^{n}\left(d_i^2 - \lambda_i(t)^2\right)$, is upper bounded by nB^2. Substituting into (13), we obtain (18). □

The proof of the upper bound of the MaxWeight policy in Theorem 3 illustrates a key limitation of MaxWeight and, more broadly, of drift-based approaches. Specifically, these methods focus solely on optimizing the first-order component of the Lyapunov drift, overlooking the effects of the second-order term. This is particularly significant because MaxWeight consistently chooses extreme points of the scheduling set, which may reduce its sensitivity to the nuanced direction and length of the arrival rate vector as captured by the second-order term. As a result, it may potential lead to queue accumulation.

4.2.2 Lower Bound of MaxWeight for Dimension 2

The upper bound in Theorem 3 suggests that MaxWeight's performance depends on the number of queues n, the variance parameters C_{a} and C_{d}, and the capacity parameter B. In the heavy-traffic regime, while the lower bound in Theorem 1 highlights the roles of n, C_{a} and C_{d}, the following lower bound further reveals the inherent dependence of the MaxWeight policy on B.

Theorem 4 (Lower Bound of MaxWeight Policy). *Let $n = 2$, and consider the family of model classes $\{\mathcal{M}^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}}) : \rho \in (0, 1]\}$. Let $c_1 > 0$ and $\psi_1(n, B, C_{\mathrm{a}}, C_{\mathrm{d}})$ be the constants defined in Theorem 1. Then for all $T \geq \psi_1(n, B, C_{\mathrm{a}}, C_{\mathrm{d}})^2 \mathbb{1}_{\{C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2 > 0\}} + 9$, the MaxWeight policy satisfies*

$$\liminf_{\rho \uparrow 1} \sup_{(\mathcal{D},A,S) \in \mathcal{M}^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}})} \mathbb{E}_{(A,S)}^{\text{MaxWeight}}\left[\sum_{i=1}^{n} Q_i(T)\right] \geq c_1\sqrt{(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)(T-1)} + \frac{BT^{\frac{1}{3}}}{4\sqrt{3}}. \tag{19}$$

Proof Sketch. The formal proof is provided in Appendix A.4.2. Here we outline the main idea. We construct a family of deterministic instances $\mathcal{I}_{b,q} := (\mathcal{D}, A, S) \in \mathcal{M}^1(B, C_{\mathrm{a}}, C_{\mathrm{d}})$. The scheduling sequence $\mathcal{D}$ is time-invariant with $\mathcal{D}_t = \mathcal{D}_\star$ for all t; the arrival process A is deterministic; and the random departure field satisfies $S_d(t) = d$ for each t. That is, if the scheduling policy selects $d(t) \in \mathcal{D}_t$, then the actual departure is $D(t) = d(t)$. The set $\mathcal{D}_\star$ and the process A are illustrated in Figure 2. Formally,

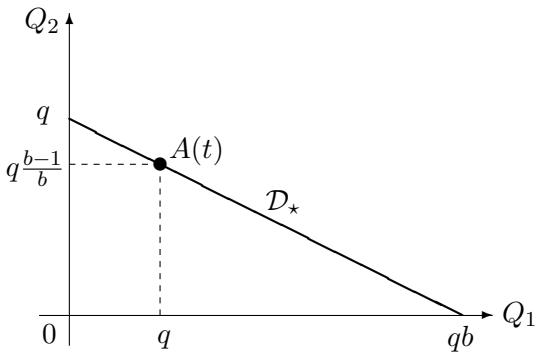


Figure 2: Construction of the instance $\mathcal{I}_{b,q}$

$$\begin{aligned}
\mathcal{D}_t = \mathcal{D}_\star &= \left\{d \in \mathbb{R}^2 : d = x(qb, 0) + (1-x)(0, q), 0 \le x \le 1\right\},\ \forall t \in \mathbb{N}_0; \\
A(t) &= q\Big(1, \frac{b-1}{b}\Big) \text{ for all } t \ge 2, \text{ and } A(1) = q\Big(1, \frac{b-1}{b} - \varepsilon\Big); \\
S_d(t) &= d,\ \forall d \in \mathcal{D}_\star,
\end{aligned} \tag{20}$$

where $b > 1$ is a rational constant and $\varepsilon > 0$ is a small irrational constant (introduced to avoid tie-breaking). To ensure that the instance lies within the model class $\mathcal{M}^1(B, C_\mathrm{a}, C_\mathrm{d})$, setting $q = \frac{\sqrt{2}B}{b}$. We establish the lower bound for the family of instances $\{\mathcal{I}_{b, \frac{\sqrt{2}B}{b}} : b \ge 6\}$.

Next, we prove the continuity of the expected total queue length under the MaxWeight policy in ρ in a neighborhood around 1, where ρ is a scalar that scales the arrival process. By combining this continuity result with the lower bound obtained from the deterministic construction above and with Theorem 1, we obtain the desired finite-time lower bound for the MaxWeight policy. □

The $BT^{\frac{1}{3}}$ term in Theorem 4 highlights that, in the heavy-traffic regime ($\rho \uparrow 1$), there exists hard instances, especially those with asymmetric scheduling sets, within $\mathcal{M}^\rho(B, C_\mathrm{a}, C_\mathrm{d})$ for which the performance of the MaxWeight policy exhibits an intrinsic dependence on the scheduling set's capacity parameter B. This establishes a fundamental performance limitation of MaxWeight in the finite-time, heavy-traffic regime.

This result does not contradict MaxWeight's classical optimality in the diffusion-scaled, heavy-traffic regime [43]. First, our analysis characterizes its performance over a model class rather than a specific instance considered in [43]. Second, even for a fixed instance, as shown in the following corollary, the dependence on B appears only in the finite horizon and therefore vanishes in the diffusion-scaled regime [40].

In order to quantify the finite-time gap between LyapOpt and MaxWeight, we focus on instances with $C_\mathrm{a} = C_\mathrm{d} = 0$. The following corollary analyzes this gap by comparing the lower bound for MaxWeight (from Theorem 4) with the performance guarantee of LyapOpt. The proof is provided in Appendix A.4.3.

Corollary 1 (Performance Gap between LyapOpt and MaxWeight)**.** *There exists a family of instances in* $\mathcal{M}^1(B, 0, 0)$ *with* $B \ge 3\sqrt{2}$*, for which the expected total queue lengths under* LyapOpt

and MaxWeight *satisfy:*

$$\sum_{i=1}^{2} Q_i^{\text{LyapOpt}}(T) \le 2 - \frac{1}{\sqrt{2B}}, \qquad T \ge 1;$$
$$\sum_{i=1}^{2} Q_i^{\text{MaxWeight}}(T) \ge \frac{\sqrt{BT}}{2^{\frac{5}{4}}}, \qquad \left\lceil \frac{2B^2}{\sqrt{2B}-1} \right\rceil \le T \le \left\lceil \left(\frac{\sqrt{2}B}{2} - 1\right)^3 \right\rceil + 1.$$

Proof Sketch. We consider the instance $\mathcal{I}_{\sqrt{2}B,1}$ given in (20), where $b = \sqrt{2}B$ and $q = 1$. Since $\lambda(t) = A(t) \in \mathcal{D}_\star$, We apply the bound in (15) for LyapOpt, together with the bound for MaxWeight derived in the proof of Theorem 4, which is stated in (44) in Appendix A.4.2. □

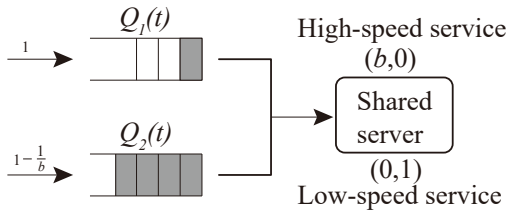


Figure 3: Asymmetric-service shared queueing model

The model used in the proof of the lower bound in (20) ($q = 1$) and illustrated in Figure 3 is not contrived. Instead, it frequently arises in real-world systems such as wireless networks and data centers [34, 12], where two queues—$Q_1(t)$ and $Q_2(t)$—have similar arrival rates $(1,\ 1 - \frac{1}{b})$ with $b \gg 1$, and share a resource whose service shifts between rate b for Q_1 and rate 1 for Q_2. MaxWeight tends to over-prioritize Q_1 by selecting extreme points, causing backlog in the harder-to-serve Q_2 and revealing its finite-time inefficiency. In contrast, LyapOpt adapts to the arrival rate and avoids backlog in the deterministic setting.

Consequently, the MaxWeight policy, which focuses solely on minimizing the first-order Lyapunov drift, inherently overlooks higher-order terms that are essential for optimal finite-time performance. This oversight results in excessively long queues, especially in scenarios with asymmetric scheduling sets over a finite horizon. Our theoretical findings directly inspire the creation of improved scheduling algorithms that explicitly account for second-order Lyapunov drift terms, leading to better finite-time outcomes.

5 Stability of LyapOpt in the Interior of the Capacity Region

In this section, we demonstrate that LyapOpt inherits the stability traditionally guaranteed by MaxWeight. Specifically, stability means that the queue length process does not blow up to infinity; it is positive recurrent in the stationary setting and uniformly bounded in expectation over time in the nonstationary setting.

To establish this result, we employ a standard Lyapunov drift analysis. Intuitively, LyapOpt directly minimizes a derandomized one-step Lyapunov drift, which is therefore no greater than the

first-order drift term under MaxWeight, up to additional variance and capacity terms. Consequently, all stability results obtained for MaxWeight through drift analysis also apply to LyapOpt. Formally, under the LyapOpt policy, within the model class $\mathcal{M}^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}})$, the one-step Lyapunov drift satisfies the following sequence of inequalities (the detailed proof is provided in Appendix A.5.1).

$$
\begin{aligned}
&\mathbb{E}^{\mathrm{LyapOpt}}_{(A,S)}\left[V(Q(t+1)) - V(Q(t)) \mid Q(t) = x\right] \\
&\leq \min_{d \in \mathcal{D}_\star} \underbrace{\sum_{i=1}^{n} \left((\max\{x_i(t) - d_i, 0\})^2 - x_i(t)^2 + 2x_i(t)\lambda_i(t+1) + \lambda_i(t+1)^2 \right)}_{\text{derandomized one-step Lyapunov drift}} + n(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2) \\
&\leq \min_{d \in \mathcal{D}_\star} \sum_{i=1}^{n} \left((x_i(t) - d_i)^2 - x_i(t)^2 + 2x_i(t)\lambda_i(t+1) + \lambda_i(t+1)^2 \right) + n(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2) \\
&\leq \underbrace{\min_{d \in \mathcal{D}_\star} \sum_{i=1}^{n} 2x_i(t)(\lambda_i(t+1) - d_i)}_{\text{first-order drift term (MaxWeight)}} + 2nB^2 + n(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2).
\end{aligned}
$$

We first consider the classical stationary setting, where the scheduling set $\mathcal{D}_t = \mathcal{D}_\star$ for all t, $A(t)$ is i.i.d. across t, and $S_d(t)$ is i.i.d. across t for each $d \in \mathcal{D}_\star$. In this setting, the induced queue-length process $\{Q(t)\}_{t \in \mathbb{N}_0}$ evolves as a Markov chain. To prove that $\{Q(t)\}_{t \in \mathbb{N}_0}$ is stable, we apply the classical Foster–Lyapunov criterion (see Lemma 5 in Appendix A.5.1), which is commonly used in analyzing the stability of MaxWeight.

The stability result is stated in the following theorem. To avoid unnecessary complexity, we focus on a countable state space in which all processes take integer values. The general case also holds under mild assumptions using standard arguments. The formal proof is deferred to Appendix A.5.1.

Theorem 5 (Stability of the LyapOpt Policy in the Stationary Setting)**.** *Suppose that all processes take integer values and are stationary; namely, $\mathcal{D}_t = \mathcal{D}_\star$ for all t, $A(t)$ is i.i.d. across t, and $S_d(t)$ is i.i.d. across t for each $d \in \mathcal{D}_\star$. Then, within the model class $\mathcal{M}^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}})$ for $\rho \in (0, 1)$, the queue length process $\{Q(t)\}_{t \in \mathbb{N}_0}$ under the* LyapOpt *policy is stochastically stable, in the sense that it reaches the recurrent states with probability 1 and every recurrent state is positive recurrent.*

This theorem establishes that when the traffic intensity $\rho < 1$, LyapOpt retains the positive-recurrence property of MaxWeight in the stationary regime. This behavior arises from the inherent nature of LyapOpt, as it minimizes the full one-step Lyapunov drift.

We next establish the stability of LyapOpt in the nonstationary setting. To this end, we introduce the following assumption, which ensures that for each queue, there exists at least one schedule capable of serving it at a minimum rate of α, uniformly over all time slots.

Assumption 1 (Uniform coordinate coverage)**.** There exists a constant $\alpha > 0$ such that

$$
\forall\, t \in \mathbb{N}_0, \forall\, j \in \{1, \ldots, n\}, \exists\, d(t) \in \mathcal{D}_t \text{ such that } d_j(t) \geq \alpha.
$$

With this assumption, we can now establish the stability of LyapOpt in the nonstationary setting. The proof is presented in Appendix A.5.2.

Proposition 1 (Stability of the LyapOpt Policy in the Nonstationary Setting)**.** *Under Assumption 1, within the model class* $\mathcal{M}^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}})$ *for* $\rho \in (0,1)$*, the queue-length process* $\{Q(t)\}_{t\in\mathbb{N}_0}$ *under the* LyapOpt *policy is stable, in the sense that for* $T \geq 1$*,*

$$\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{(A,S)}^{\mathrm{LyapOpt}}\left[\sum_{i=1}^{n}Q_i(t)\right] \leq \frac{n^2(B^2+C_{\mathrm{a}}^2+C_{\mathrm{d}}^2)}{2(1-\rho)\alpha}. \tag{21}$$

Proposition 1 shows that, in the general nonstationary setting, the LyapOpt policy ensures that the time-averaged expected queue lengths remain uniformly bounded, a property that also holds for the MaxWeight policy.

Together, Theorem 5 and Proposition 1 establish the stability of LyapOpt in both stationary and nonstationary settings when the traffic intensity $\rho < 1$, using a drift analysis argument similar to that of MaxWeight. Combined with its finite-time minimax optimality, these results demonstrate that LyapOpt provides a unified framework ensuring both short-term optimality and long-term stability.

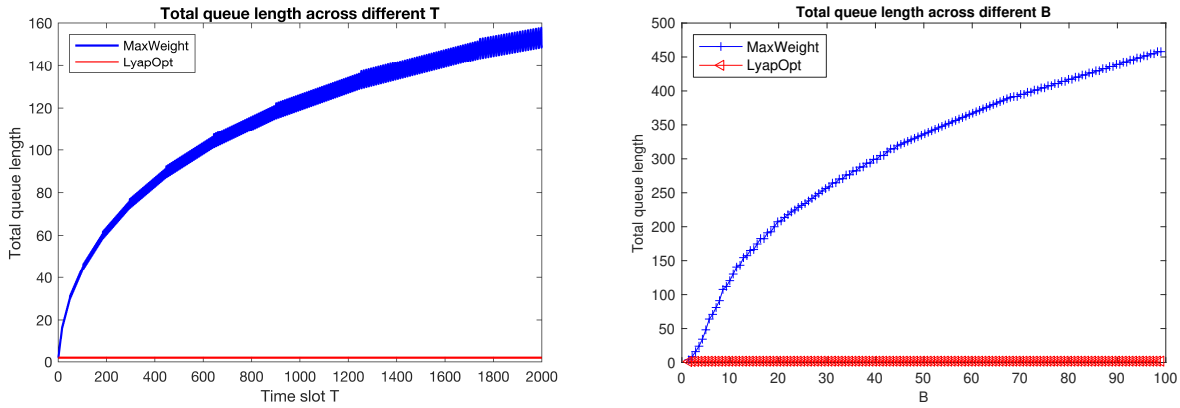
6 Numerical Experiments

In this section, we present numerical results that validate the theoretical findings and illustrate the finite-time performance differences between LyapOpt and MaxWeight across a broad range of settings. The simulation scripts are publicly accessible at `https://github.com/LYJ-aa/lyapopt-queueing-code`.

6.1 Validation of the Hard Instance

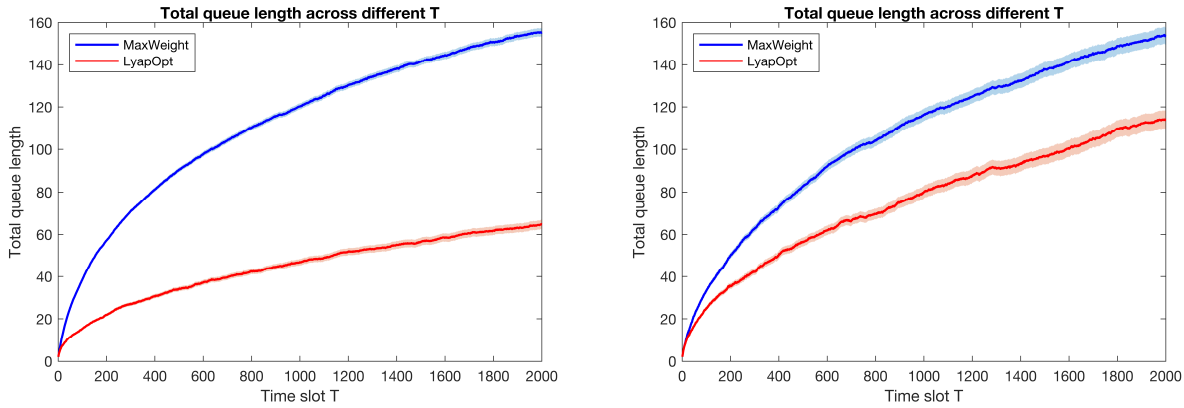
We demonstrate that the $\sqrt{BT}$ gap established in Corollary 1, for the construction $\mathcal{I}_{\sqrt{2}B,1}$ given in (20), is evident for practical values of T and B, as shown in Figures 4a and 4b, where $C_{\mathrm{a}} = C_{\mathrm{d}} = 0$ and $b = \sqrt{2}B$. However, as shown in Figures 9a and 9b in Appendix A.6, the total queue length under MaxWeight eventually becomes bounded (i.e., $O(1)$ in T) as T increases.

