

# Delay Stability Regions of the Max-Weight Policy under Heavy-Tailed Traffic \*

Mihalis G. Markakis, Eytan Modiano, and John N. Tsitsiklis †

## Abstract

We carry out a delay stability analysis (i.e., determine conditions under which expected steady-state delays at a queue are finite) for a simple 3-queue system operated under the Max-Weight scheduling policy, for the case where one of the queues is fed by heavy-tailed traffic (i.e., when the number of arrivals at each time slot has infinite second moment). This particular system exemplifies an intricate phenomenon whereby heavy-tailed traffic at one queue may or may not result in the delay instability of another queue, depending on the arrival rates.

While the ordinary stability region (in the sense of convergence to a steady-state distribution) is straightforward to determine, the determination of the delay stability region is more involved: (i) we use “fluid-type” sample path arguments, combined with renewal theory, to prove delay instability outside a certain region; (ii) we use a piecewise linear Lyapunov function to prove delay stability in the interior of that same region; (iii) as an intermediate step in establishing delay stability, we show that the expected workload of a stable M/GI/1 queue scales with time as  $\mathcal{O}(t^{1/(1+\gamma)})$ , assuming that service times have a finite  $1 + \gamma$  moment, where  $\gamma \in (0, 1)$ .

---

\*This work was supported by NSF Grants CNS-0915988 and CCF-0728554, and ARO MURI Grant W911NF-08-1-0238. The material in this paper was presented in part at the 31<sup>st</sup> IEEE INFOCOM Conference, Orlando, FL, USA.

†The authors are with the Laboratory for Information and Decision Systems, at the Massachusetts Institute of Technology, Cambridge, MA, USA. Emails: {mihalis,modiano,jnt}@mit.edu.

# 1 Introduction

We consider a simple 3-queue system operated under the Max-Weight scheduling policy, for the case where one of the queues is fed by heavy-tailed traffic (i.e, when the number of arrivals at each time slot has infinite second moment). The performance metric that we are interested in is **delay stability**: a queue is delay stable if the expected steady-state delay in that queue is finite, and delay unstable otherwise.

Under the Max-Weight scheduling policy, the notions of stability and delay stability coincide when all primitive stochastic processes are light-tailed; that is, all queues are delay stable for all arrival rate vectors within the stability region of the system (see, e.g., [21]). On the other hand, for a system of parallel queues served by a single server, it is known that if one of the arrival streams is heavy-tailed, then all queues are delay unstable, for all positive arrival rates [14]. The system that we study in this paper exhibits an intermediate behavior: we show that in the presence of heavy-tailed traffic, there may be nontrivial parts of the stability region where a queue is delay stable, and nontrivial parts of the stability region where the same queue is delay unstable. This exemplifies an intricate “delay instability propagation” whereby heavy-tailed traffic at one queue may or may not result in the delay instability of another queue, depending on the numerical values of the arrival rates.

The study of queueing systems with heavy-tailed traffic is motivated by significant evidence that traffic in real-world networks exhibits strong correlations and statistical self-similarity over different time scales. This observation was first made by Leland *et al.* [13] through an analysis of Ethernet traffic traces. Subsequent empirical studies have documented this phenomenon in other networks, while accompanying theoretical studies have associated it with heavy-tailed arrival processes; see [16] for an overview. At the same time, the performance analysis of Max-Weight scheduling is of independent interest, with a long history and rich literature, see, e.g., [1, 6, 20, 21]. The feature that makes the Max-Weight policy appealing is its throughput optimality, i.e., its ability to stabilize a queueing system whenever this is possible. In other words, dynamic instability phenomena, such as the ones reported in [12, 17], are guaranteed not to occur under Max-Weight scheduling.