Moreover, although we do not have a theoretical guarantee for the same construction when the arrival process and random departure field are extended to include variances (i.e., from $C_{\mathrm{a}} = C_{\mathrm{d}} = 0$ to $C_{\mathrm{a}}, C_{\mathrm{d}} > 0$), Figures 5a and 5b indicate that the performance advantage of LyapOpt over MaxWeight remains substantial.



(a) Total queue length when $B = 10$, , $C_{\mathrm{a}} = C_{\mathrm{d}} = 0$ (b) Total queue length when $T = 1000$, $C_{\mathrm{a}} = C_{\mathrm{d}} = 0$

Figure 4: Performance comparison of MaxWeight and LyapOpt policies versus B



(a) Total queue length when $B = 10$, $C_{\mathrm{a}} = C_{\mathrm{d}} = 1$ (b) Total queue length when $B = 10$, $C_{\mathrm{a}} = C_{\mathrm{d}} = 2$

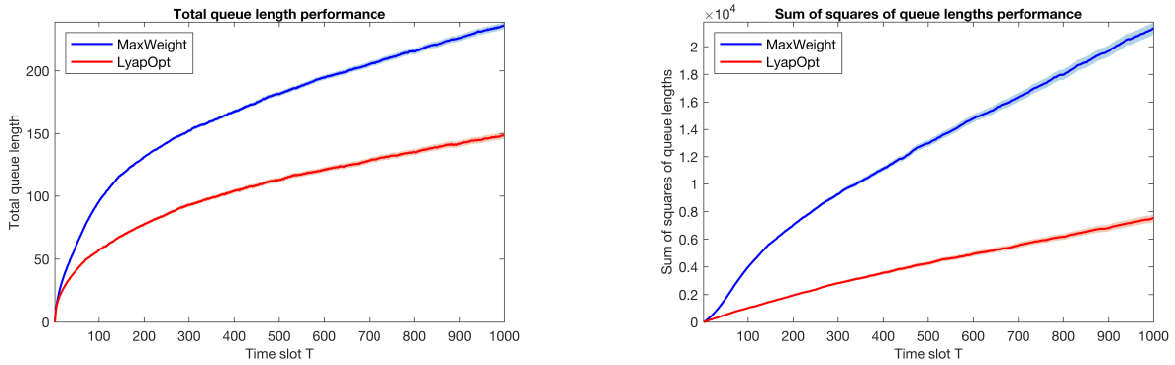
Figure 5: Performance comparison of MaxWeight and LyapOpt policies when $C_{\mathrm{a}}, C_{\mathrm{d}} > 0$

6.2 Comprehensive Experiments

We compare the finite-time performance of LyapOpt and MaxWeight as the number of parallel queues n increases from 2 to 8. For each n, the time-invariant scheduling set $\mathcal{D}_\star$ is generated by uniformly sampling $10n$ vectors with integer entries between 1 and 10. Subsequently, 2000 arrival rate vectors are uniformly drawn from the boundary of the capacity region $\rho\Pi(\mathcal{D}_\star)$, each representing a distinct scenario (a total of 2000 scenarios per n). In each scenario, arrivals evolve i.i.d. over time according to Gamma distributions with fixed variance 1 per queue, calibrated to match the corresponding sampled arrival rate vector, while departures are modeled as Gamma random variables with fixed variance 1 per queue, calibrated to match the departure vectors selected by the policy.

Figures 6a and 6b show a representative case with $n = 8$ and an arrival rate on the boundary of the capacity region $\Pi(\mathcal{D}_\star)$, corresponding to the heavy-traffic regime. The results are averaged over 1000 independent realizations, and the accompanying error bars indicate 95% confidence intervals.

Both figures demonstrate that LyapOpt significantly outperforms MaxWeight. In Figure 6a, both policies appear to exhibit $\sqrt{T}$ growth, but LyapOpt maintains consistently lower total queue length. Figure 6b further shows reduced queue imbalance under LyapOpt, as reflected in the lower sum of squared queue lengths.

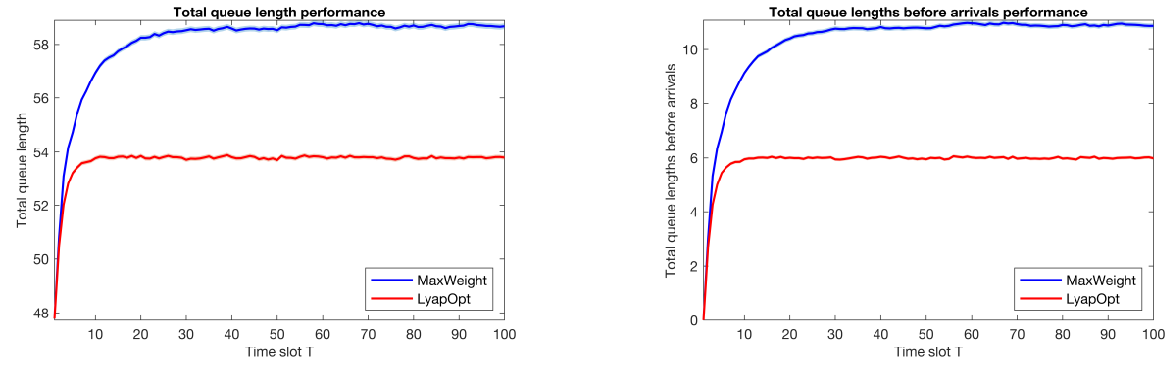


(a) Total queue length

(b) Sum of squares of queue lengths.

Figure 6: Finite-time comparison of MaxWeight and LyapOpt policies ($n = 8$, $\rho = 1$).

Figures 7a and 7b illustrate a representative case with $n = 8$ and an arrival rate on the boundary of the capacity region $\rho\Pi(\mathcal{D}_\star)$ for $\rho = 0.9$, corresponding to the interior regime. Specifically, the vertical axis in Figure 7a is $\sum_{i=1}^{n} Q_i(T)$, while in Figure 7b it is $\sum_{i=1}^{n}(Q_i(T) - A_i(T))$. The results are averaged over 10,000 independent realizations, and the accompanying error bars indicate 95% confidence intervals. Both figures demonstrate that LyapOpt outperforms MaxWeight and that the system reach stability quickly. In Figure 7b, after removing the common new arrivals, we can see that the queue length under the LyapOpt policy is noticeably smaller than that under MaxWeight.



(a) Total queue length

(b) Total queue length before arrivals.

Figure 7: Finite-time comparison of MaxWeight and LyapOpt policies ($n = 8$, $\rho = 0.9$).

Table 3 reports, for each number of queues n ($2 \leq n \leq 8$), the proportion of scenarios where the

ratio

$$\text{ratio} = \frac{\text{Expected Total Queue Length (LyapOpt) at } t = 1000}{\text{Expected Total Queue Length (MaxWeight) at } t = 1000}$$

falls below 1, 0.9, and 0.8, respectively—indicating cases where LyapOpt outperforms MaxWeight. The expected total queue length is averaged over 100 independent realizations in each scenario. The accompanying error bars represent 95% confidence intervals based on 2000 independently generated arrival scenarios. These results show that for a wide range of queues, LyapOpt consistently outperforms MaxWeight, sometimes substantially so.

Table 3: Proportion of scenarios with ratio below 1, 0.9, and 0.8 ($\rho = 1$)

Number of Queues (n)	ratio ≤ 1	ratio ≤ 0.9	ratio ≤ 0.8
2	63.30% $\pm$ 2.11%	9.90% $\pm$ 1.31%	0.10% $\pm$ 0.14%
3	97.30% $\pm$ 0.71%	78.75% $\pm$ 1.79%	38.30% $\pm$ 2.13%
4	99.70% $\pm$ 0.24%	95.60% $\pm$ 0.90%	85.85% $\pm$ 1.53%
5	99.35% $\pm$ 0.35%	92.35% $\pm$ 1.16%	61.25% $\pm$ 2.14%
6	97.40% $\pm$ 0.70%	87.40% $\pm$ 1.45%	72.00% $\pm$ 1.97%
7	100.00% $\pm$ 0.00%	100.00% $\pm$ 0.00%	100.00% $\pm$ 0.00%
8	100.00% $\pm$ 0.00%	99.50% $\pm$ 0.31%	92.00% $\pm$ 1.19%

7 Extensions

7.1 Minimax Lower Bound–Independent Case

Some classical queueing models assume independence across queues in both arrivals and departures to facilitate asymptotic analyses [45, 19, 43]. To complement our general analysis, we also present the bound under this assumption.

We consider the subclass capturing independence across queues:

$$\mathcal{M}^{\rho}_{\mathcal{I}}(B, C_{\mathrm{a}}, C_{\mathrm{d}}) = \left\{ (\mathcal{D}, A, S) \in \mathcal{M}^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}}) \,:\, \begin{array}{l} \forall\, t \in \mathbb{N},\ \{A_i(t)\}_{i=1}^{n} \text{ indpt across } i \\ \forall\, t \in \mathbb{N},\ \{S_d(t)_i\}_{i=1}^{n} \text{ indpt across } i, \forall\, d \in \mathcal{D}_t \end{array} \right\}.$$

Let

$$\Gamma^{\rho}_{\mathcal{I}}(B, C_{\mathrm{a}}, C_{\mathrm{d}}, T) := \inf_{\Phi} \sup_{(\mathcal{D},A,S)\in\mathcal{M}^{\rho}_{\mathcal{I}}(B,C_{\mathrm{a}},C_{\mathrm{d}})} \mathbb{E}^{\Phi}_{(A,S)} \left[\sum_{i=1}^{n} Q_i(T) \right].$$

The following proposition characterizes the minimax expected total queue length $\Gamma^{\rho}_{\mathcal{I}}(B, C_{\mathrm{a}}, C_{\mathrm{d}}, T)$. The proof is presented in Appendix A.7.1.

Proposition 2 (Minimax Lower Bound–Independent Case)**.** *For any scheduling policy, and for scheduling sets, arrival processes and random departure fields within the model class* $\mathcal{M}^{\rho}_{\mathcal{I}}(B, C_{\mathrm{a}}, C_{\mathrm{d}})$*,*

there exist an absolute constant $c_1 > 0$ *and a constant* $\psi_2(n, B, C_\mathrm{a}, C_\mathrm{d}) > 0$ *such that, for*

$$\rho \in \begin{cases} \left[\left(1 - \frac{\sqrt{C_\mathrm{a}^2 + C_\mathrm{d}^2}}{B\psi_2(n,B,C_\mathrm{a},C_\mathrm{d})}\right) \vee 1/2, 1\right], & \text{if } C_\mathrm{a}^2 + C_\mathrm{d}^2 > 0, \\ (0, 1], & \text{if } C_\mathrm{a}^2 + C_\mathrm{d}^2 = 0, \end{cases}$$

the following lower bound holds for all $T \geq \psi_2(n, B, C_\mathrm{a}, C_\mathrm{d})^2 + 1$,

$$\Gamma_{\mathcal{I}}^{\rho}(B, C_\mathrm{a}, C_\mathrm{d}, T) \geq c_1 \min\left\{\sqrt{n(C_\mathrm{a}^2 + C_\mathrm{d}^2)(T-1)}, \frac{C_\mathrm{a}^2 + C_\mathrm{d}^2}{B(1-\rho)}\right\} + n\rho B. \tag{22}$$

Compared with Theorem 1, the two lower bounds exhibit the same dependence on T, ρ, B, C_a, and C_d, the only difference lies in their scaling with respect to the number of queues n.

This lower bound relates to the asymptotic results of Stolyar [43] through its dependence on n, C_a, and C_d, although the two are not directly comparable. In Stolyar [43], the analysis is performed for individual problem instances. By contrast, our results address the worst-case performance over an entire model class, yielding a minimax characterization. Moreover, the two works differ in the underlying performance metrics.

7.2 Other Model Classes

The previous sections considered the model class $\mathcal{M}^{\rho}(B, C_\mathrm{a}, C_\mathrm{d})$, in which the scheduling set is allowed to vary over time to capture a more general and challenging problem. In this section, we focus on the time-invariant scheduling set, as is common in the classical queueing literature [45, 33]. This setting simplifies the model and further highlights that the independence of the lower bound from the scheduling set capacity parameter B in the finite-time regime, together with the dependence of the MaxWeight bound on B, are both critical and fundamental characteristics of the system.

Fix the scheduling set $\mathcal{D}_\star$ for all t, and consider a reduced model class that does not involve scheduling sets explicitly:

$$\mathcal{M}_{\mathcal{D}_\star}^{\rho}(C_\mathrm{a}, C_\mathrm{d}) = \left\{(A, S) \,:\, \begin{array}{l} \lambda(t) \in \rho\Pi(\mathcal{D}_\star),\ \frac{1}{n}\sum_{i=1}^{n} \mathrm{Var}(A_i(t)) \leq C_\mathrm{a}^2,\ \ \forall\, t \in \mathbb{N}; \\ \frac{1}{n}\sum_{i=1}^{n} \mathrm{Var}((S_d(t))_i) \leq C_\mathrm{d}^2,\ \ \forall\, d \in \mathcal{D}_\star, t \in \mathbb{N}_0 \end{array}\right\}$$

where the arrival process A, and the random departure field S are those defined in Section 2.1.

Our objective is to characterize the fundamental minimax expected total queue length at time T:

$$\Gamma_{\mathcal{D}_\star}^{\rho}(C_\mathrm{a}, C_\mathrm{d}, T) := \inf_{\Phi} \sup_{(A,S)\in\mathcal{M}_{\mathcal{D}_\star}^{\rho}(C_\mathrm{a},C_\mathrm{d})} \mathbb{E}_{(A,S)}^{\Phi}\left[\sum_{i=1}^{n} Q_i(T)\right].$$

The following proposition presents the minimax lower bound for the model class $\mathcal{M}_{\mathcal{D}_\star}^{\rho}(C_\mathrm{a}, C_\mathrm{d})$. Fixing $\mathcal{D}_\star$ allows the lower bound to be characterized in terms of the geometric properties of $\mathcal{D}_\star$. In particular, we introduce three parameters, M_1, M_2, and M_3, associated with $\mathcal{D}_\star$, which represent the maximum total capacity, the maximum ℓ_2 capacity norm, and the minimum ℓ_2 norm of the

vector aligned with the capacity, respectively.

$$M_1 := \max_{d\in\mathcal{D}_\star} \sum_{i=1}^n d_i; \quad M_2 := \max_{d\in\mathcal{D}_\star} \sqrt{\sum_{i=1}^n d_i^2};$$
$$M_3 := \min\left\{ \sqrt{\sum_{i=1}^n \gamma_i^2} \ : \ \gamma\in\Pi(\mathcal{D}_\star),\ \sum_{i=1}^n \gamma_i = M_1 \right\}.$$

Proposition 3 (Minimax Lower Bound–Fixed Scheduling Set)**.** *Fix a scheduling set $\mathcal{D}_\star$. For any scheduling policy, and for arrival processes and random departure fields in the model class $\mathcal{M}^\rho_{\mathcal{D}_\star}(C_\mathrm{a}, C_\mathrm{d})$, there exist an absolute constant $c_1 > 0$ and a finite constant $\psi_3(n, C_\mathrm{a}, C_\mathrm{d}, \mathcal{D}_\star) > 0$ such that, for*

$$\rho \in \begin{cases} \left[\left(1 - \frac{\sqrt{C_\mathrm{a}^2 + C_\mathrm{d}^2}}{B\psi_3(n, C_\mathrm{a}, C_\mathrm{d}, \mathcal{D}_\star)}\right) \vee 1/2, 1 \right], & \text{if } C_\mathrm{a}^2 + C_\mathrm{d}^2 > 0, \\ (0,1], & \text{if } C_\mathrm{a}^2 + C_\mathrm{d}^2 = 0, \end{cases}$$

the following lower bound holds for all $T \geq \psi_3(n, C_\mathrm{a}, C_\mathrm{d}, \mathcal{D}_\star)^2 \mathbb{1}_{\{C_\mathrm{a}^2 + C_\mathrm{d}^2 > 0\}} + 1$,

$$\Gamma^\rho_{\mathcal{D}_\star}(C_\mathrm{a}, C_\mathrm{d}, T) \geq c_1 \min\left\{ \sqrt{\left(\frac{M_1^2}{M_3^2} C_\mathrm{a}^2 + \frac{M_1^2}{M_2^2} C_\mathrm{d}^2\right) n(T-1)},\ \frac{n}{1-\rho}\left(\frac{M_1}{M_3^2} C_\mathrm{a}^2 + \frac{M_1}{M_2^2} C_\mathrm{d}^2\right) \right\} + \rho M_1.$$

Proof. By repeating the proof of Theorem 1, the desired result follows immediately. □

This proposition reveals that the lower bound for the model class $\mathcal{M}^\rho_{\mathcal{D}_\star}(C_\mathrm{a}, C_\mathrm{d})$ depends on the geometric discrepancy between different norms associated with the scheduling set $\mathcal{D}_\star$. To simplify this, we introduce the following variant of the model class:

$$\widetilde{\mathcal{M}}^\rho_{\mathcal{D}_\star}(C_\mathrm{a}, C_\mathrm{d}) = \left\{ (A,S) \ : \ \begin{array}{l} \lambda(t) \in \rho\Pi(\mathcal{D}_\star),\ \frac{1}{n}\sum_{i=1}^n \sqrt{\mathrm{Var}(A_i(t))} \leq C_\mathrm{a},\ \ \forall\, t\in\mathbb{N}; \\ \frac{1}{n}\sum_{i=1}^n \sqrt{\mathrm{Var}((S_d(t))_i)} \leq C_\mathrm{d},\ \ \forall\, d\in\mathcal{D}_\star, t\in\mathbb{N}_0 \end{array} \right\}$$

where the arrival process A, and the random departure field S are those defined in Section 2.1. This variant differs from $\mathcal{M}^\rho_{\mathcal{D}_\star}(C_\mathrm{a}, C_\mathrm{d})$ in that the constraints are placed on average *standard deviations* rather than *variances*.

Let $\widetilde{\Gamma}^\rho_{\mathcal{D}_\star}(C_\mathrm{a}, C_\mathrm{d}, T)$ denote the minimax expected total queue length at time T under the model class $\widetilde{\mathcal{M}}^\rho_{\mathcal{D}_\star}(C_\mathrm{a}, C_\mathrm{d})$. The following proposition establishes a lower bound of $\widetilde{\Gamma}^\rho_{\mathcal{D}_\star}(C_\mathrm{a}, C_\mathrm{d}, T)$. Within the model class $\widetilde{\mathcal{M}}^\rho_{\mathcal{D}_\star}(C_\mathrm{a}, C_\mathrm{d})$, the analysis simplifies and yields a cleaner expression, while remaining consistent with the lower bound in Theorem 1.

Proposition 4 (Minimax Lower Bound–Variant Model Class)**.** *Fix a scheduling set $\mathcal{D}_\star$. For any scheduling policy, and for arrival processes and random departure fields in the model class $\widetilde{\mathcal{M}}^\rho_{\mathcal{D}_\star}(C_\mathrm{a}, C_\mathrm{d})$, there exist an absolute constant $c_1 > 0$ and a finite constant $\psi_4(n, C_\mathrm{a}, C_\mathrm{d}, \mathcal{D}_\star) > 0$*

such that, for

$$\rho \in \begin{cases} \left[\left(1 - \frac{\sqrt{C_{\rm a}^2+C_{\rm d}^2}}{B\psi_4(n,C_{\rm a},C_{\rm d},\mathcal{D}_\star)}\right) \vee 1/2, 1\right], & \text{if } C_{\rm a}^2 + C_{\rm d}^2 > 0, \\ (0,1], & \text{if } C_{\rm a}^2 + C_{\rm d}^2 = 0, \end{cases}$$

the following lower bound holds for all $T \geq \psi_4(n, C_{\rm a}, C_{\rm d}, \mathcal{D}_\star)^2 \mathbb{1}_{\{C_{\rm a}^2+C_{\rm d}^2>0\}} + 1$,

$$\widetilde{\Gamma}^{\rho}_{\mathcal{D}_\star}(C_{\rm a}, C_{\rm d}, T) \geq c_1 \min\left\{n\sqrt{\left(C_{\rm a}^2 + C_{\rm d}^2\right)(T-1)},\ \frac{n^2(C_{\rm a}^2 + C_{\rm d}^2)}{M_1(1-\rho)}\right\} + \rho M_1. \tag{23}$$

Proof. The argument follows the proof of Theorem 1, with the tuning parameters in the arrival and departure distributions chosen as

$$K_{\rm a} := \left(\frac{nC_{\rm a}}{\rho M_1}\right)^2 + 1, \qquad K_{\rm d} := \left(\frac{nC_{\rm d}}{M_1}\right)^2 + 1.$$

□

In Corollary 1, we consider an instance with a time-invariant scheduling set and $C_{\rm a} = C_{\rm d} = 0$. In this case, the lower bound for MaxWeight still depends on the capacity parameter B within a finite time horizon. This dependence shows that the MaxWeight policy cannot achieve the lower bound given in (23). To close this gap and attain the lower bound, we introduce the following Empirical Projection (EP) policy:

$$d^{\rm EP}(t) \in \operatorname*{argmin}_{d \in \mathcal{D}_\star} \sum_{i=1}^{n} \max\{\widehat{A}_i(t) - d_i, 0\},$$

where $\widehat{A}(t) = \frac{1}{t}\sum_{s=1}^{t} A(s)$. The EP policy leverages empirical averages for scheduling. By projecting the empirical arrival rates onto the feasible scheduling set, it balances service with observed arrivals. The following theorem establishes its finite-time performance guarantee.

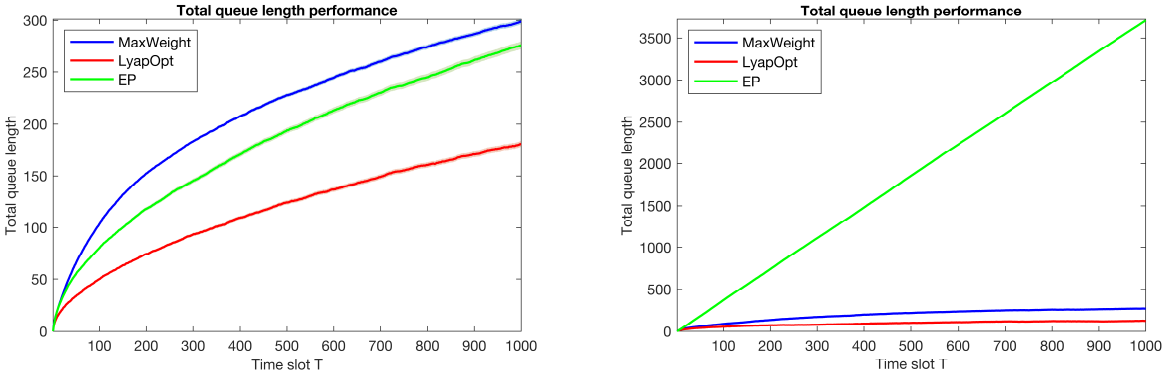
Proposition 5 (Finite-Time Performance of the EP Policy)**.** *Assume that* $A(t)$ *is i.i.d. across* t *with mean* $\mathbb{E}[A(t)] = \lambda \in \mathcal{D}_\star$*. Within the model class* $\widetilde{\mathcal{M}}^{\rho}_{\mathcal{D}_\star}(C_{\rm a}, C_{\rm d})$ *for* $\rho \in (0,1]$*, the* EP *policy guarantees the following bound on the expected total queue length: for all* $T \geq 1$*,*

$$\mathbb{E}^{\rm EP}_{(A,S)}\left[\sum_{i=1}^{n} Q_i(T)\right] \leq n\sqrt{T-1}(6C_{\rm a} + 2C_{\rm d}) + \sum_{i=1}^{n} \mathbb{E}_{(A,S)}[A_i(T)]. \tag{24}$$

The proof is provided in Appendix A.7.2. This result shows that the EP policy achieves the same $O\big(n(C_{\rm a}+C_{\rm d})\sqrt{T}\big)$ scaling as the first term in the min in the lower bound (23), which is therefore tight within the model class $\widetilde{\mathcal{M}}^{\rho}_{\mathcal{D}_\star}(C_{\rm a}, C_{\rm d})$.

However, minimax optimality of the EP policy (up to universal factors) in the finite-time regime requires the strong assumption that the arrival process $A(t)$ is i.i.d. across t with mean

$\mathbb{E}[A(t)] = \lambda \in \mathcal{D}_\star$. When this condition fails, the policy can perform poorly (see Figure 8a); and even under the assumption, it performs no better than LyapOpt or MaxWeight (see Figure 8b), under the same experimental setting as in Section 6.2.



(a) Total queue length when $\lambda \in \mathcal{D}_\star$

(b) Total queue length when $\lambda \notin \mathcal{D}_\star$

Figure 8: Performance comparison of EP, MaxWeight and LyapOpt policies

8 Conclusion and Future Directions

We established a minimax framework to study the scheduling problem in parallel queueing systems in the finite-time regime. Through this framework, we identified a fundamental finite-time gap between the MaxWeight policy and the minimax lower bound, which depends explicitly on the capacity of the scheduling set. To address this gap, we introduced LyapOpt, a second-order Lyapunov policy that attains the lower bound under certain structural conditions while retaining stability guarantees.

Beyond the finite-time regime, we conjecture that LyapOpt exhibits state space collapse (see, e.g., [5]) in the fluid-scaled regime, meaning that the high-dimensional queue length process evolves along a lower-dimensional manifold. This behavior is typically established through Lyapunov drift and fluid-limit analysis, and LyapOpt directly minimizes the full one-step Lyapunov drift. This, in turn, suggests that LyapOpt may also achieve diffusion-scaled workload optimality in the heavy-traffic regime, since state space collapse is a key step toward establishing such optimality. This constitutes one direction for future research. Another important direction is to extend the finite-time optimality of LyapOpt to multi-hop networks via a backpressure-style generalization, which would enhance its applicability to practical systems and introduce additional challenges arising from routing and interference among queues.

Another practical direction is to develop efficient methods for implementing LyapOpt in large-scale systems, especially when the scheduling constraints admit combinatorial structure such as bipartite matching graphs [31, 32]. Such structures, long central to Internet routing, are now equally important in data centers and cloud platforms. This motivates exploiting the matching geometry to design fast quadratic optimizers or structured approximations for LyapOpt, enabling its real-time advantages in practice.

References

[1] J. Abate and W. Whitt. Transient behavior of the M/M/1 queue: Starting at the origin. *Queueing systems*, 2(1):41–65, 1987.

[2] M. Andrews, K. Kumaran, K. Ramanan, A. L. Stolyar, R. Vijayakumar, and P. Whiting. Scheduling in a queuing system with asynchronously varying service rates. *Probability in the Engineering and Informational Sciences*, 18(2):191–217, 2004.

[3] R. Ao, G. Luo, D. Simchi-Levi, and X. Wang. Optimizing LLM inference: Fluid-guided online scheduling with memory constraints. *arXiv preprint arXiv:2504.11320*, 2025.

[4] D. P. Bertsekas. *Dynamic Programming and Optimal Control: Volume I*. Athena scientific, 4th edition, 2012.

[5] M. Bramson. State space collapse with application to heavy traffic limits for multiclass queueing networks. *Queueing Systems*, 30(1):89–140, 1998.

[6] N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

[7] F. Chen, D. J. Foster, Y. Han, J. Qian, A. Rakhlin, and Y. Xu. Assouad, fano, and le cam with interaction: A unifying lower bound framework and characterization for bandit learnability. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

[8] H. Chen and D. D. Yao. *Fundamentals of Queueing Networks: Performance, Asymptotics, and Optimization*, volume 46. Springer Science & Business Media, 2001.

[9] J. G. Dai and W. Lin. Maximum pressure policies in stochastic processing networks. *Operations Research*, 53(2):197–218, 2005.

[10] J. G. Dai and W. Lin. Asymptotic optimality of maximum pressure policies in stochastic processing networks. *The Annals of Applied Probability*, 18(6):2239–2299, 2008.

[11] Y. Dai and L. Huang. Adversarial network optimization under bandit feedback: Maximizing utility in non-stationary multi-hop networks. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 8(3):1–48, 2024.

[12] C. D'Apice, A. Dudin, S. Dudin, and R. Manzo. Priority queueing system with many types of requests and restricted processor sharing. *Journal of Ambient Intelligence and Humanized Computing*, 14(9):12651–12662, 2023.

[13] R. Dar and E. Karabulut. NVIDIA open sources run: AI scheduler to foster community collaboration, 2025. URL `https://developer.nvidia.com/blog/nvidia-open-sources-runai-scheduler-to-foster-community-collaboration/`.

[14] A. Eryilmaz and R. Srikant. Asymptotically tight steady-state queue length bounds implied by drift conditions. *Queueing Systems*, 72:311–359, 2012.

[15] D. Freund, T. Lykouris, and W. Weng. Quantifying the cost of learning in queueing systems. *Advances in Neural Information Processing Systems*, 36:6532–6544, 2023.

[16] D. Gamarnik and D. A. Goldberg. On the rate of convergence to stationarity of the M/M/N queue in the Halfin–Whitt regime. *The Annals of Applied Probability*, 23(5):1845–1882, 2013.

[17] G. R. Gupta and N. B. Shroff. Delay analysis for multi-hop wireless networks. In *Proceedings of the 28th IEEE International Conference on Computer Communications (INFOCOM)*, pages 2356–2364. IEEE, 2009.

[18] J. M. Harrison. *Brownian Models of Performance and Control*. Cambridge University Press, 2013.

[19] J. M. Harrison and J. A. Van Mieghem. Dynamic control of Brownian networks: State-space collapse and equivalent workload formulations. *The Annals of Applied Probability*, 7(3):747–771, 1997.

[20] T. Hirayama, M. Kijima, and S. Nishimura. Further results for dynamic scheduling of multiclass G/G/1 queues. *Journal of Applied Probability*, 26(3):595–603, 1989.

[21] W.-K. Hsu, J. Xu, X. Lin, and M. R. Bell. Integrated online learning and adaptive control in queueing systems with uncertain payoffs. *Operations Research*, 70(2):1166–1181, 2022.

[22] S. Krishnasamy, A. Arapostathis, R. Johari, and S. Shakkottai. On learning the cμ rule in single and parallel server networks. In *2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 153–154. IEEE, 2018.

[23] S. Krishnasamy, R. Sen, R. Johari, and S. Shakkottai. Learning unknown service rates in queues: A multiarmed bandit approach. *Operations Research*, 69(1):315–330, 2021.

[24] T. Lattimore and C. Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2020.

[25] E. L. Lehmann and G. Casella. *Theory of Point Estimation*. Springer, 1998.

[26] Y. Li, J. G. Dai, and T. Peng. Throughput-optimal scheduling algorithms for LLM inference and AI agents. *arXiv preprint arXiv:2504.07347*, 2025.

[27] Q. Liang and E. Modiano. Minimizing queue length regret under adversarial network models. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 2(1):1–32, 2018.

[28] M. J. Luczak and C. McDiarmid. On the maximum queue length in the supermarket model. *The Annals of Probability*, 34(2):538–598, 2006.

[29] Y. Ma and S. T. Maguluri. Convergence rate analysis of the Join-the-Shortest-Queue system. *arXiv preprint arXiv:2503.15736*, 2025.

[30] S. T. Maguluri and R. Srikant. Heavy traffic queue length behavior in a switch under the Maxweight algorithm. *Stochastic Systems*, 6(1):211–250, 2016.

[31] S. T. Maguluri and R. Srikant. Heavy-traffic queue length behavior in a switch under the MaxWeight algorithm. *Stochastic Systems*, 6(1):1–34, 2016. doi: 10.1214/15-SSY193.

[32] S. T. Maguluri, S. K. Burle, and R. Srikant. Optimal heavy-traffic queue length scaling in an incompletely saturated switch. *Queueing Systems*, 88(3–4):279–309, 2018. doi: 10.1007/s11134-017-9562-x.

[33] A. Mandelbaum and A. L. Stolyar. Scheduling flexible servers with convex delay costs: Heavy-traffic optimality of the generalized cμ-rule. *Operations Research*, 52(6):836–855, 2004.

[34] L. Massoulié and J. Roberts. Bandwidth sharing: Objectives and algorithms. In *IEEE INFOCOM'99. Conference on Computer Communications. Proceedings. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. The Future is Now (Cat. No. 99CH36320)*, volume 3, pages 1395–1403. IEEE, 1999.

[35] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer Science & Business Media, 2012.

[36] M. Mundhenk, J. Goldsmith, C. Lusena, and E. Allender. The complexity of finite-horizon Markov decision process problems. *Journal of the ACM*, 47(4):681–720, 2000.

[37] M. J. Neely. *Stochastic Network Optimization with Application to Communication and Queueing Systems*. Morgan & Claypool Publishers, 2010.

[38] C. H. Papadimitriou and J. N. Tsitsiklis. The complexity of Markov decision processes. *Mathematics of Operations Research*, 12(3):441–450, 1987.

[39] Y. Qiao, S. Anzai, S. Yu, H. Ma, Y. Wang, M. Kim, and H. Xu. Conserve: Harvesting GPUs for low-latency and high-throughput large language model serving. *arXiv preprint arXiv:2410.01228*, 2024.

[40] M. I. Reiman. Open queueing networks in heavy traffic. *Mathematics of Operations Research*, 9 (3):441–458, 1984.

[41] P. Robert. *Stochastic Networks and Queues*, volume 52. Springer Science & Business Media, 2013.

[42] T. Stahlbuhk, B. Shrader, and E. Modiano. Learning algorithms for minimizing queue length regret. *IEEE Transactions on Information Theory*, 67(3):1759–1781, 2021.

[43] A. L. Stolyar. Maxweight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic. *The Annals of Applied Probability*, 14(1):1–53, 2004.

[44] L. Tassiulas and P. P. Bhattacharya. Allocation of interdependent resources for maximal throughput. *Stochastic Models*, 16(1):27–48, 2000.

[45] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. In *29th IEEE Conference on Decision and Control*, 1990.

[46] J. S. Van Leeuwaarden and C. Knessl. Transient behavior of the Halfin–Whitt diffusion. *Stochastic Processes and their Applications*, 121(7):1524–1545, 2011.

[47] A. Verma, L. Pedrosa, M. Korupolu, D. Oppenheimer, E. Tune, and J. Wilkes. Large-scale cluster management at Google with Borg. In *Proceedings of the Tenth European Conference on Computer Systems (EuroSys)*, 2015.

[48] M. J. Wainwright. *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2019.

[49] A. Wald. Statistical decision functions which minimize the maximum risk. *Annals of Mathematics*, 46(2):265–280, 1945.

[50] N. Walton and K. Xu. Learning and information in stochastic networks and queues. In *Tutorials in Operations Research: Emerging Optimization Methods and Modeling Techniques with Applications*, pages 161–198. INFORMS, 2021.