The impact of heavy-tailed traffic has been analyzed extensively in the context of single and multi-server queues; see the survey papers [3, 5] and the references therein. However, there is little work on more complex queueing systems and scheduling policies. A notable exception is the work by Borst *et al.* [4], which studies the Generalized Processor Sharing policy in the presence of heavy-tailed traffic. Related results, which come closer to the subject of this paper, are also given in [14], which reports a nontrivial delay stability region for a system of parallel queues served by a single server under the Round-Robin policy. In contrast to the aforementioned scheduling policies, Max-Weight is a queue length-based policy, whose dynamics are more complex and harder to analyze. Finally, the work by Jagannathan *et al.* [11] presents a queueing system with heavy-tailed traffic, and under the Max-Weight policy, with a nontrivial delay stability region. However, what affects delay stability in that study is the intermittent access of servers to queues, arising, e.g., in wireless networks. Consequently, the underlying mechanism that determines the delay stability region there is very different than the one analyzed here.

The main contributions of this paper are as follows:

- (i) we show that under the Max-Weight policy, heavy-tailed arrivals at one queue may or may not cause delay instability at other queues, as determined by an intricate instability propagation phenomenon;
- (ii) we use “fluid-type” sample path arguments, combined with renewal theory, to prove delay instability if the arrival rate to a specific queue is greater than a certain threshold (Proposition 2);
- (iii) we use drift analysis (over a sufficiently long time interval) of a suitably constructed piecewise linear Lyapunov function to prove delay stability if the arrival rate to a specific queue is less than that same threshold (Proposition 3);
- (iv) as an intermediate step in establishing delay stability, we show that the expected workload of a stable M/GI/1 queue with heavy-tailed service times scales sublinearly in time, assuming just the existence of the  $1 + \gamma$  moment of service times (see the proof of Proposition 3).

The rest of the paper is organized as follows. In Section 2 we give a detailed description of

the queueing system under consideration and the probabilistic assumptions that we make, together with some necessary definitions and notation. In Sections 3 and 4 we prove our delay instability and stability results, respectively. We conclude in Section 5 with a brief discussion and future directions of research.

## 2 Model and Definitions

Throughout the paper we denote by  $\mathbb{Z}_+$  and  $\mathbb{N}$  the sets of nonnegative and positive integers, respectively. Similarly, the Cartesian product of  $M$  copies of  $\mathbb{Z}_+$  is denoted by  $\mathbb{Z}_+^M$ . We use  $1_E$  for the indicator variable of event  $E$ . Finally,  $[x]^+$  stands for  $\max\{x, 0\}$ , the nonnegative part of a scalar  $x$ . With few exceptions, we follow the usual convention of using lower case letters to denote real numbers or vectors and upper case letters to denote random variables or events.

### 2.1 The basic model and dynamics

We consider the queueing system depicted in Figure 1, consisting of three single-class, single-server queues, with infinite capacity. Time is slotted and traffic consists of constant size packets. We define  $A_i(t)$  as the (nonnegative) number of packets (a random variable) that queue  $i \in \{1, 2, 3\}$  receives at slot  $t \in \mathbb{Z}_+$ , and let  $A(t) = (A_1(t), A_2(t), A_3(t))$ . We assume that arrivals occur at the end of the time slots. We define  $Q_i(t)$  as the number of packets in queue  $i$  at the beginning of time slot  $t$ , and let  $Q(t) = (Q_1(t), Q_2(t), Q_3(t))$ .

We define  $S_i(t)$  as the number of packets that the servers attempt to remove from queue  $i$  during time slot  $t$ , and let  $S(t) = (S_1(t), S_2(t), S_3(t))$ . We assume that  $S_i(t)$  can only take values in  $\{0, 1\}$ . An attempt will be successful if and only if queue  $i$  has a packet, so that the number of packets actually removed from queue  $i$  during time slot  $t$  is  $S_i(t) \cdot 1_{\{Q_i(t) > 0\}}$ . Accordingly, the **queue length dynamics** take the form

$$Q_i(t+1) = Q_i(t) + A_i(t) - S_i(t) \cdot 1_{\{Q_i(t) > 0\}}, \quad t \in \mathbb{Z}_+, \quad i \in \{1, 2, 3\},$$

initialized with an arbitrary vector of initial queue lengths  $Q(0)$  in  $\mathbb{Z}_+^3$ .