[51] I. Waudby-Smith, M. Larsson, and A. Ramdas. Nonasymptotic and distribution-uniform Komlós–Major–Tusnády approximation. *arXiv preprint arXiv:2502.06188*, 2025.

[52] Z. Yang, R. Srikant, and L. Ying. Learning while scheduling in multi-server systems with unknown statistics: Maxweight with discounted UCB. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 206, pages 4275–4312. PMLR, 2023.

A Technical Appendices and Supplementary Material

A.1 Quadratic Objective for Queue Lengths

We now show that all the results established for the total queue length objective can be extended to the quadratic objective

$$\inf_{\Phi} \sup_{(\mathcal{D},A,S)\in\mathcal{M}^{\rho}(B,C_{\mathrm{a}},C_{\mathrm{d}})} \mathbb{E}^{\Phi}_{(A,S)}\left[\sum_{i=1}^{n} Q_i(T)^2\right].$$

For the lower bound, applying the Cauchy–Schwarz and Jensen's inequalities yields

$$\mathbb{E}^{\Phi}_{(A,S)}\left[\sum_{i=1}^{n} Q_i(T)^2\right] \geq \frac{1}{n}\,\mathbb{E}^{\Phi}_{(A,S)}\left[\sum_{i=1}^{n} Q_i(T)\right]^2 \geq \frac{1}{n}\left(\mathbb{E}^{\Phi}_{(A,S)}\left[\sum_{i=1}^{n} Q_i(T)\right]\right)^2.$$

Thus, the corresponding lower bound follows directly from the results for the total queue-length objective. For the upper bound, both the LyapOpt and MaxWeight policies are analyzed using a uniform Lyapunov method with the quadratic Lyapunov function $V(x) = \|x\|_2^2$. Hence, the corresponding results are directly obtained for $\sum_{i=1}^{n} Q_i(T)^2$.

Consequently, even under the quadratic objective, the LyapOpt policy achieves the finite-time lower bound up to a constant factor, provided that the conditions of Theorem 2 are satisfied, thereby confirming that the optimality of LyapOpt extends beyond the total queue-length objective.

A.2 Lower Bound Inequalities

This section presents several auxiliary results supporting the analysis in Section 3. Among them, Lemma 3 is used directly in the main proof, while the first two lemmas are utilized in its derivation.

We begin with a lower bound for the expected maximum of a drifted Brownian motion.

Lemma 1 (Lower Bound for the Maximum of a Drifted Brownian Motion)**.** *Let $Y_t = \sigma W_t - \theta t$ with $\theta > 0$, $\sigma > 0$, and W a standard Brownian motion. For $t > 0$, define $M_t := \sup_{0\leq s\leq t} Y_s$. Then*

$$\mathbb{E}[M_t] \geq \underline{c}\min\left\{\sigma\sqrt{t}, \sigma^2/\theta\right\}, \qquad \underline{c} = f_{\mathcal{N}}(1) + 2F_{\mathcal{N}}(1) - \tfrac{3}{2} \approx 0.425, \tag{25}$$

where $f_{\mathcal{N}}$ and $F_{\mathcal{N}}$ denote the standard normal density and cumulative distribution functions, respectively.

Proof. The distribution of the maximum of a Brownian motion with drift (see, e.g., Section 1.9 of [18]) is given by

$$\mathbb{P}(M_t \leq x) = F_{\mathcal{N}}\Big(\frac{x+\theta t}{\sigma\sqrt{t}}\Big) - e^{-\frac{2\theta x}{\sigma^2}} F_{\mathcal{N}}\Big(\frac{-x+\theta t}{\sigma\sqrt{t}}\Big), \quad x \geq 0.$$

Integrating the tail probability yields, for $r = \theta\sqrt{t}/\sigma$,

$$\mathbb{E}[M_t] = \sigma\sqrt{t}A(r) + \frac{\sigma^2}{\theta}B(r), \text{ where } A(r) := f_{\mathcal{N}}(r) - r(1 - F_{\mathcal{N}}(r)),\ B(r) := F_{\mathcal{N}}(r) - 1/2.$$

Define the functions

$$g(r) := \frac{\mathbb{E}[M_t]}{\sigma\sqrt{t}} = A(r) + \frac{B(r)}{r}, \qquad h(r) := \frac{\mathbb{E}[M_t]}{\sigma^2/\theta} = r\,A(r) + B(r), \qquad r > 0.$$

We will show that $g(r)$ is decreasing on the interval $(0, 1]$ and $h(r)$ is increasing on the interval $[1, \infty)$.

Since $f_{\mathcal{N}}$ decreases on $[0,\infty)$,

$$B(r) = \int_0^r f_{\mathcal{N}}(u)\mathrm{d}u \geq r f_{\mathcal{N}}(r), \qquad r > 0. \tag{26}$$

Differentiating $g(r)$ gives

$$g'(r) = A'(r) + \frac{B'(r)r - B(r)}{r^2} = -(1 - F_{\mathcal{N}}(r)) + \frac{f_{\mathcal{N}}(r)r - B(r)}{r^2},$$

where the last equality follows from the identity $f'_{\mathcal{N}}(r) = -r f_{\mathcal{N}}(r)$. By (26), $f_{\mathcal{N}}(r)r - B(r) \leq 0$. Thus $g(r)$ is decreasing on $(0,1]$, and $\inf_{0<r\leq 1} g(r) = g(1)$.

Let the Mills ratio be $m(r) := \frac{1-F_{\mathcal{N}}(r)}{f_{\mathcal{N}}(r)}$ for $r > 0$. Then

$$h'(r) = A(r) + rA'(r) + B'(r) = 2f_{\mathcal{N}}(r)(1 - rm(r)).$$

By the Mills inequality $rm(r) \leq 1$, we have $h'(r) \geq 0$ for all $r > 0$. Hence $h(r)$ is increasing on $[1,\infty)$, and $\inf_{r\geq 1} h(r) = h(1)$. A direct computation gives

$$g(1) = h(1) = A(1) + B(1) = f_{\mathcal{N}}(1) + 2F_{\mathcal{N}}(1) - \tfrac{3}{2}.$$

Set

$$\underline{c} := g(1) = h(1) = f_{\mathcal{N}}(1) + 2F_{\mathcal{N}}(1) - \tfrac{3}{2} \approx 0.425.$$

For $r \leq 1$, i.e., $t \leq (\sigma/\theta)^2$,

$$\mathbb{E}[M_t] \geq \sigma\sqrt{t} g(r) \geq \underline{c}\sigma\sqrt{t}.$$

For $r \geq 1$, i.e., $t \geq (\sigma/\theta)^2$,

$$\mathbb{E}[M_t] \geq (\sigma^2/\theta) h(r) \geq \underline{c}\sigma^2/\theta.$$

Combining the two regimes yields (25). □

The following lemma follows from the Komlós–Major–Tusnády (KMT) coupling theorem, which couples partial sums of i.i.d. random variables with a Brownian motion. For a random variable Z, define

$$\bar{\delta} := \sup\left\{\delta \geq 0 : \delta\mathbb{E}\left[|Z|^3 e^{\delta|Z|}\right] \leq \mathrm{Var}(Z)\right\}.$$

If $\bar{\delta} > 0$, we say that Z admits a Sakhanenko parameter $\bar{\delta}$.

Lemma 2 (KMT Coupling)**.** *Let $\{Z(k)\}_{k\in\mathbb{N}}$ be a sequence of i.i.d. random variables with a non-normal distribution satisfying*

$$\mathbb{E}[Z(1)] = 0, \quad \mathbb{E}[Z(1)^2] = \sigma^2 > 0,$$

and suppose that $Z(1)$ admits a Sakhanenko parameter $\bar{\delta} > 0$. Then there exist i.i.d. random variables $\{\widetilde{Z}(k)\}_{k\in\mathbb{N}}$ each with the same distribution as $Z(1)$, and a standard Brownian motion $W = \{W_t\}_{t\geq 0}$

such that the corresponding partial sum $\widetilde{S}(k) = \sum_{i=1}^{k} \widetilde{Z}(i)$ *satisfy,*

$$\mathbb{E}\left[\max_{0\le k\le t} \left|\widetilde{S}(k) - \sigma W_k\right|\right] \le \bar{c}\log(1+t), \tag{27}$$

where $\bar{c}$ *is a positive constant depending only on* σ *and* $\bar{\delta}$*.*

Moreover, if for each $\alpha \in \mathcal{A}$*, the sequence* $Z^\alpha = \{Z^\alpha(k)\}_{k\in\mathbb{N}}$ *satisfies the above conditions with uniformly bounded variance and a uniformly positive lower bound on the Sakhanenko parameter across* α*, then the same constant* $\bar{c}$ *can be chosen uniformly over* α*.*

Proof of Lemma 2. By Theorem 2.4 in [51], there exist random variables $\{\widetilde{Z}(k)\}_{k\in\mathbb{N}}$ and i.i.d. standard normal random variables $\{V(k)\}_{k\in\mathbb{N}}$ such that, for any $z > 0$ and $0 < \delta < \bar{\delta}$,

$$\mathbb{P}\left(\sup_{k\ge 4} \frac{|\widetilde{S}(k) - \sum_{i=1}^{k} \sigma V(k)|}{\log k} \ge 4z\right) \le 2\sum_{n\ge 1}(1 + \delta\sigma 2^{2^n-1})\exp\{-c\delta z 2^n \log 2\}, \tag{28}$$

where $c > 0$ is an absolute constant. Define

$$\Lambda := \sup_{k\ge 4} \frac{|\widetilde{S}(k) - \sum_{i=1}^{k} \sigma V(k)|}{\log k}.$$

Let $z_0 = \frac{2}{c\delta}$. Then

$$\mathbb{E}[\Lambda] = 4\int_0^\infty \mathbb{P}(\Lambda \ge 4z)\mathrm{d}z \le 4z_0 + 4\int_{z_0}^\infty \mathbb{P}(\Lambda \ge 4z)\mathrm{d}z.$$

By (28) and an application of Tonelli's theorem,

$$\begin{aligned} 4\int_{z_0}^\infty \mathbb{P}(\Lambda \ge 4z)\mathrm{d}z &\le 8\sum_{n\ge 1}\int_{z_0}^\infty (1 + \delta\sigma 2^{2^n-1})e^{-c\delta z 2^n \log 2}\mathrm{d}z \\ &= \frac{8}{c\delta\log 2}\sum_{n\ge 1}\frac{1}{2^n}(1 + \delta\sigma 2^{2^n-1})\, e^{-c\delta z_0 2^n \log 2}. \end{aligned}$$

With $z_0 = \frac{2}{c\delta}$, we have $e^{-c\delta z_0\, 2^n \log 2} = e^{-2\cdot 2^n \log 2} = 2^{-2^{n+1}}$. Hence,

$$\begin{aligned} 4\int_{z_0}^\infty \mathbb{P}(\Lambda \ge 4z)\mathrm{d}z &\le \frac{8}{c\delta\log 2}\sum_{n\ge 1}\left(\frac{2^{-2^{n+1}}}{2^n} + \frac{\delta\sigma}{2}\cdot\frac{2^{2^n}2^{-2^{n+1}}}{2^n}\right) \\ &= \frac{8}{c\delta\log 2}\sum_{n\ge 1}\frac{2^{-2^{n+1}}}{2^n} + \frac{4\sigma}{c\log 2}\sum_{n\ge 1}\frac{2^{-2^n}}{2^n}. \end{aligned} \tag{29}$$

Both series converge due to the super-exponential decay of 2^{-2^n}. Therefore, there exists a constant $\widetilde{c} > 0$ (depending only on σ and δ) such that

$$\mathbb{E}[\Lambda] \le \widetilde{c}.$$

Since $\log k \le \log t$ for all $4 \le k \le t$,

$$\sup_{4\le k\le t}\left|\widetilde{S}(k) - \sum_{i=1}^{k}\sigma V(k)\right| \le \Lambda \log t.$$

Taking expectations, we obtain

$$\mathbb{E}\left[\sup_{4\le k\le t}\left|\widetilde{S}(k) - \sum_{i=1}^{k}\sigma V(k)\right|\right] \le \widetilde{c}\log t.$$

We now control the first three terms. Note that

$$\mathbb{E}\left[\max_{0\le k\le 3}\left|\widetilde{S}(k) - \sum_{i=1}^{k}\sigma V(k)\right|\right] \le \mathbb{E}\left[\max_{0\le k\le 3}|\widetilde{S}(k)|\right] + \mathbb{E}\left[\max_{0\le k\le 3}\left|\sum_{i=1}^{k}\sigma V(i)\right|\right].$$

The partial sums $\{\widetilde{S}_k\}_{k\in\mathbb{N}}$ and $\{\sum_{i=1}^{k} V(i)\}_{k\in\mathbb{N}}$ form square-integrable martingales. Applying Jensen's inequality and Doob's maximal inequality yields

$$\mathbb{E}\left[\max_{0\le k\le 3}|\widetilde{S}(k)|\right] \le \sqrt{\mathbb{E}\left[\max_{0\le k\le 3}\widetilde{S}(k)^2\right]} \le \sqrt{4\mathbb{E}[\widetilde{S}(3)^2]} = 2\sqrt{3}\sigma,$$

$$\mathbb{E}\left[\max_{0\le k\le 3}\left|\sum_{i=1}^{k}V(i)\right|\right] \le \sqrt{\mathbb{E}\left[\max_{0\le k\le 3}\left(\sum_{i=1}^{k}V(i)\right)^2\right]} \le \sqrt{4\mathbb{E}\left[\sum_{i=1}^{3}V(i)^2\right]} = 2\sqrt{3}.$$

Summing the two parts gives

$$\mathbb{E}\left[\max_{0\le k\le 3}\left|\widetilde{S}_k - \sum_{i=1}^{k}\sigma V(i)\right|\right] \le 4\sqrt{3}\sigma.$$

Thus, there exists a constant $\bar{c}$ depending only on σ and $\bar{\delta}$ such that (27) holds, using the fact that the sequence $\{V(k)\}_{k\in\mathbb{N}}$ can be embedded into a standard Brownian motion $W = \{W_t\}_{t\ge 0}$. The second part follows directly from the choice of $\bar{c}$. □

Finally, combining Lemmas 1 and 2, we derive the following lower bound for the maximum of a drifted random walk. This is the key result used in the main proof of Section 3.

Lemma 3 (Lower Bound for the Maximum of a Drifted Random Walk)**.** *Let $\{Z(k)\}_{k\in\mathbb{N}}$ satisfy the same conditions as in Lemma 2. Define the partial sum $S(t) = \left(\sum_{k=1}^{t} Z(k)\right) - \theta t$ with $\theta \ge 0$. Then, for all $t \ge \psi^2$ and for all θ satisfying $\theta \le \sigma/\psi$, the expected maximum satisfies*

$$\mathbb{E}\left[\max_{0\le k\le t} S(k)\right] \ge \frac{\underline{c}}{2}\min\left\{\sigma\sqrt{t}, \frac{\sigma^2}{\theta}\right\}, \tag{30}$$

where ψ is the unique positive solution to $\frac{4(\bar{c}/\sigma+1)}{\underline{c}}\log(1+x) + \frac{8}{\underline{c}} = x$; $\underline{c}$ and $\bar{c}$ are the constants

defined in Lemmas 1 and 2 respectively.

Remark 1. The equation $\frac{4(\bar{c}/\sigma+1)}{\underline{c}}\log(1+x)+\frac{8}{\underline{c}}=x$ admits a unique positive root.

Proof of Lemma 3. Let $X_s := \sigma W_s - \theta s$ and $M_s := \sup_{0\le u\le s} X_u$ for $s>0$, where W denotes a standard Brownian motion. In the following, the index k ranges over integers, while s and u vary over real numbers. For an integer $t\ge 1$, note that

$$\max_{0\le k\le t} S(k) \ge \max_{0\le k\le t} X_k - \max_{0\le k\le t}\left|S(k)-X_k\right| = \max_{0\le k\le t} X_k - \max_{0\le k\le t}\left|\sum_{i=1}^{k} Z(i) - \sigma W_k\right|.$$

Taking expectations and using Lemma 2 yields

$$\mathbb{E}\left[\max_{0\le k\le t} S(k)\right] \ge \mathbb{E}\left[\max_{0\le k\le t} X_k\right] - \bar{c}\log(1+t).$$

Moreover, for all $t\ge 1$,

$$\mathbb{E}\left[\max_{0\le k\le t} X_k\right] \ge \mathbb{E}\left[\max_{0\le s\le t} X_s\right] - \sigma\left(\sqrt{2\log(4t)} + \frac{1}{\sqrt{2\log(4t)}}\right), \tag{31}$$

whose proof will be completed later. Since $\sqrt{2y}\le y+1$ for all $y\ge 0$, we obtain

$$\mathbb{E}\left[\max_{0\le k\le t} S(k)\right] \ge \mathbb{E}[M_t] - ((\bar{c}+\sigma)\log(1+t)+4\sigma).$$

We first consider $\theta>0$. Let $u := t\wedge(\sigma/\theta)^2$. Then

$$\mathbb{E}\left[\max_{0\le k\le t} S(k)\right] \ge \mathbb{E}\left[\max_{0\le k\le u} S(k)\right] \ge \mathbb{E}[M_u] - ((\bar{c}+\sigma)\log(1+u)+4\sigma), \quad u = t\wedge(\sigma/\theta)^2. \tag{32}$$

By Lemma 1, for every $u>0$,

$$\mathbb{E}[M_u] \ge \underline{c}\,\min\left\{\sigma\sqrt{u}, \frac{\sigma^2}{\theta}\right\} = \underline{c}\,\min\left\{\sigma\sqrt{t}, \frac{\sigma^2}{\theta}\right\}. \tag{33}$$

Combining (32) and (33) gives the additive-error inequality

$$\mathbb{E}\left[\max_{0\le k\le t} S(k)\right] \ge \underline{c}\,\min\left\{\sigma\sqrt{t}, \frac{\sigma^2}{\theta}\right\} - \left((\bar{c}+\sigma)\log\left(1+t\wedge(\tfrac{\sigma}{\theta})^2\right)+4\sigma\right).$$

To obtain (30), it suffices to ensure

$$(\bar{c}/\sigma+1)\log\left(1+t\wedge(\tfrac{\sigma}{\theta})^2\right)+4 \le \frac{\underline{c}}{2}\min\left\{\sqrt{t}, \frac{\sigma}{\theta}\right\}.$$

By the definition of ψ, for $t \geq \psi^2$, we have

$$\frac{4(\bar{c}/\sigma + 1)}{\underline{c}} \log(1 + \sqrt{t}) + \frac{8}{\underline{c}} \leq \sqrt{t}. \tag{34}$$

Hence,

$$(\bar{c}/\sigma + 1) \log\left(1 + t \wedge (\tfrac{\sigma}{\theta})^2\right) + 4 \leq (\bar{c}/\sigma + 1) \log(1 + t) + 4 \leq 2(\bar{c}/\sigma + 1) \log(1 + \sqrt{t}) + 4 \leq \frac{\underline{c}}{2}\sqrt{t}.$$

where the second inequality follows from $(1 + \sqrt{t})^2 \geq 1 + t$.