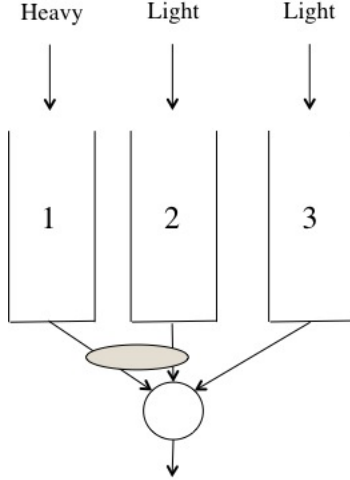


Figure 1: A simple system of three single-class, single-server queues. A scheduler decides, at each time slot, whether queues 1 and 2 will be served simultaneously, or whether queue 3 will be served alone. Queue 1 receives heavy-tailed traffic, whereas queues 2 and 3 receive light-tailed traffic.

## 2.2 Scheduling constraints and the Max-Weight policy

We assume that the servers of queues 1 and 2 can be active simultaneously, whereas the server of queue 3 can only be active alone. “Scheduling constraints” of this type are common in communication networks, such as wireless networks or data switches. Formally, the sets of queues  $\{1, 2\}$  and  $\{3\}$ , as well as the empty set, are called the **feasible schedules**. Furthermore,  $S(t)$  is constrained to take values in the set  $\mathcal{S} = \{(0, 0, 0), (1, 1, 0), (0, 0, 1)\}$ , for all  $t \in \mathbb{Z}_+$ . The service discipline of packets within each queue is assumed to be “**First Come, First Served**.”

The choice of a service vector  $S(t) \in \mathcal{S}$  is made at each time slot according to a scheduling policy. In this paper, we focus exclusively on the Max-Weight policy, which chooses a schedule with the maximum total workload, at each time slot. Formally,

$$S(t) \in \arg \max_{S \in \mathcal{S}} \left\{ \sum_{i=1}^3 Q_i(t) \cdot S_i \right\}, \quad t \in \mathbb{Z}_+.$$

In case of a tie, each one of the feasible schedules is chosen with equal probability.

### 2.3 Assumptions on the arrival processes

We assume that the discrete time stochastic arrival processes  $\{A_i(t); t \in \mathbb{Z}_+\}$ ,  $i \in \{1, 2, 3\}$  are independent and identically distributed (IID) over time, and mutually independent. We denote by  $\lambda_i = E[A_i(0)]$  the arrival rate of the  $i^{th}$  process, and let  $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ . For the model to be interesting, all rates are assumed to be strictly positive.

We say that a random variable  $X$  is **heavy-tailed** (respectively, **light-tailed**) if  $E[X^2]$  is infinite (respectively, finite). Similarly, we say that the  $i^{th}$  arrival process is heavy or light-tailed depending on whether  $E[A_i^2(t)]$  is infinite or finite, respectively. Throughout the paper we assume that the first arrival process is heavy-tailed, and the other two arrival processes are light-tailed.

### 2.4 The stability region

A batch of packets arriving to a queue at any given time slot can be viewed as a single entity, and will be referred to as a **file**. We define the **end-to-end delay of a file** to be the number of time slots that the file spends in the system, starting from the time slot right after it arrives, until the time slot that its last packet exits the system. (For example, a two-packet file that arrives during time slot  $t$  and is served during the next two time slots has an end-to-end delay of 2.) For  $k \in \mathbb{N}$ , we denote by  $D_i(k)$  the end-to-end delay of the  $k^{th}$  file arriving at queue  $i$ , and let  $D(k) = (D_1(k), D_2(k), D_3(k))$ .

The following definition gives the precise notion of stability that we use in this paper.

**Definition 2: (Stability)** The earlier defined queueing system is stable under a particular scheduling policy if the vector-valued sequences  $\{Q(t); t \in \mathbb{Z}_+\}$  and  $\{D(k); k \in \mathbb{N}\}$  converge in distribution, and their limiting distributions do not depend on  $Q(0)$ .

Notice that our definition of stability is slightly different than the commonly used one (positive recurrence of the Markov chain of queue lengths), because it includes the convergence of the sequence of file delays  $\{D(k); k \in \mathbb{N}\}$ . The reason is that we are interested in properties of the limiting distribution of  $\{D(k); k \in \mathbb{N}\}$  and, naturally, we need to ensure that this limiting distribution exists.

An arrival rate vector  $\lambda$  is said to belong to the **stability region** of the system, if there exist nonnegative real numbers  $\mu_{12}$  and  $\mu_3$ , such that

$$\max\{\lambda_1, \lambda_2\} \leq \mu_{12},$$

$$\lambda_3 \leq \mu_3,$$

$$\mu_{12} + \mu_3 < 1.$$

Intuitively, one can think of  $\mu_{12}$  and  $\mu_3$  as the fractions of time that the schedules  $(1, 1, 0)$  and  $(0, 0, 1)$ , respectively, are applied. If a positive arrival vector  $\lambda$  is outside (and at a positive distance) from the stability region, it is easily shown that there is no scheduling policy that makes the system stable. In contrast, when  $\lambda$  belongs to the stability region, the system can be made stable, e.g., by using the Max-Weight policy; see Lemma 1 below. Because of these considerations, we will assume throughout the rest of the paper that  $\lambda$  belongs to the stability region.

**Lemma 1: (Stability under Max-Weight)** The queueing system under consideration is stable under the Max-Weight scheduling policy, for all positive arrival rate vectors in the stability region.

*Proof.* It can be verified that the sequence  $\{Q(t); t \in \mathbb{Z}_+\}$  is a time-homogeneous, irreducible, and aperiodic Markov chain on the countable state-space  $\mathbb{Z}_+^3$ . Proposition 2 of [20] implies that this Markov chain is also positive recurrent. Hence,  $\{Q(t); t \in \mathbb{Z}_+\}$  converges in distribution, and its limiting distribution does not depend on  $Q(0)$ . Based on this, it can be verified that the sequence  $\{D(k); k \in \mathbb{N}\}$  is a (possibly delayed) aperiodic and positive recurrent regenerative process. Therefore, it also converges in distribution, and its limiting distribution does not depend on  $Q(0)$ ; see [19].  $\square$

We let  $Q = (Q_1, Q_2, Q_3)$  and  $D = (D_1, D_2, D_3)$  be random vectors distributed according to the limiting distributions of  $\{Q(t); t \in \mathbb{Z}_+\}$  and  $\{D(k); k \in \mathbb{N}\}$ , respectively, under the Max-Weight scheduling policy, and refer to them as the steady-state queue lengths and delays.

## 2.5 Delay stability

We now define formally, and then discuss, the property that we will be focusing on.

**Definition 3: (Delay Stability)** Queue  $i \in \{1, 2, 3\}$  is delay stable under a particular scheduling policy if the system is stable under that policy, and  $\mathbb{E}[D_i]$  is finite; otherwise, queue  $i$  is delay unstable.

The following lemma relates the steady-state quantities  $\mathbb{E}[Q_i]$  and  $\mathbb{E}[D_i]$ , and will help us prove delay stability results.

**Lemma 2 [15]:** For the system under consideration, and under Max-Weight scheduling,

$$\mathbb{E}[Q_i] < \infty \iff \mathbb{E}[D_i] < \infty, \quad i \in \{1, 2, 3\}.$$

In the remainder of the paper we will focus on the delay stability of queue 2. This is because under the Max-Weight scheduling policy, queues 1 and 3 are always delay unstable, as established formally in Theorems 1 and 2 of the companion paper [15].