Let $r := \sigma/\theta$. Under the assumption on θ, we have $r \geq \psi$, and thus

$$\frac{4(\bar{c}/\sigma + 1)}{\underline{c}} \log(1 + r) + \frac{8}{\underline{c}} \leq r.$$

Then

$$(\bar{c}/\sigma + 1) \log\left(1 + t \wedge (\tfrac{\sigma}{\theta})^2\right) + 4 \leq (\bar{c}/\sigma + 1) \log(1 + r^2) + 4 \leq 2(\bar{c}/\sigma + 1) \log(1 + r) + 4 \leq \frac{\underline{c}}{2} r.$$

Combining the two cases completes the proof for case $\theta > 0$. When $\theta = 0$,

$$\mathbb{E}[M_t] = \sigma\sqrt{2t/\pi} \geq \sigma\underline{c}\sqrt{t},$$

and, by applying (34), the proof proceeds in the same manner.

Now we prove (31). Let $\widetilde{X}_s := X_k$ and $\widetilde{W}_s := W_k$ for $k \leq s < k + 1$. Note that

$$\max_{0 \leq k \leq t} \widetilde{X}_k = \max_{0 \leq s \leq t} \widetilde{X}_s \geq \max_{0 \leq s \leq t} X_s - \max_{0 \leq s \leq t} \left|X_s - \widetilde{X}_s\right|.$$

It therefore suffices to show that

$$\max_{0 \leq s \leq t} \left|W_s - \widetilde{W}_s\right| \leq \sqrt{2\log(4t)} + \frac{1}{\sqrt{2\log(4t)}}.$$

On each interval $[k, k+1)$, write $s = k + u$ with $u \in [0, 1)$. Let $t \geq 1$ be an integer. By the stationary and independent increments of Brownian motion,

$$\sup_{0 \leq s \leq t} |\widetilde{W}_s - W_s| = \max_{0 \leq k \leq t-1} \sup_{0 \leq u \leq 1} |W_{k+u} - W_k| \stackrel{d}{=} \max_{1 \leq j \leq t} Y_j,$$

where $Y_1, \ldots, Y_t$ are i.i.d. copies of $Y := \sup_{0 \leq u \leq 1} |W_u|$. By the reflection principle and the standard Gaussian tail bound $1 - F_{\mathcal{N}}(x) \leq e^{-x^2/2}$ which holds for all $x \geq 0$, we have

$$\mathbb{P}(Y > x) = \mathbb{P}\left(\sup_{0 \leq u \leq 1} |W_u| > x\right) \leq 2\mathbb{P}\left(\sup_{0 \leq u \leq 1} W_u > x\right) = 4(1 - F_{\mathcal{N}}(x)) \leq 4e^{-x^2/2}, \quad x \geq 0.$$

Let $\widetilde{Y}_t := \max_{1\le j\le t} Y_j$. For any $a > 0$, we decompose

$$\mathbb{E}[\widetilde{Y}_t] = \int_0^\infty \mathbb{P}(\widetilde{Y}_t \ge x)\mathrm{d}x \le a + \int_a^\infty \mathbb{P}(\widetilde{Y}_t \ge x)\mathrm{d}x.$$

Since

$$\mathbb{P}(\widetilde{Y}_t \ge x) \le \sum_{k=1}^{t} \mathbb{P}(Y_k \ge x) = t\mathbb{P}(Y \ge x) \le 4te^{-x^2/2},$$

we have

$$\mathbb{E}[\widetilde{Y}_t] \le a + 4t\int_a^\infty e^{-x^2/2}\mathrm{d}x \le a + \frac{4t}{a}e^{-a^2/2},$$

where the last inequality follows from integration by parts. Choosing $a := \sqrt{2\log(4t)}$ yields

$$\mathbb{E}[\widetilde{Y}_t] \le \sqrt{2\log(4t)} + \frac{1}{\sqrt{2\log(4t)}}.$$

Therefore, we complete the proof. □

A.3 Proofs of Theorems and Supporting Results for Section 3

A.3.1 Auxiliary Lemma

Lemma 4 (Lower Bound on Total Queue Length)**.** *Given the arrival process $\{A(t)\}_{t\in\mathbb{N}}$ and the departure process $\{D(t)\}_{t\in\mathbb{N}_0}$, the total queue length at time $T \ge 1$ satisfies*

$$\sum_{i=1}^n Q_i(T) \ge \max\Bigg\{\sum_{t=1}^{T-1}\Big(\sum_{i=1}^n A_i(t) - \sum_{i=1}^n D_i(t)\Big), \sum_{t=2}^{T-1}\Big(\sum_{i=1}^n A_i(t) - \sum_{i=1}^n D_i(t)\Big), \ldots, \sum_{i=1}^n A_i(T-1) - \sum_{i=1}^n D_i(T-1), 0\Bigg\} + \sum_{i=1}^n A_i(T). \tag{35}$$

Proof. Let $\Delta(t) = A(t) - D(t)$ for all $t \ge 1$. By (1), for $1 \le i \le n$ and $T \ge 1$, we have

$$\begin{aligned}
&Q_i(T) - A_i(T)\\
&= \max\{Q_i(T-1) - D_i(T-1), 0\}\\
&= \max\{\max\{Q_i(T-2) - D_i(T-2), 0\} + A_i(T-1) - D_i(T-1), 0\}\\
&= \max\{Q_i(T-2) - D_i(T-2) + \Delta_i(T-1), \Delta_i(T-1), 0\}\\
&= \max\Big\{Q_i(1) - D_i(1) + \sum_{t=2}^{T-1}\Delta_i(t), \sum_{t=2}^{T-1}\Delta_i(t), \ldots, \Delta_i(T-1), 0\Big\}\\
&= \max\Big\{\sum_{t=1}^{T-1}\Delta_i(t), \sum_{t=2}^{T-1}\Delta_i(t), \ldots, \Delta_i(T-1), 0\Big\},
\end{aligned}$$

where the last equation follows that $Q_i(0) = 0$ and then $Q_i(1) = A_i(1)$. Therefore, we have

$$\begin{aligned}\sum_{i=1}^{n} Q_i(T) &= \sum_{i=1}^{n} \max\left\{ \sum_{t=1}^{T-1} \Delta_i(t), \sum_{t=2}^{T-1} \Delta_i(t), \ldots, \Delta_i(T-1), 0 \right\} + \sum_{i=1}^{n} A_i(T) \\ &\geq \max\left\{ \sum_{t=1}^{T-1} \sum_{i=1}^{n} \Delta_i(t), \sum_{t=2}^{T-1} \sum_{i=1}^{n} \Delta_i(t), \ldots, \sum_{i=1}^{n} \Delta_i(T-1), 0 \right\} + \sum_{i=1}^{n} A_i(T),\end{aligned}$$

which completes the proof. □

A.3.2 Proof of Theorem 1

Proof of Theorem 1. Fix a time-invariant scheduling sequence $\mathcal{D}$ with $\mathcal{D}_t = \mathcal{D}_\star$ for all $t \in \mathbb{N}_0$, and assume that $\mathcal{D}_\star$ satisfies the condition in the model class $\mathcal{M}^\rho(B, C_\mathrm{a}, C_\mathrm{d})$. Let

$$M = \max\left\{ \sum_{i=1}^{n} d_i : d \in \mathcal{D}_\star \right\}$$

denote the maximum total service capacity in a single time slot. In the following, we derive a lower bound for the right-hand-side of inequality (35) by constructing a family of hard instances (A, S). Here the arrival process A and the random departure field S are constructed independently. We focus on the case $C_\mathrm{a}^2 + C_\mathrm{d}^2 > 0$; as the case $C_\mathrm{a}^2 + C_\mathrm{d}^2 = 0$ is straightforward.

Arrival Process: We construct an i.i.d. sequence $\{A(t)\}_{t\in\mathbb{N}}$ over time. At time t,

$$\begin{aligned}&\mathbb{P}(A_1(t) = K_\mathrm{a}\lambda_1, A_2(t) = K_\mathrm{a}\lambda_2, \ldots, A_n(t) = K_\mathrm{a}\lambda_n) = 1/K_\mathrm{a}, \\ &\mathbb{P}(A_1(t) = 0, A_2(t) = 0, \ldots, A_n(t) = 0) = (K_\mathrm{a} - 1)/K_\mathrm{a},\end{aligned}$$

where

$$\lambda/\rho \in \operatorname*{argmin}_{\substack{\gamma\in\Pi(\mathcal{D}_\star) \\ \sum_{i=1}^n \gamma_i = M}} \left\{ \sum_{i=1}^{n} \gamma_i^2 \right\} \quad \text{and} \quad K_\mathrm{a} = \frac{nC_\mathrm{a}^2}{\sum_{i=1}^n \lambda_i^2} + 1. \tag{36}$$

Note that

$$\mathbb{E}[A(t)] = \lambda, \quad \text{and} \quad \mathrm{Var}(A_i(t)) = \frac{(K_\mathrm{a}-1)^2\lambda_i^2 + \lambda_i^2(K_\mathrm{a}-1)}{K_\mathrm{a}} = (K_\mathrm{a}-1)\lambda_i^2. \tag{37}$$

The choice of K_a ensures that $\frac{1}{n}\sum_{i=1}^n \mathrm{Var}(A_i(t)) = C_\mathrm{a}^2$.

Random Departure Field: We construct an independent sequence $\{S_t\}_{t\in\mathbb{N}_0}$ over time. At time t, for each $d \in \mathcal{D}_t$,

$$\mathbb{P}(S_d(t) = K_\mathrm{d} d) = 1/K_\mathrm{d}, \text{ and } \mathbb{P}(S_d(t) = 0) = (K_\mathrm{d} - 1)/K_\mathrm{d},$$

where

$$K_{\mathrm{d}} = \frac{nC_{\mathrm{d}}^2}{\max_{d\in\mathcal{D}_\star}\sum_{i=1}^n d_i^2} + 1. \tag{38}$$

If the scheduling policy selects $d(t) \in \mathcal{D}_t$, the actual departure $D(t)$ is distributed as

$$\begin{aligned}
&\mathbb{P}(D_1(t) = K_{\mathrm{d}}d_1(t), D_2(t) = K_{\mathrm{d}}d_2(t), \ldots, D_n(t) = K_{\mathrm{d}}d_n(t)) = 1/K_{\mathrm{d}},\\
&\mathbb{P}(D_1(t) = 0, D_2(t) = 0, \ldots, D_n(t) = 0) = (K_{\mathrm{d}} - 1)/K_{\mathrm{d}}.
\end{aligned}$$

Similarly, the choice of K_{d} ensures S satisfies the conditions of the model class $\mathcal{M}^\rho(B, C_{\mathrm{a}}, C_{\mathrm{d}})$.

The departures satisfy $\mathbb{E}[D(t)] = d(t)$, and $\{\sum_{i=1}^n D_i(t)\}_{t\in\mathbb{N}_0}$ forms a sequence of binary random variables:

$$\mathbb{P}\left(\sum_{i=1}^n D_i(t) = K_{\mathrm{d}}\sum_{i=1}^n d_i(t)\right) = 1/K_{\mathrm{d}}, \text{ and } \mathbb{P}\left(\sum_{i=1}^n D_i(t) = 0\right) = (K_{\mathrm{d}} - 1)/K_{\mathrm{d}}.$$

Let $\{G(t)\}_{t\in\mathbb{N}}$ be an i.i.d. sequence of random variables defined by

$$G(t) = \frac{M}{\sum_{i=1}^n d_i(t)}\sum_{i=1}^n D_i(t).$$

Since $\sum_{i=1}^n d_i(t) \le M$ for any scheduling policy, then $G(t) \ge \sum_{i=1}^n D_i(t)$. Let $X(t) = \sum_{i=1}^n A_i(t)$. By (35), the queue length satisfies

$$\sum_{i=1}^n Q_i(T) \ge \max\left\{\sum_{t=1}^{T-1}\Big(X(t) - G(t)\Big), \sum_{t=2}^{T-1}\Big(X(t) - G(t)\Big), \ldots, X(T-1) - G(T-1), 0\right\} + X(T).$$

Observe that

$$X(t) - G(t) = (X(t) - \rho M) + (M - G(t)) - (1 - \rho)M.$$

Note that $\{(X(t) - \rho M) - (M - G(t))\}_{t\in\mathbb{N}}$ forms an i.i.d. sequence of random variables, satisfying

$$\begin{aligned}
\mathbb{E}[(X(t) - \rho M) - (M - G(t))] &= 0,\\
\mathrm{Var}((X(t) - \rho M) + (M - G(t))) &= M^2(\rho^2(K_{\mathrm{a}} - 1) + (K_{\mathrm{d}} - 1)),
\end{aligned}$$

where the variance follows from the independence between A and S. Thus, by Lemma 3, we obtain

the following lower bound: for all $T \geq \psi_1^2 + 1$ and $(1-\rho)M \leq \sigma/\psi_1$,

$$\begin{aligned}
&\mathbb{E}\left[\max\left\{\sum_{t=1}^{T-1}\Big(X(t)-G(t)\Big), \sum_{t=2}^{T-1}\Big(X(t)-G(t)\Big), \ldots, X(T-1)-G(T-1), 0\right\}\right] \\
&= \mathbb{E}\left[\max\left\{0, X(1)-G(1), \sum_{t=1}^{2}\Big(X(t)-G(t)\Big), \ldots, \sum_{t=1}^{T-1}\Big(X(t)-G(t)\Big)\right\}\right] \\
&\geq \frac{\underline{c}}{2}\min\left\{M\sqrt{(\rho^2(K_{\mathrm{a}}-1)+(K_{\mathrm{d}}-1))(T-1)}, \frac{M(\rho^2(K_{\mathrm{a}}-1)+(K_{\mathrm{d}}-1))}{(1-\rho)}\right\}, \qquad (39)
\end{aligned}$$

where ψ_1 is the unique positive solution to $\frac{4(\bar{c}/\sigma+1)}{\underline{c}}\log(1+x)+\frac{8}{\underline{c}} = x$; $\sigma = M\sqrt{\rho^2(K_{\mathrm{a}}-1)+(K_{\mathrm{d}}-1)}$, $\underline{c}$ and $\bar{c}$ are the constants defined in Lemmas 1 and 2 respectively.

We now construct a specific scheduling set:

$$\mathcal{D}_\star = \{(x_1, x_2, \ldots, x_n) : 0 \leq x_i \leq B \text{ for } 1 \leq i \leq n\}. \qquad (40)$$

Note that this construction of $\mathcal{D}_\star$ ensures that scheduling sequence $\mathcal{D}$ satisfies the condition in the model class $\mathcal{M}^\rho(B, C_{\mathrm{a}}, C_{\mathrm{d}})$. In this case, we choose $\lambda = (\rho B, \rho B, \ldots, \rho B)$ so that

$$M = nB, \;\; K_{\mathrm{a}} - 1 = \frac{C_{\mathrm{a}}^2}{\rho^2 B^2}, \text{ and } K_{\mathrm{d}} - 1 = \frac{C_{\mathrm{d}}^2}{B^2}. \qquad (41)$$

Substitute (41) into (39). Observe that for $\rho \in [1/2, 1]$, the random variables $(X(t)-\rho M)+(M-G(t))$ have uniformly bounded support across all ρ. Indeed,

$$|(X(t)-\rho M)+(M-G(t))| \leq \big(\rho((K_{\mathrm{a}}-1)\vee 1)+(K_{\mathrm{d}}-1)\vee 1\big)M \leq n\left(\frac{2C_{\mathrm{a}}^2}{B}\vee 1 + \frac{C_{\mathrm{d}}^2}{B}\vee 1\right)$$

and the variance of $(X(t)-\rho M)+(M-G(t))$ is $\sigma^2 = n^2(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)$. Consequently, their Sakhanenko parameters are uniformly bounded away from zero by a positive constant depending only on n, C_{a}, C_{d} and B. By Lemma 2, the constant $\bar{c}$ can therefor be chosen to depend only on n, C_{a}, C_{d} and B.

Let $c_1 = \underline{c}/2$ and define $\psi_1(n, B, C_{\mathrm{a}}, C_{\mathrm{d}})$ as the unique positive solution to $\frac{4(\bar{c}/\sigma+1)}{\underline{c}}\log(1+x)+\frac{8}{\underline{c}} = x$, we recover the result in (4). □

A.4 Proofs of Theorems and Supporting Results for Section 4

A.4.1 Formal Proof of the Upper Bound in (13)

For any scheduling sets, arrival processes and random departure fields in $\mathcal{M}^\rho(B, C_{\mathrm{a}}, C_{\mathrm{d}})$, and any scheduling policy Φ, consider the one-step Lyapunov drift. Here for notational simplicity and without risk of ambiguity, we write $\mathbb{E}^\Phi_{(A,S)}$ as $\mathbb{E}$. First, observe that

$$\mathbb{E}\left[\sum_{i=1}^n (\max\{Q_i(t)-D_i(t),0\})^2 \middle| \mathcal{F}_t\right] \leq \mathbb{E}\left[\sum_{i=1}^n (\max\{Q_i(t)-d_i(t),0\})^2 \middle| \mathcal{F}_t\right] + \sum_{i=1}^n \mathrm{Var}(D_i(t)). \quad (42)$$

To see this, define $g(d_i) = (\max\{x_i - d_i, 0\})^2$. Since $|g'(d_i) - g'(s_i)| \le 2|d_i - s_i|$, the descent lemma yields for any D_i,

$$g(D_i) \le g(d_i) + g'(d_i)(D_i - d_i) + (D_i - d_i)^2.$$

Taking conditional expectations given $\mathcal{F}_t$ gives

$$\mathbb{E}[g(D_i)|\mathcal{F}_t] \le \mathbb{E}[g(d_i)|\mathcal{F}_t] + \mathrm{Var}(D_i).$$

Taking the summation over i completes the proof of (42).

The one-step Lyapunov drift satisfies

$$\begin{aligned}
&\mathbb{E}\left[V\big(Q(t+1) - A(t+1)\big) - V\big(Q(t) - A(t)\big) \mid \mathcal{F}_t\right] \\
&= \mathbb{E}\left[\sum_{i=1}^n (\max\{Q_i(t) - D_i(t), 0\})^2 \middle| \mathcal{F}_t\right] - \sum_{i=1}^n (Q_i(t) - A_i(t))^2 \\
&\le \sum_{i=1}^n \mathbb{E}\left[(\max\{Q_i(t) - d_i(t), 0\})^2 \middle| \mathcal{F}_t\right] + \sum_{i=1}^n \mathrm{Var}(D_i(t)) - \sum_{i=1}^n (Q_i(t) - A_i(t))^2 \\
&\le \sum_{i=1}^n \left(\mathbb{E}\left[(Q_i(t) - d_i(t))^2 \middle| \mathcal{F}_t\right] - Q_i(t)^2 + 2Q_i(t)A_i(t) - A_i(t)^2\right) + \sum_{i=1}^n \mathrm{Var}(D_i(t)) \\
&= \sum_{i=1}^n \mathbb{E}\left[d_i(t)^2 - A_i(t)^2 + 2Q_i(t)\big(A_i(t) - d_i(t)\big) \middle| \mathcal{F}_t\right] + \sum_{i=1}^n \mathrm{Var}(D_i(t)) \\
&= f(Q(t), d(t)) + r(Q(t), A(t)) + \sum_{i=1}^n \mathrm{Var}(D_i(t)),
\end{aligned}$$

where

$$f(Q(t), d) := \underbrace{\mathbb{E}\left[2\sum_{i=1}^n Q_i(t)\big(\lambda_i(t) - d_i\big) \middle| \mathcal{F}_t\right]}_{\text{first-order term}} + \underbrace{\sum_{i=1}^n \big(d_i^2 - \lambda_i(t)^2\big)}_{\text{second-order term}},$$

and

$$r(Q(t), A(t)) = \sum_{i=1}^n \left[2Q_i(t)(A_i(t) - \lambda_i(t)) + \lambda_i(t)^2 - A_i(t)^2\right].$$

Note that

$$\begin{aligned}
&\mathbb{E}\big[Q_i(t)A_i(t)\big] \\
&= \mathbb{E}\left[\big(Q_i(t) - A_i(t)\big)A_i(t)\right] + \mathbb{E}\big[A_i(t)^2\big] \\
&\overset{(a)}{=} \mathbb{E}\big[Q_i(t) - A_i(t)\big]\mathbb{E}\big[A_i(t)\big] + \mathbb{E}\big[A_i(t)^2\big] \\
&\overset{(b)}{=} \mathbb{E}\big[Q_i(t)\big]\lambda_i(t) + \mathrm{Var}(A_i(t)),
\end{aligned}$$

where (a) follows from the independence between $Q_i(t) - A_i(t)$ and $A_i(t)$, and (b) follows from the relation $\mathbb{E}\big[A_i(t)^2\big] = \lambda_i(t)^2 + \mathrm{Var}(A_i(t))$.

Then we have

$$\mathbb{E}[r(Q(t), A(t))] = \sum_{i=1}^{n} \mathrm{Var}(A_i(t)).$$

By taking expectation and summing over for $1 \le t \le T-1$, we have

$$\begin{aligned}
&\mathbb{E}[V(Q(T) - A(T)] \\
&\le \mathbb{E}[V(Q(1) - A(1))] + \sum_{t=1}^{T-1} \mathbb{E}[f(Q(t), d(t))] + \sum_{t=1}^{T-1}\sum_{i=1}^{n} (\mathrm{Var}(A_i(t)) + \mathrm{Var}(D_i(t))) \\
&\le \sum_{t=1}^{T-1} \mathbb{E}[f(Q(t), d(t))] + (T-1)n(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2),
\end{aligned} \tag{43}$$

where the last inequality follows from the relation $Q(1) = A(1)$. Note that

$$\begin{aligned}
\left(\mathbb{E}\Big[\sum_{i=1}^{n} \big(Q_i(T) - A_i(T)\big)\Big]\right)^2 &\overset{(c)}{\le} \mathbb{E}\left[\Big(\sum_{i=1}^{n} \big(Q_i(T) - A_i(T)\big)\Big)^2\right] \\
&\overset{(d)}{\le} n\mathbb{E}\left[\sum_{i=1}^{n} \big(Q_i(T) - A_i(T)\big)^2\right] \\
&= n\mathbb{E}\big[V\big(Q(T) - A(T)\big)\big].
\end{aligned}$$

Here (c) follows from Jensen's inequality and (d) follows from the Cauchy–Schwartz inequality. Substituting this expression into (43), we have

$$\mathbb{E}\left[\sum_{i=1}^{n} Q_i(T)\right] \le n\sqrt{\sum_{t=1}^{T-1} \frac{1}{n}\, \mathbb{E}[f(Q(t), d(t))] + (T-1)(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)} + \sum_{i=1}^{n} \mathbb{E}[A_i(T)].$$

A.4.2 Formal Proof of Theorem 4

Proof of Theorem 4. **Step 1.** Fix an instance $(\mathcal{D}, A, S)$ in the model class $\mathcal{M}^1(B, C_{\mathrm{a}}, C_{\mathrm{d}})$. For $\rho \in (0, 1]$, let $\{Q^{(\rho)}(t)\}_{t\in\mathbb{N}_0}$ and $\{d^{(\rho)}(t)\}_{t\in\mathbb{N}_0}$ denote the queue length process and the corresponding MaxWeight decisions when the arrival process $\{A(t)\}_{t\in\mathbb{N}}$ is replaced by the

scaled process $\{\rho A(t)\}_{t\in\mathbb{N}}$. For brevity, when $\rho = 1$, we write $\{Q(t)\}_{t\in\mathbb{N}_0}$ and $\{d(t)\}_{t\in\mathbb{N}_0}$ in place of $\{Q^{(1)}(t)\}_{t\in\mathbb{N}_0}$ and $\{d^{(1)}(t)\}_{t\in\mathbb{N}_0}$, respectively. Suppose that each scheduling set $\mathcal{D}_t$ is a polytope and that the MaxWeight policy yields a unique maximizer of $\langle Q(t), d\rangle$ at every time (with the exception of $t = 0$, where we fix an arbitrary schedule). Under this assumption, we establish that

$$\mathbb{E}_{(A,S)}\left[\sum_{i=1}^{n} Q_i^{(\rho)}(T)\right]$$

is continuous in ρ on an interval $(1-\delta, 1]$ for some $\delta > 0$.

By definition, the MaxWeight policy always selects an extreme point of the scheduling set. Since the extreme points of the polytope is finite, there exists $\delta > 0$ such that for $\rho \in (1-\delta, 1]$, the maximizer of $\langle Q^{(\rho)}(t), d\rangle$ remains unchanged, that is, $d^{(\rho)}(t) = d(t)$ for all $t \le T$ when $\rho \in (1-\delta, 1]$. Consequently, for every sample path, $\sum_{i=1}^{n} Q_i^{(\rho)}(t)$ is continuous in ρ on the interval $(1-\delta, 1]$. Moreover, because

$$\sum_{i=1}^{n} Q_i^{(\rho)}(T) \le \sum_{t=1}^{T}\sum_{i=1}^{n} A_i(T),$$

the family $\{\sum_{i=1}^{n} Q_i^{(\rho)}(T)\}_{\rho\in(1-\delta,1]}$ is uniformly integrable. By the dominated convergence theorem, the expectation $\mathbb{E}_{(A,S)}[\sum_{i=1}^{n} Q_i^{(\rho)}(T)]$ is continuous in ρ on the interval $(1-\delta, 1]$.

Step 2. We begin by considering a family of instances $\mathcal{I}_{b,1}$, which is given in (20) and illustrated in Figure 2 for the case $q = 1$. For the instance $\mathcal{I}_{b,1}$ with $b \ge 6$, we can show (as done in the proof following Step 5) that for all time horizons T satisfying $\left\lceil \frac{b^2}{b-1} \right\rceil \le T \le \left\lceil (\frac{b}{2} - 1)^3 \right\rceil + 1$,

$$\mathbb{E}_{(A,S)}\left[\sum_{i=1}^{2} Q_i(T)\right] \ge \frac{\sqrt{bT}}{2\sqrt{2}}. \tag{44}$$

Step 3. We now consider the family of instances $\mathcal{I}_{b,q}$, which is given in (20) and illustrated in Figure 2. To ensure that the instance lies within the model class $\mathcal{M}^1(B, C_{\mathrm{a}}, C_{\mathrm{d}})$, we require

$$\frac{q^2}{2} \le B^2, \text{ and } \frac{(qb)^2}{2} \le B^2.$$

Recalling that $b > 1$, both conditions are satisfied by setting $q = \frac{\sqrt{2}B}{b}$. It then follows that $\mathcal{I}_{b,\frac{\sqrt{2}B}{b}} \in \mathcal{M}^1(B, C_{\mathrm{a}}, C_{\mathrm{d}})$ for all $b > 1$. Consequently, the queue lengths under the MaxWeight policy for the instance $\mathcal{I}_{b,\frac{\sqrt{2}B}{b}}$ are exactly the queue lengths for the instance $\mathcal{I}_{b,1}$ multiplied by the scaling factor $q = \frac{\sqrt{2}B}{b}$, which can be shown by induction. Suppose $\widetilde{Q}(t) = qQ(t)$. Since MaxWeight selects $d(t) \in \arg\max_{d\in\mathcal{D}_t}\langle Q(t), d\rangle$, it follows that

$$qd(t) \in \operatorname*{argmax}_{d\in q\mathcal{D}_t}\langle Q(t), d\rangle = \operatorname*{argmax}_{d\in q\mathcal{D}_t}\langle \widetilde{Q}(t), d\rangle.$$

Therefore, recalling $D(t) = d(t)$, the queue dynamics satisfy

$$\widetilde{Q}(t+1) = \max\{\widetilde{Q}(t) - qD(t), \mathbf{0}\} + qA(t) = q(\max\{Q(t) - D(t), \mathbf{0}\} + A(t)) = qQ(t+1).$$

By induction, this holds for all $t \in \mathbb{N}$. As a result, applying inequality (44), for any instance $\mathcal{I}_{b,\frac{\sqrt{2}B}{b}}$ with $b \geq 6$, and for all time horizons satisfying $\left\lceil \frac{b^2}{b-1} \right\rceil \leq T \leq \left\lceil (\frac{b}{2} - 1)^3 \right\rceil + 1$, we have

$$\mathbb{E}_{(A,S)} \left[\sum_{i=1}^{2} Q_i(T) \right] \geq \frac{B\sqrt{T}}{2\sqrt{b}}. \tag{45}$$

Step 4. For any $T \geq 9$, we can choose the instance $\mathcal{I}_{b,\frac{\sqrt{2}B}{b}}$ with $b \geq 6$ such that $T = \left\lceil (\frac{b}{2} - 1)^3 \right\rceil + 1$. For this instance, the total queue length at time T satisfies inequality (45). Moreover, since $T = \left\lceil (\frac{b}{2} - 1)^3 \right\rceil + 1$, it follows that $b \leq 3T^{\frac{1}{3}}$, and hence

$$\mathbb{E}_{(A,S)} \left[\sum_{i=1}^{2} Q_i(T) \right] \geq \frac{BT^{\frac{1}{3}}}{2\sqrt{3}}.$$

Therefore, for all $T \geq 9$,

$$\sup_{\{\mathcal{I}_{b,\frac{\sqrt{2}B}{b}} : b \geq 6\}} \mathbb{E}_{(A,S)} \left[\sum_{i=1}^{2} Q_i(T) \right] \geq \frac{BT^{\frac{1}{3}}}{2\sqrt{3}}.$$

By the continuity of $\mathbb{E}_{(A,S)} \left[\sum_{i=1}^{2} Q_i^{(\rho)}(T) \right]$ in ρ on the interval $(1-\delta, 1]$ for each instance $\mathcal{I}_{b,\frac{\sqrt{2}B}{b}}$, we further obtain

$$\liminf_{\rho \uparrow 1} \sup_{\{\mathcal{I}_{b,\frac{\sqrt{2}B}{b}} : b \geq 6\}} \mathbb{E}_{(A,S)} \left[\sum_{i=1}^{2} Q_i^{(\rho)}(T) \right] \geq \sup_{\{\mathcal{I}_{b,\frac{\sqrt{2}B}{b}} : b \geq 6\}} \lim_{\rho \uparrow 1} \mathbb{E}_{(A,S)} \left[\sum_{i=1}^{2} Q_i^{(\rho)}(T) \right] \geq \frac{BT^{\frac{1}{3}}}{2\sqrt{3}}.$$

Since $\mathcal{I}_{b,\frac{\sqrt{2}B}{b}} \in \mathcal{M}^1(B, C_{\mathrm{a}}, C_{\mathrm{d}})$ for all $b > 1$, we conclude that for all $T \geq 9$,

$$\liminf_{\rho \uparrow 1} \sup_{\mathcal{M}^\rho(B, C_{\mathrm{a}}, C_{\mathrm{d}})} \mathbb{E}_{(A,S)} \left[\sum_{i=1}^{2} Q_i(T) \right] \geq \liminf_{\rho \uparrow 1} \sup_{\mathcal{M}^1(B, C_{\mathrm{a}}, C_{\mathrm{d}})} \mathbb{E}_{(A,S)} \left[\sum_{i=1}^{2} Q_i^{(\rho)}(T) \right] \geq \frac{BT^{\frac{1}{3}}}{2\sqrt{3}}, \tag{46}$$

where the first inequality uses that scaling the arrival process of any instance $(\mathcal{D}, A, S) \in \mathcal{M}^1(B, C_{\mathrm{a}}, C_{\mathrm{d}})$ by ρ yields an instance in $\mathcal{M}^\rho(B, C_{\mathrm{a}}, C_{\mathrm{d}})$.

Step 5. Combining the bound in (46) with the lower bound in (4) from Theorem 1, we obtain the

following: for all $T \geq \psi_1(n, B, C_{\mathrm{a}}, C_{\mathrm{d}})^2 + 9$, we have

$$\liminf_{\rho \uparrow 1} \sup_{\mathcal{M}^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}})} \mathbb{E}_{(A,S)} \left[\sum_{i=1}^{2} Q_i(T) \right] \geq \max \left\{ 2c_1 \sqrt{(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)(T-1)}, \frac{BT^{\frac{1}{3}}}{2\sqrt{3}} \right\}$$
$$\geq c_1 \sqrt{(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)(T-1)} + \frac{BT^{\frac{1}{3}}}{4\sqrt{3}},$$

which completes the proof of the theorem.

We now prove inequality (44). Before presenting the formal argument, we first outline the main idea.

Proof Idea for Inequality (44). The choices of b and ε ensure that no two schedules tie under MaxWeight at finite time, so the policy always selects either $(0,1)$ or $(b,0)$. Under MaxWeight, one can show: At each time t, MaxWeight selects $(b,0)$ unless $\frac{Q_2(t)}{Q_1(t)} \geq b$. As a result, $Q_2(t)$ must accumulate to b before it forces the policy to use $(0,1)$; however, doing so causes $Q_1(t)$ to increase to 2, requiring $Q_2(t)$ to build up to $2b$ before $(0,1)$ can be used again more frequently. More precisely, let $t_0 = 0$ and

$$t_k = \min\{t : Q_2(t) \geq kb\} \text{ for } 1 \leq k \leq \left\lfloor \frac{b}{2} \right\rfloor.$$

Under the MaxWeight policy, we observe the following behavior: when $t_0 \leq t < t_1$, $d(t) = (b,0)$, so the average growth rate of Q_2 is $\frac{b-1}{b}$; when $t_1 \leq t < t_2$, $Q_2(t) \geq b-1$, and $d(t)$ alternates between $(b,0)$ and $(0,1)$ until $Q_2(t) = 2b$, the average growth rate of Q_2 is $\frac{b-1}{b} - \frac{1}{2}$; when $t_2 \leq t < t_3$, $Q_2(t) \geq 2b-1$, and each use of $(b,0)$ is followed by two uses of $(0,1)$ until $Q_2(t) = 3b$, the average growth rate of Q_2 is $\frac{b-1}{b} - \frac{2}{3}$.