**Proposition 1 [15]:** For the system under consideration, and under Max-Weight scheduling, queues 1 and 3 are delay unstable, for all positive arrival rate vectors in the stability region.

We provide some intuition as to why Proposition 1 is true. For queue 1, the result follows easily from the Pollaczek-Khinchine formula for the expected delay in a stable M/GI/1 queue, and a stochastic comparison argument. Regarding queue 3, we argue as follows. Queue 1 is occasionally very long (infinite, in steady-state expectation), due to the heavy-tailed nature of the traffic that it receives. Whenever queue 1 becomes very long, queue 3 is starved and builds up, at a more or less constant rate, to a size comparable to that of queue 1. Thus, large values of  $Q_1(t)$  result in large values of  $Q_3(t)$  at subsequent times. As  $\mathbb{E}[Q_1]$  is infinite, we can argue that  $\mathbb{E}[Q_3]$  is infinite as well, and Lemma 2 implies that queue 3 is delay unstable.

Our main results will concern queue 2. We will show in the next two sections that it is delay unstable (respectively, delay stable) if  $\lambda_2$  is greater (respectively, less) than  $(1 + \lambda_1 - \lambda_3)/2$ .



### 3 Delay Instability Region

Lemma 1 and Proposition 1 in the previous section establish that the queueing system under consideration is stable, and that queues 1 and 3 are delay unstable for every positive arrival rate vector in the stability region. Regarding queue 2, the situation is more complicated. One could argue that queue 1, by virtue of its receiving heavy-tailed traffic, would tend to claim service more often, and that queue 2 would benefit from the resulting additional service opportunities. Somewhat surprisingly though, there is a more intricate indirect mechanism at play that can render queue 2 delay unstable.

**Proposition 2:** For the system under consideration, with positive arrival rates within the stability region, and under Max-Weight scheduling, if the arrival rates satisfy  $\lambda_2 > (1 + \lambda_1 - \lambda_3)/2$ , then queue 2 is delay unstable.

Before proceeding with the formal proof of Proposition 2, we provide an intuitive outline of the argument, also aimed at explaining the threshold value  $(1 + \lambda_1 - \lambda_3)/2$ . Our approach is based on tracking the evolution of the system on a particular set of “fluid” sample paths:<sup>1</sup> assume that at time slot 0, queue 1 receives a very large file, consisting of  $b$  packets. For a long period of time after that, queue 3 does not receive service under the Max-Weight policy, and builds up. If the arrival processes of all traffic flows are close to their “average behavior” (in the Strong Law of Large Numbers sense), then at the time slot when the service switches from schedule  $\{1, 2\}$  to schedule  $\{3\}$ , the lengths of both queues 1 and 3 are proportional to  $b$ , whereas queue 2 is still small. From that point on, the Max-Weight policy will drain the weights of the two schedules at roughly the same rate, until one of the weights becomes zero.

Let  $\mu_i$  be the average departure rate from queue  $i$  during the latter period. For the weights of the two schedules to be drained at the same rate, the departure rates have to satisfy:

$$\lambda_1 + \lambda_2 - \mu_1 - \mu_2 = \lambda_3 - \mu_3.$$

---

<sup>1</sup>In essence, our argument employs a “fluid approximation”; however, instead of establishing formal convergence to a fluid model (which would be unduly complicated for our purposes), we proceed in an elementary manner, arguing from first principles.