More generally, for all $k \leq \lfloor b/2 \rfloor$, during the interval $t_{k-1} \leq t < t_k$, $Q_2(t) \geq (k-1)b - 1$, and each use of $(b,0)$ is followed by k uses of $(0,1)$ until $Q_2(t) = kb$, the average growth rate of Q_2 is $\frac{b-1}{b} - \frac{k-1}{k}$. Moreover, since the interval length satisfies $t_k - t_{k-1} = \Theta(kb)$, it follows that $t_k = \Theta(k^2 b)$. Note that for $t_k \leq t < t_{k+1}$, $Q_2(t) \geq kb - 1$. Combining these observations yields $Q_2(t) = \Theta(\sqrt{bt})$.

Formal Proof for Inequality (44). Now we present the formal proof. Consider the MaxWeight policy under the instance (20) with $q = 1$, where $b \geq 6$ is rational and we choose $0 < \varepsilon < \frac{b-1}{b}$ to be a small irrational constant.

By the queue dynamics, $Q(1) = (1, \frac{b-1}{b} - \varepsilon)$. MaxWeight repeatedly selects $(b,0)$ until $Q_2(t) \geq b$. Thus $Q(t) = (1, \frac{b-1}{b}t - \varepsilon)$ for $1 \leq t < \lceil \frac{b(b+\varepsilon)}{b-1} \rceil$. When $t = \lceil \frac{b(b+\varepsilon)}{b-1} \rceil$, we have $Q_2(t) > b = bQ_1(t)$ so MaxWeight choose $(0,1)$. There is no tie in this decision, since $Q_2(t)$ is irrational while $Q_1(t)$ is rational. Note that once $Q_2(t)$ falls below b, MaxWeight will always choose $(b,0)$. Consequently, for all $t \geq \lceil \frac{b(b+\varepsilon)}{b-1} \rceil$, $Q_2(t)$ can never drop below $b-1$. This ensures that the $-\varepsilon$ term is always preserved in $Q_2(t)$, making it irrational for all t. In the following, we consider only time horizons where $Q_1(t) < b$ holds throughout, ensuring that $Q_1(t)$ always remains an integer. Consequently,

there is never a tie in MaxWeight decisions, and it either selects $(b,0)$ or $(0,1)$. Therefore, we henceforth restrict our analysis to the MaxWeight policy under

$$\begin{aligned}
&\mathcal{D}_\star = \{(b,0),(0,1)\};\\
&A(t) = \left(1, \frac{b-1}{b}\right) \text{ for all } t \geq 2, \text{ and } A(1) = \left(1, \frac{b-1}{b} - \varepsilon\right);\\
&S_d(t) = d, \qquad \forall d \in \mathcal{D}_\star,
\end{aligned} \tag{47}$$

where $b \geq 6$ is rational and $0 < \varepsilon < \frac{b-1}{b}$ is an irrational constant. In addition, we choose ε small enough so that $\left\lceil \frac{b(b+\varepsilon)}{b-1} \right\rceil = \left\lceil \frac{b^2}{b-1} \right\rceil$. This is possible because $\frac{b^2}{b-1}$ is not an integer for any $b \geq 6$.

At each time t, the MaxWeight policy will choose $(0,1)$ only when $Q_2(t) > bQ_1(t)$ and $(b,0)$ when $bQ_1(t) > Q_2(t)$. Let $t_0 = 0$ and

$$t_k = \min\{t : Q_2(t) \geq kb\} \text{ for } 1 \leq k \leq \left\lfloor \frac{b}{2} \right\rfloor.$$

We will show, by induction, that

(i) t_k is well defined;

(ii) $Q_1(t_k) = 1$;

(iii) For $t_{k-1} \leq t < t_k$, $1 \leq Q_1(t) \leq k$; $(k-1)b - (k-1)/b \leq Q_2(t) < kb$.

For $k = 1$, by the above discussion, $t_1 = \lceil \frac{b(b+\varepsilon)}{b-1} \rceil$ is well defined, $Q_1(t_1) = 1$, and (iii) naturally holds.

Now suppose that (i)-(iii) hold for some $k-1$ with $k \geq 2$. We show (i)-(iii) also hold for k. Note the fact that at any time t, choosing $(0,1)$ gives

$$Q_1(t+1) = Q_1(t) + 1, Q_2(t+1) = Q_2(t) - \tfrac{1}{b},$$

whereas choosing $(b,0)$ yields

$$Q_1(t+1) = \max\{Q_1(t) - b, 0\} + 1 = 1, Q_2(t+1) = Q_2(t) + 1 - \tfrac{1}{b}.$$

Thus Q_1 increases strictly only when $(0,1)$ is used, and Q_2 increases strictly only when $(b,0)$ is used. Starting from $t \geq t_{k-1}$, and noting $Q_1(t_{k-1}) = 1$, consider the next k time slots. There must be at least one use of $(b,0)$, because if the first $k-1$ departures were all $(0,1)$, then $Q_1(t_{k-1}+k-1) = k$, while $Q_2(t_{k-1}+k-1) < (k-1)b + 1 - (k-1)/b < kb$, forcing the kth departure to be $(b,0)$. Moreover, if only consider $k-1$ time slots after t_{k-1}, at most one $(b,0)$ can be used, because for $s \leq k$, $Q_2(t_{k-1}+s) \geq (k-1)b - s/b > (k-2)b$. Note that $Q_2(t_{k-1}+k) \geq (k-1)b + 1 - k/b > (k-1)b$. By the same reasoning and the fact that $Q_2(t_{k-1}+k+s) \geq (k-1)b + 1 - (k+s)/b > (k-1)b$ for $s \leq k$, in each subsequent block of k slots, Q_2 increases by exactly $1 - k/b > 0$. Hence t_k is well defined, and since Q_2 only increases when $(b,0)$ is applied, we conclude $Q_1(t_k) = 1$. Finally, (iii)

follows directly from the analysis above. And we complete the induction. In addition, we also have the following properties:

1. For all $t_{k-1} \le t \le t_k$,

$$Q_1(t) + Q_2(t) \ge (k-1)b. \tag{48}$$

2. Each interval length satisfies

$$t_1 - t_0 = \left\lceil \frac{b(b+\varepsilon)}{b-1} \right\rceil,$$
$$\left\lceil \frac{k(b-1)}{1-k/b} \right\rceil - 1 \le t_k - t_{k-1} \le \left\lceil \frac{kb}{1-k/b} \right\rceil,\ 2 \le k \le \left\lfloor \frac{b}{2} \right\rfloor.$$

By inequality (48), for all $t_{k-1} \le t \le t_k$,

$$Q_1(t) + Q_2(t) \ge (k-1)b = \sqrt{t}b\frac{k-1}{\sqrt{t}} \ge \sqrt{t}b\frac{k-1}{\sqrt{t_k}}.$$

Note that

$$t_k = \sum_{i=1}^{k}(t_i - t_{i-1}) \le \left\lceil \frac{b(b+\varepsilon)}{b-1} \right\rceil + \sum_{i=2}^{k}\left\lceil \frac{ib}{1-i/b} \right\rceil \le \sum_{i=1}^{k}\left\lceil \frac{ib}{1-i/b} \right\rceil + 1.$$

The remainder of the proof relies on two technical claims.

Claim 1. For $b \ge 4$ and $k \ge 2$,

$$\frac{k-1}{\sqrt{t_k}} \ge \frac{1}{2\sqrt{2b}}.$$

Claim 2. For $b \ge 4$,

$$t_{\lfloor \frac{b}{2} \rfloor} \ge (\frac{b}{2} - 1)^3 + 1.$$

With the help of Claim 1, for $t_1 \le t \le t_{\lfloor \frac{b}{2} \rfloor}$,

$$Q_1(t) + Q_2(t) \ge \frac{\sqrt{bt}}{2\sqrt{2}}.$$

Since $t_1 = \left\lceil \frac{b(b+\varepsilon)}{b-1} \right\rceil \le \frac{b^2}{b-1} + 1$, and when $b \ge 6$, we have $\frac{b^2}{b-1} + 1 \le \left\lceil (\frac{b}{2} - 1)^3 \right\rceil + 1$. Thus, for $b \ge 6$, for all $\left\lceil \frac{b^2}{b-1} \right\rceil \le t \le \left\lceil (\frac{b}{2} - 1)^3 \right\rceil + 1$,

$$Q_1(t) + Q_2(t) \ge \frac{\sqrt{bt}}{2\sqrt{2}}.$$

Substituting $b = \sqrt{2}B$ completes the proof.

□

We now proceed to prove these two claims.

Proof of Claim 1. For $2 \le k \le \lfloor b/2 \rfloor$ and each $1 \le i \le k$, we have $1 - i/b \ge 1/2$. Hence

$$\left\lceil \frac{ib}{1 - i/b} \right\rceil \le \frac{ib}{1 - i/b} + 1 \le 2ib + 1.$$

Summing over $i = 1, 2, \ldots, k$ yields

$$t_k \le \sum_{i=1}^{k} (2ib + 1) + 1 = bk(k+1) + k + 1 \le 2bk^2.$$

The last inequality holds under the assumptions $b \ge 4$ and $k \ge 2$. Hence

$$\frac{k-1}{\sqrt{t_k}} \ge \frac{k-1}{k\sqrt{2b}} = \frac{1 - \frac{1}{k}}{\sqrt{2b}} \ge \frac{1}{2\sqrt{2b}},$$

where the final inequality uses $k \ge 2$, which implies $1 - \frac{1}{k} \ge \frac{1}{2}$. □

Proof of Claim 2. Note that

$$t_{\lfloor \frac{b}{2} \rfloor} = \sum_{i=1}^{\lfloor \frac{b}{2} \rfloor} (t_i - t_{i-1}) \ge \sum_{i=1}^{\lfloor \frac{b}{2} \rfloor} \left(\left\lceil \frac{i(b-1)}{1 - i/b} \right\rceil - 1 \right) + 1.$$

And

$$\sum_{i=1}^{k} \left(\left\lceil \tfrac{i(b-1)}{1-i/b} \right\rceil - 1 \right) + 1 \ge \sum_{i=1}^{k} \left(\frac{ib(b-1)}{b-i} - 1 \right) + 1 \ge \frac{bk(k+1)}{2} - k + 1.$$

Then

$$\sum_{i=1}^{\lfloor b/2 \rfloor} \left(\left\lceil \tfrac{i(b-1)}{1-i/b} \right\rceil - 1 \right) + 1 \ge \frac{b\lfloor b/2 \rfloor (\lfloor b/2 \rfloor + 1)}{2} - \lfloor b/2 \rfloor + 1 \ge (\frac{b}{2} - 1)^3 + 1.$$

□

A.4.3 Formal Proof of Corollary 1

Proof of Corollary 1. We continue to use the same construction as in (20) with $q = 1$ and $B = b/\sqrt{2}$, under which the lower bound for MaxWeight is given by (44). For LyapOpt, since the arrival rate lies in the scheduling set $\mathcal{D}_\star$ and $C_\mathrm{a} = C_\mathrm{d} = 0$, Theorem 2 and (15) yield

$$\sum_{i=1}^{2} Q_i^{\mathrm{LyapOpt}}(T) = \sum_{i=1}^{2} A_i(T) = 2 - \frac{1}{b} - \varepsilon \le 2 - \frac{1}{\sqrt{2}B},$$

which completes the proof. □

A.5 Proofs of Theorems and Supporting Results for Section 5

A.5.1 Proof of Theorem 5

We invoke the following Foster–Lyapunov lemma (see, e.g., [35]).

Lemma 5 (Foster–Lyapunov Criterion). *Let $\{X_t\}_{t\in\mathbb{N}_0}$ be a Markov chain on a countable state space $\mathcal{X}$. Suppose there exists a function $V:\mathcal{X}\to[0,\infty)$, a finite subset $W\subset\mathcal{X}$, and constants $\epsilon>0$ and $b<\infty$ such that*

$$\mathbb{E}[V(X_{t+1})-V(X_t)\mid X_t=x]\le -\epsilon\mathbb{1}_{\{x\notin W\}}+b\mathbb{1}_{\{x\in W\}}.$$

Then the Markov chain $\{X_t\}_{t\in\mathbb{N}_0}$ is stochastically stable in the sense that it reaches the recurrent states with probability 1, and every recurrent state is positive recurrent.

Proof of Theorem 5. Under the LyapOpt policy, the one-step Lyapunov drift

$$\begin{aligned}
&\mathbb{E}^{\text{LyapOpt}}_{(A,S)}\left[V(Q(t+1))-V(Q(t))\mid Q(t)=x\right]\\
&=\mathbb{E}_{(A,S)}\left[\sum_{i=1}^n(\max\{x_i(t)-D_i^{\text{LyapOpt}}(t),0\}+A_i(t+1))^2\,\middle|\,Q(t)=x\right]-\sum_{i=1}^n x_i(t)^2\\
&\le\sum_{i=1}^n\left(\mathbb{E}\left[(\max\{x_i(t)-D_i^{\text{LyapOpt}}(t),0\})^2\middle|Q(t)=x\right]+2x_i(t)\mathbb{E}[A_i(t+1)]+\mathbb{E}[A_i(t+1)^2]-x_i(t)^2\right),
\end{aligned}$$

where the inequality uses the independence between $A(t+1)$ and $Q(t)$. Use the same argument as in (42) of Appendix A.4.1, we have

$$\begin{aligned}
&\sum_{i=1}^n\mathbb{E}\left[(\max\{x_i(t)-D_i^{\text{LyapOpt}}(t),0\})^2\middle|Q(t)=x\right]\\
&\le\sum_{i=1}^n\mathbb{E}\left[(\max\{x_i(t)-d_i^{\text{LyapOpt}}(t),0\})^2\middle|Q(t)=x\right]+\sum_{i=1}^n\mathrm{Var}(D_i^{\text{LyapOpt}}(t))\\
&=\min_{d\in\mathcal{D}_\star}\sum_{i=1}^n(\max\{x_i(t)-d_i,0\})^2+\sum_{i=1}^n\mathrm{Var}(D_i^{\text{LyapOpt}}(t))\\
&\le\min_{d\in\mathcal{D}_\star}\sum_{i=1}^n(x_i(t)-d_i)^2+\sum_{i=1}^n\mathrm{Var}(D_i^{\text{LyapOpt}}(t))\\
&\le\min_{d\in\mathcal{D}_\star}\sum_{i=1}^n-2x_i(t)d_i+\max_{d\in\mathcal{D}_\star}\sum_{i=1}^n d_i^2+\sum_{i=1}^n x_i^2+\sum_{i=1}^n\mathrm{Var}(D_i^{\text{LyapOpt}}(t)),
\end{aligned}$$

where the equality uses the definition of LyapOpt, and the last inequality is by the fact that

$\min(f+g) \le \min f + \max g$. Therefore,

$$\begin{aligned}
&\mathbb{E}^{\text{LyapOpt}}_{(A,S)}\left[V(Q(t+1)) - V(Q(t)) \mid Q(t) = x\right] \\
&\le \min_{d\in\mathcal{D}_\star} \sum_{i=1}^{n} 2x_i(t)(\lambda_i(t+1) - d_i) + 2nB^2 + nC_{\text{a}}^2 + nC_{\text{d}}^2 \\
&\le -2(1-\rho)\max_{d\in\mathcal{D}_\star} \sum_{i=1}^{n} x_i(t)d_i + 2nB^2 + nC_{\text{a}}^2 + nC_{\text{d}}^2,
\end{aligned}$$

where the last inequality follows from $\lambda(t+1) \in \rho\Pi(\mathcal{D}_\star)$, which implies

$$\sum_{i=1}^{n} x_i(t)\lambda_i(t+1) \le \max_{d\in\mathcal{D}_\star} \rho \sum_{i=1}^{n} x_i(t)d_i.$$

Then we choose

$$W = \left\{ x \in \mathbb{Z}_+^n : \max_{d\in\mathcal{D}_t} \sum_{i=1}^{n} x_i d_i \le \frac{\epsilon + n(2B^2 + C_{\text{a}}^2 + C_{\text{d}}^2)}{2(1-\rho)} \right\}.$$

Since W is finite, it satisfies the condition in Lemma 5, which completes the proof. □

A.5.2 Proof of Proposition 1 and Related Lemma

The next lemma follows directly from Assumption 1.