Moreover, the fact that Max-Weight is a work-conserving policy implies that

$$\mu_1 + \mu_3 = 1.$$

Finally, since queues 1 and 2 are served simultaneously, and queue 2 may be empty through parts of the draining period, we have that

$$\mu_1 \geq \mu_2.$$

The above equations and some simple algebra imply that

$$\mu_2 \leq \frac{1 + \lambda_1 + \lambda_2 - \lambda_3}{3}.$$

Suppose that the arrival rates satisfy

$$\lambda_2 > \frac{1 + \lambda_1 - \lambda_3}{2}.$$

Then,

$$\lambda_2 > \frac{1 + \lambda_1 + \lambda_2 - \lambda_3}{3} \geq \mu_2.$$

This implies that queue 2 builds up at a roughly constant rate, to size  $O(b)$ , during a period of time whose duration is proportional to  $b$ , and the integral of  $Q_2$  over a busy period of the process becomes of order  $O(b^2)$ . Because  $b$  is drawn from a heavy-tailed distribution (infinite second moment), it follows that  $\mathbb{E}[Q_2]$  is infinite.

*Proof.* We start by defining a shorthand notation that we will be using in the course of the proof. We say that a random variable  $X$  *scales at least linearly* with  $b$  on the event  $H$ , and write  $X = \Omega_H(b)$ , if there exist positive constants  $k$  and  $k'$  (possibly depending on the event  $H$ ), such that  $X \geq k \cdot b - k'$ , for all sample paths in  $H$ .

We break the proof into four steps, which follow the various stages in our earlier proof outline.

**Step 1: buildup of queue 3 following a large arrival to queue 1.**

Because the system is stable, it empties infinitely often, and the times at which  $Q(t) = 0$  are renewal

epochs. Let us consider the system at a typical renewal epoch which, for simplicity of notation, we assume to happen at time 0. Consider the set of sample paths for which, at time slot 0, queue 1 receives a file that consists of  $b$  packets, and all other queues receive no traffic; we denote this set of sample paths by  $H(b)$ . Since  $\lambda_i < 1$  (stability), and since  $1 - \mathbb{P}(A_i(0) = 0) = \mathbb{P}(A_i(0) > 0) \leq \mathbb{E}[A_i(0)] = \lambda_i < 1$ , we have  $\mathbb{P}(A_i(0) = 0) > 0$ , for all  $i$ . Let  $\mathcal{B}$  be the support of the distribution of  $A_1(0)$ . Using the independence of the arrival processes, we have, for every  $b \in \mathcal{B}$ ,

$$\mathbb{P}(H(b)) = \mathbb{P}(A_1(0) = b) \cdot \mathbb{P}(A_2(0) = 0) \cdot \mathbb{P}(A_3(0) = 0) > 0.$$

For sample paths in  $H(b)$ , we denote by  $T_b^1$  the first time slot, starting from 0, when the length of queue 3 becomes greater than or equal to the sum of the lengths of queues 1 and 2:

$$T_b^1 = \min\{t > 0 \mid Q_3(t) \geq Q_1(t) + Q_2(t)\} \cdot 1_{H(b)}.$$

The first part of the proof is to show that  $Q_1(T_b^1)$  and  $Q_3(T_b^1)$  scale at least linearly with  $b$ , provided all arrival processes are close to their “average behavior.”

By the definition of the stopping time  $T_b^1$ , we have

$$Q_1(T_b^1) \leq Q_1(T_b^1) + Q_2(T_b^1) \leq Q_3(T_b^1).$$

A direct consequence of the Strong Law of Large Numbers is that for every  $\epsilon > 0$  there exists  $\delta > 0$ , such that the set of sample paths

$$\Delta = \left\{ (\lambda_i - \epsilon)t - \delta \leq \sum_{\tau=1}^t A_i(\tau) \leq (\lambda_i + \epsilon)t + \delta, \quad \forall t \in \mathbb{N}, \quad \forall i \in \{1, 2, 3\} \right\},$$

has positive probability. Consequently, the set of sample paths

$$\Delta(b) = \left\{ (\lambda_i - \epsilon)t - \delta \leq \sum_{\tau=1}^t A_i(\tau) \leq (\lambda_i + \epsilon)t + \delta, \quad \forall t \in \{1, \dots, T_b^1 - 1\}, \quad \forall i \in \{1, 2, 3\} \right\},$$

has positive probability, uniformly over all  $b$ .

From now on we fix a small  $\epsilon > 0$ . (How small it will have to be will become apparent in the course of the proof.) We then fix a corresponding  $\delta$  such that  $\inf_b \mathbb{P}(\Delta(b)) \geq \mathbb{P}(\Delta) > 0$ . Let  $\tilde{H}(b)$  be the set of sample paths  $H(b) \cap \Delta(b)$ , and observe that  $H(b) \cap \Delta(b) \supset H(b) \cap \Delta$ . Then, the IID nature of the arriving traffic implies that  $H(b)$  and  $\Delta$  are independent, so that

$$\mathbb{P}(\tilde{H}(b)) = \mathbb{P}(H(b) \cap \Delta(b)) \geq \mathbb{P}(H(b) \cap \Delta) = \mathbb{P}(H(b)) \cdot \mathbb{P}(\Delta) > 0.$$

Recall that at most one packet can be removed from each queue at any given time slot. So, for sample paths in  $\tilde{H}(b)$ , we have that

$$Q_1(T_b^1) \geq b - (T_b^1 - 1) + (\lambda_1 - \epsilon) \cdot (T_b^1 - 1) - \delta.$$

Moreover, queue 3 receives no service before time slot  $T_b^1$  under the Max-Weight scheduling policy, which implies that

$$Q_3(T_b^1) = \sum_{t=1}^{T_b^1-1} A_3(t) \leq (\lambda_3 + \epsilon) \cdot (T_b^1 - 1) + \delta.$$

Since  $Q_1(T_b^1) \leq Q_3(T_b^1)$ , the last two inequalities and some algebra yield

$$(1) \quad T_b^1 - 1 \geq \frac{b - 2\delta}{1 + \lambda_3 - \lambda_1 + 2\epsilon}.$$

(This argument requires the last denominator to be positive. This will be the case as long as  $\epsilon$  has been chosen small enough.) Therefore,

$$(2) \quad Q_3(T_b^1) = \sum_{t=1}^{T_b^1-1} A_3(t) \geq (\lambda_3 - \epsilon) \cdot (T_b^1 - 1) - \delta \geq (\lambda_3 - \epsilon) \cdot \frac{b - 2\delta}{1 + \lambda_3 - \lambda_1 + 2\epsilon} - \delta,$$

which implies that  $Q_3(T_b^1) = \Omega_{\tilde{H}(b)}(b)$ , since we can chose  $\epsilon$  to be less than  $\lambda_3$ .

Coming to queue 2, it can be verified that for sample paths in  $\Delta$  (and hence in  $\tilde{H}(b)$ ), and for

any subinterval  $\{\tau_0, \dots, \tau_1\}$  of  $\{1, \dots, T_b^1\}$ ,

$$\sum_{t=\tau_0}^{\tau_1-1} A_2(t) \leq (\lambda_2 + \epsilon) \cdot (\tau_1 - \tau_0) + 2\epsilon\tau_0 + 2\delta.$$

We assume that  $\epsilon$  has been chosen so that  $\lambda_2 + \epsilon < 1$ . Recall also that queue 2 gets served whenever it is nonempty throughout the period  $\{1, \dots, T_b^1 - 1\}$ . We use Lindley's formula, the above upper bound on the arrivals to queue 2, and the fact that since queue 2 gets served whenever it is nonempty throughout the period  $\{1, \dots, T_b^1 - 1\}$ , to conclude that

$$(3) \quad Q_2(T_b^1) \leq A_2(T_b^1 - 1) + 2\epsilon(T_b^1 - 1) + 2\delta \leq \lambda_2 + 4\epsilon(T_b^1 - 1) + 4\delta.$$

This shows that, essentially,  $Q_2(T_b^1)$  does not scale with  $b$ .