Lemma 6. *Under Assumption 1, for all $t \in \mathbb{N}_0$ and all $x \in \mathbb{R}_+^n$,*

$$\max_{d\in\mathcal{D}_t} \langle x, d\rangle \ge \frac{\alpha}{n} \sum_{i=1}^{n} x_i. \tag{49}$$

Proof of Lemma 6. Fix $t \in \mathbb{N}_0$ and $x \in \mathbb{R}_+^n$. By Assumption 1, for each j there exists a schedule vector $d^{(j)}(t) \in D_t$ such that its j-th component $d_j^{(j)}(t) \ge \alpha$. Hence

$$\max_{d\in\mathcal{D}_t} \langle x, d\rangle \ge \max_{1\le j\le n} \langle x, d^{(j)}(t)\rangle \ge \max_{1\le j\le n} x_j d_j^{(j)}(t) \ge \alpha \max_{1\le j\le n} x_j.$$

Since $\max_i x_i \ge \frac{1}{n}\sum_{i=1}^n x_i$, inequality (49) follows. □

Proof of Proposition 1. By inequality (10) and the definition of LyapOpt, the one-step Lyapunov

drift under the LyapOpt policy satisfies

$$
\begin{aligned}
&\mathbb{E}_{(A,S)}^{\mathrm{LyapOpt}}\left[V(Q(t+1)-A(t+1))-V(Q(t)-A(t))\mid \mathcal{F}_t\right]\\
&\le \min_{d\in\mathcal{D}_t} f(Q(t),d)+r(Q(t),A(t))+\sum_{i=1}^{n}\mathrm{Var}(D_i^{\mathrm{LyapOpt}}(t))\\
&= \min_{d\in\mathcal{D}_t}\left(2\sum_{i=1}^{n}Q_i(t)\big(\lambda_i(t)-d_i\big)+\sum_{i=1}^{n}\big(d_i^2-\lambda_i(t)^2\big)\right)+r(Q(t),A(t))+\sum_{i=1}^{n}\mathrm{Var}(D_i^{\mathrm{LyapOpt}}(t))\\
&\le \min_{d\in\mathcal{D}_t} 2\sum_{i=1}^{n}Q_i(t)\big(\lambda_i(t)-d_i\big)+nB^2+r(Q(t),A(t))+\sum_{i=1}^{n}\mathrm{Var}(D_i^{\mathrm{LyapOpt}}(t))\\
&\le -2(1-\rho)\max_{d\in\mathcal{D}_t}\sum_{i=1}^{n}Q_i(t)d_i+nB^2+r(Q(t),A(t))+\sum_{i=1}^{n}\mathrm{Var}(D_i^{\mathrm{LyapOpt}}(t)),
\end{aligned}
$$

where the last inequality follows from $\lambda(t)\in\rho\Pi(\mathcal{D}_t)$, which implies

$$
\sum_{i=1}^{n}Q_i(t)\lambda_i(t)\le\max_{d\in\mathcal{D}_t}\rho\sum_{i=1}^{n}Q_i(t)d_i.
$$

Taking expectations and summing over t yields

$$
\mathbb{E}_{(A,S)}^{\mathrm{LyapOpt}}[V(Q(T)-A(T))]\le -2(1-\rho)\sum_{t=1}^{T-1}\max_{d\in\mathcal{D}_t}\sum_{i=1}^{n}\mathbb{E}_{(A,S)}^{\mathrm{LyapOpt}}[Q_i(t)]d_i+n(T-1)(B^2+C_{\mathrm{a}}^2+C_{\mathrm{d}}^2).
$$

By Lemma 6,

$$
\max_{d\in\mathcal{D}_t}\sum_{i=1}^{n}\mathbb{E}_{(A,S)}^{\mathrm{LyapOpt}}[Q_i(t)]d_i\ge\frac{\alpha}{n}\sum_{i=1}^{n}\mathbb{E}_{(A,S)}^{\mathrm{LyapOpt}}[Q_i(t)].
$$

Then we have the uniform stability bound (21). □

A.6 Supplementary Experiments in Section 6.1

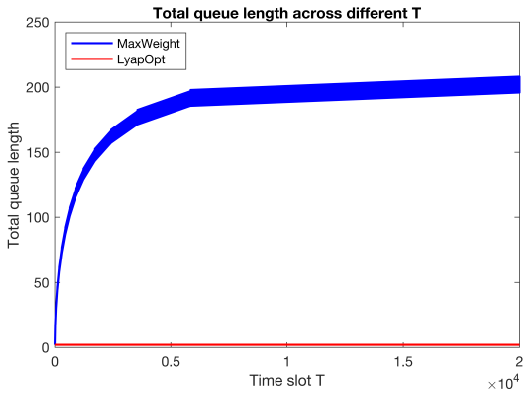
The following two figures confirm that, in the setting of the Corollary 1, the total queue length under MaxWeight eventually becomes bounded (i.e., $O(1)$ in T) as T increases.

A.7 Formal Proofs of Results in Section 7

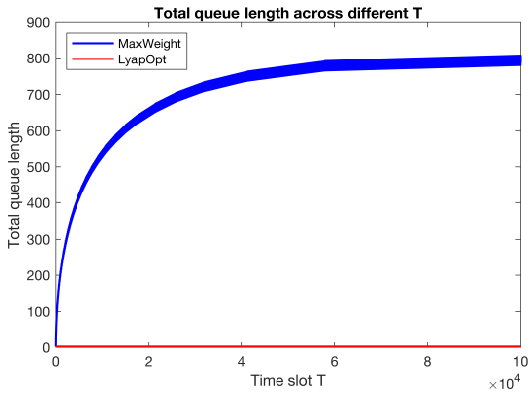
A.7.1 Proof of Proposition 2

Proof of Proposition 2. We focus on the case $C_{\mathrm{a}}^2+C_{\mathrm{d}}^2>0$; as the case $C_{\mathrm{a}}^2+C_{\mathrm{d}}^2=0$ is straightforward. Fix a time-invariant scheduling sequence $\mathcal{D}$ with $\mathcal{D}_t=\mathcal{D}_\star$ for all $t\in\mathbb{N}$, where $\mathcal{D}_\star$ is defined in (40). Let

$$
M=\max\left\{\sum_{i=1}^{n}d_i : d\in\mathcal{D}_\star\right\}=nB
$$



(a) Total queue length when $B = 10$, $C_{\mathrm{a}} = C_{\mathrm{d}} = 0$



(b) Total queue length when $B = 20$, $C_{\mathrm{a}} = C_{\mathrm{d}} = 0$

Figure 9: Performance comparison of MaxWeight and LyapOpt policies versus B

denote the maximum total service capacity in a single time slot. We construct arrival process A and random departure field S are mutually independent. For each t, the components $A_i(t)$ are independent across i and distributed as

$$\mathbb{P}\Big(A_i(t) = K_{\mathrm{a}}\lambda_i\Big) = \frac{1}{K_{\mathrm{a}}}, \quad \mathbb{P}\Big(A_i(t) = 0\Big) = \frac{K_{\mathrm{a}} - 1}{K_{\mathrm{a}}},$$

where the choices of λ and K_{a} follow those specified in (36). Thus, the mean and variance conditions for A in (37) still holds. In particular, the total variance across all queues is given by

$$\sum_{i=1}^{n} \mathrm{Var}(A_i(t)) = \sum_{i=1}^{n} (K_{\mathrm{a}} - 1)\lambda_i^2 = nC_{\mathrm{a}}^2.$$

For each t and $d \in \mathcal{D}_t$, the components of $(S_d(t))_i$ are independent across i and distributed as

$$\mathbb{P}\Big((S_d(t))_i = K_{\mathrm{d}} d_i\Big) = \frac{1}{K_{\mathrm{d}}}, \quad \mathbb{P}\Big((S_d(t))_i = 0\Big) = \frac{K_{\mathrm{d}} - 1}{K_{\mathrm{d}}},$$

where the choice of K_{d} follows those specified in (38). This confirms that the constructed pair (A, S) belongs to the model class $\mathcal{M}_{\mathcal{I}}^{\rho}(B, C_{\mathrm{a}}, C_{\mathrm{d}})$. Recall that $D(t) = S_{d(t)}(t)$. So the components $D_i(t)$ are independent across i and distributed as

$$\mathbb{P}\Big(D_i(t) = K_{\mathrm{d}} d_i(t)\Big) = \frac{1}{K_{\mathrm{d}}}, \quad \mathbb{P}\Big(D_i(t) = 0\Big) = \frac{K_{\mathrm{d}} - 1}{K_{\mathrm{d}}}.$$

Let $\{G(t)\}_{t\in\mathbb{N}}$ be an i.i.d. sequence of random variables defined by $G(t) = \sum_{i=1}^{n} Y_i(t)$, where $Y_i(t)$ are independent across i and distributed as

$$\mathbb{P}\Big(Y_i(t) = K_{\mathrm{d}} B\Big) = \frac{1}{K_{\mathrm{d}}}, \quad \mathbb{P}\Big(Y_i(t) = 0\Big) = \frac{K_{\mathrm{d}} - 1}{K_{\mathrm{d}}},$$

so that

$$\mathbb{P}(G(t) \geq y) \geq \mathbb{P}(\sum_{i=1}^{n} D_i(t) \geq y) \text{ for each } y, \ \mathbb{E}[G(t)] = M, \text{ and } \mathrm{Var}(G(t)) = nC_{\mathrm{d}}^2.$$

The following argument follows the same idea as in the proof of Theorem 1. Let $X(t) = \sum_{i=1}^{n} A_i(t)$. By (35), the queue length satisfies

$$\sum_{i=1}^{n} Q_i(T) \geq \max\left\{\sum_{t=1}^{T-1}\Big(X(t) - G(t)\Big), \sum_{t=2}^{T-1}\Big(X(t) - G(t)\Big), \ldots, X(T-1) - G(T-1), 0\right\} + X(T).$$

Observe that

$$X(t) - G(t) = (X(t) - \rho M) + (M - G(t)) - (1-\rho)M.$$

Note that $\{(X(t) - \rho M) - (M - G(t))\}_{t\in\mathbb{N}}$ forms an i.i.d. sequence of random variables, satisfying

$$\mathbb{E}[(X(t) - \rho M) - (M - G(t))] = 0, \mathrm{Var}((X(t) - \rho M) + (M - G(t))) = nC_{\mathrm{a}}^2 + nC_{\mathrm{d}}^2,$$

where the variance follows from the independence of all components. Thus, by Lemma 3, we obtain the following lower bound: for all $T \geq \psi_2^2 + 1$ and $(1-\rho)M \leq \sigma/\psi_2$,

$$\begin{aligned}
&\mathbb{E}\left[\max\left\{\sum_{t=1}^{T-1}\Big(X(t) - G(t)\Big), \sum_{t=2}^{T-1}\Big(X(t) - G(t)\Big), \ldots, X(T-1) - G(T-1), 0\right\}\right] \\
&= \mathbb{E}\left[\max\left\{0, X(1) - G(1), \sum_{t=1}^{2}\Big(X(t) - G(t)\Big), \ldots, \sum_{t=1}^{T-1}\Big(X(t) - G(t)\Big)\right\}\right] \\
&\geq \frac{\underline{c}}{2}\min\left\{\sqrt{n(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)(T-1)}, \frac{C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2}{B(1-\rho)}\right\},
\end{aligned}$$

where ψ_2 is the unique positive solution to $\frac{4(\bar{c}/\sigma+1)}{\underline{c}}\log(1+x) + \frac{8}{\underline{c}} = x$; $\sigma = \sqrt{nC_{\mathrm{a}}^2 + nC_{\mathrm{d}}^2}$; $\underline{c}$ and $\bar{c}$ are the constants defined in Lemmas 1 and 2 respectively.

Observe that for $\rho \in [1/2, 1]$, the random variables $(X(t) - \rho M) + (M - G(t))$ have uniformly bounded support across ρ,

$$|(X(t) - \rho M) + (M - G(t))| \leq n\left(\frac{2C_{\mathrm{a}}^2}{B} \vee 1 + \frac{C_{\mathrm{d}}^2}{B} \vee 1\right)$$

and variance $\sigma^2 = n(C_{\mathrm{a}}^2 + C_{\mathrm{d}}^2)$. Consequently, their Sakhanenko parameters are uniformly bounded away from zero by a positive constant depending only on n, C_{a}, C_{d} and B. By Lemma 2, the constant $\bar{c}$ can therefore be chosen to depend only on n, C_{a}, C_{d} and B.

Let $c_1 = \underline{c}/2$ and define $\psi_2(n, B, C_{\mathrm{a}}, C_{\mathrm{d}})$ as the unique positive solution to $\frac{4(\bar{c}/\sigma+1)}{\underline{c}}\log(1+x) + \frac{8}{\underline{c}} = x$, we recover the result in (22).

□

A.7.2 Proof of Proposition 5

Proof of Proposition 5. Under the assumption that $\mathbb{E}[A(t)] = \lambda \in \mathcal{D}_\star$, we have

$$\sum_{i=1}^{n} \max\{\widehat{A}_i(t) - d_i^{\mathrm{EP}}(t), 0\} \le \sum_{i=1}^{n} \max\{\widehat{A}_i(t) - \lambda_i, 0\}. \tag{50}$$

Let $\Delta(t) = A(t) - D^{\mathrm{EP}}(t)$. Then

$$\begin{aligned}\Delta(t) &= (A(t) - \lambda) + (\lambda - \widehat{A}(t)) + (\widehat{A}(t) - d^{\mathrm{EP}}(t)) + (d^{\mathrm{EP}}(t) - D^{\mathrm{EP}}(t)) \\ &\triangleq \Delta^{(1)}(t) + \Delta^{(2)}(t) + \Delta^{(3)}(t) + \Delta^{(4)}(t).\end{aligned}$$

Under this policy, for each i,

$$\begin{aligned}Q_i(T) - A_i(T) &= \max\left\{\sum_{t=1}^{T-1} \Delta_i(t), \sum_{t=2}^{T-1} \Delta_i(t), \ldots, \Delta_i(T-1), 0\right\} \\ &\le \sum_{j=1}^{4} \max\left\{\sum_{t=1}^{T-1} \Delta_i^{(j)}(t), \sum_{t=2}^{T-1} \Delta_i^{(j)}(t), \ldots, \Delta_i^{(j)}(T-1), 0\right\}.\end{aligned}$$

For the first term, let $S_i^{(1)}(t) = \sum_{s=T-t}^{T-1} \Delta_i^{(1)}(s)$ and $M_i^{(1)}(t) = \max_{1\le s\le t} S_i^{(1)}(s)$ for $t \in \mathbb{N}$. Specifically, let $S_i^{(1)}(0) = M_i^{(1)}(0) = 0$. since the variables $\{\Delta_i^{(1)}(s)\}_{s\ge 1}$ are independent across s and satisfy $\mathbb{E}[\Delta_i^{(1)}(s)] = 0$, $\{S_i^{(1)}(t)\}_{t=0}^{T-1}$ is a martingale. Thus, we have

$$\begin{aligned}&\mathbb{E}\left[\max\left\{\sum_{t=1}^{T-1} \Delta_i^{(1)}(t), \sum_{t=2}^{T-1} \Delta_i^{(1)}(t), \ldots, \Delta_i^{(1)}(T-1), 0\right\}\right] \\ &= \mathbb{E}\Big[M_i^{(1)}(T-1)\Big] \le \sqrt{\mathbb{E}\Big[\big(M_i^{(1)}(T-1)\big)^2\Big]} \\ &\le \sqrt{4\mathbb{E}\Big[\big(S_i^{(1)}(T-1)\big)^2\Big]} = 2\sqrt{\sum_{t=1}^{T-1} \mathrm{Var}(A_i(1))},\end{aligned}$$

where the first inequality follows from the Jensen's inequality and the second inequality follows from Doob's maximal inequality. Using the same argument, for the forth term, we also have

$$\mathbb{E}\left[\max\left\{\sum_{t=1}^{T-1} \Delta_i^{(4)}(t), \sum_{t=2}^{T-1} \Delta_i^{(4)}(t), \ldots, \Delta_i^{(4)}(T-1), 0\right\}\right] \le 2\sqrt{\sum_{t=1}^{T-1} \mathrm{Var}(D_i(t))}.$$

For the second term,

$$\max\left\{\sum_{t=1}^{T-1} \Delta_i^{(2)}(t), \sum_{t=2}^{T-1} \Delta_i^{(2)}(t), \ldots, \Delta_i^{(2)}(T-1), 0\right\} \le \sum_{t=1}^{T-1} \max\{\Delta_i^{(2)}(t), 0\} \le \sum_{t=1}^{T-1} \big|\Delta_i^{(2)}(t)\big|.$$

Note that

$$\mathbb{E}\Big[\big|\Delta_i^{(2)}(t)\big|\Big] \le \sqrt{\mathbb{E}\Big[\big(\Delta_i^{(2)}(t)\big)^2\Big]} = \sqrt{\frac{\mathrm{Var}(A_i(1))}{t}}.$$

Then

$$\sum_{t=1}^{T-1} \big|\Delta_i^{(2)}(t)\big| \le \sqrt{\mathrm{Var}(A_i(1))} \sum_{s=1}^{T-1} \frac{1}{\sqrt{t}} \le \sqrt{\mathrm{Var}(A_i(1))} \int_{s=1}^{T-1} \frac{1}{\sqrt{t}} \mathrm{d}t \le 2\sqrt{(T-1)\,\mathrm{Var}(A_i(1))}.$$

For the third term,

$$\begin{aligned}
&\sum_{i=1}^{n} \max\left\{ \sum_{t=1}^{T-1} \Delta_i^{(3)}(t), \sum_{t=2}^{T-1} \Delta_i^{(3)}(t), \ldots, \Delta_i^{(3)}(T-1), 0 \right\} \\
&\le \sum_{i=1}^{n} \sum_{t=1}^{T-1} \max\{\Delta_i^{(3)}(t), 0\} = \sum_{t=1}^{T-1} \sum_{i=1}^{n} \max\{\widehat{A}_i(s) - d_i^{\mathrm{EP}}(t), 0\} \\
&\le \sum_{t=1}^{T-1} \sum_{i=1}^{n} \max\{\widehat{A}_i(s) - \lambda_i, 0\} \le \sum_{t=1}^{T-1} \sum_{i=1}^{n} |\widehat{A}_i(s) - \lambda_i| \\
&= \sum_{t=1}^{T-1} \sum_{i=1}^{n} \big|\Delta_i^{(2)}(t)\big| \le \sum_{i=1}^{n} 2\sqrt{(T-1)\,\mathrm{Var}(A_i(t))}.
\end{aligned}$$

In sum, we have

$$\begin{aligned}
\mathbb{E}\Big[\sum_{i=1}^{n} Q_i(T)\Big] - \mathbb{E}\Big[\sum_{i=1}^{n} A_i(T)\Big] &\le \sum_{i=1}^{n} 6\sqrt{(T-1)\,\mathrm{Var}(A_i(1))} + 2\sum_{i=1}^{n} \sqrt{\sum_{t=1}^{T-1} \mathrm{Var}(D_i(t))} \\
&\le n\sqrt{T-1}(6C_{\mathrm{a}} + 2C_{\mathrm{d}}).
\end{aligned}$$

□