We finally turn our attention to  $Q_1(T_b^1)$ . By definition of the stopping time  $T_b^1$ ,

$$(4) \quad Q_3(T_b^1 - 1) < Q_1(T_b^1 - 1) + Q_2(T_b^1 - 1).$$

By arguing similar to the derivation of Eqs. (2) and (3), it can be verified that

$$(5) \quad Q_2(T_b^1 - 1) \leq \lambda_2 + 4\epsilon(T_b^1 - 2) + 4\delta,$$

and that  $Q_3(T_b^1 - 1) = \Omega_{\tilde{H}(b)}(b)$ . Then, Eqs. (4) and (5) readily imply that  $Q_1(T_b^1) = \Omega_{\tilde{H}(b)}(b)$ , when  $\epsilon$  is chosen sufficiently small.

To summarize, at time slot  $T_b^1$ , and for sample paths in  $\tilde{H}(b)$ , the lengths of queues 1 and 3 are proportional to  $b$ , while queue 2 is still small.

**Step 2: draining down until queue 1 or 3 empties.**

Let  $T_b^2$  be the first time slot after  $T_b^1$  that either queue 1 or queue 3 becomes empty:

$$T_b^2 = \min\{t > T_b^1 \mid Q_1(t) \cdot Q_3(t) = 0\} \cdot 1_{\tilde{H}(b)}.$$

We will show that if the arrival processes stay close to their ‘‘average behavior’’ (i.e., the event

$\tilde{H}(b)$  occurs), then, at time slot  $T_b^2$ , the length of queue 3 cannot be much larger than the sum of the lengths of queues 1 and 2. The reason is that  $Q_1(t) + Q_2(t)$  and  $Q_3(t)$  are kept roughly equal by the Max-Weight policy throughout the interval  $[T_b^1, T_b^2]$ .

For the same constants  $\epsilon$  and  $\delta$  as Step 1, the set of sample paths

$$\Delta'(b) = \left\{ (\lambda_i - \epsilon)t - \delta \leq \sum_{\tau=T_b^1}^t A_i(\tau) \leq (\lambda_i + \epsilon)t + \delta, \forall t \in \{T_b^1, \dots, T_b^2 - 1\}, \forall i \in \{1, 2, 3\} \right\}$$

contains  $\Delta$ . Let  $\hat{H}(b) = \tilde{H}(b) \cap \Delta'(b)$ . The events  $\Delta(b)$  and  $\Delta'(b)$  are determined by the arrivals over disjoint time intervals. Hence, due to the IID nature of the arrival processes,  $\Delta(b)$  and  $\Delta'(b)$  are independent. Since they both contain the positive probability event  $\Delta$ , we have

$$\mathbb{P}(\hat{H}(b)) \geq \mathbb{P}(H(b)) \cdot (\mathbb{P}(\Delta))^2 > 0.$$

We will show that for sample paths in  $\hat{H}(b)$ ,

$$(6) \quad Q_3(T_b^2) \leq Q_1(T_b^2) + Q_2(T_b^2) + 2\epsilon(T_b^2 - T_b^1) + 2\delta + 3.$$

We first notice that queues 1 and 3 cannot empty at the same time slot, since they cannot be served simultaneously. Therefore, we have two possible cases: either  $Q_3(T_b^2) = 0$ , in which case Eq. (6) is trivially satisfied, or  $Q_1(T_b^2) = 0$ , which we henceforth assume. In the latter case,  $S_1(T_b^2 - 1) = S_2(T_b^2 - 1) = 1$ , and  $S_3(T_b^2 - 1) = 0$ . For sample paths in  $\hat{H}(b)$ , we have that

$$(7) \quad Q_3(T_b^2) = Q_3(T_b^2 - 1) + A_3(T_b^2 - 1) \leq Q_3(T_b^2 - 1) + \lambda_3 + 2\epsilon \cdot (T_b^2 - T_b^1) + 2\delta.$$

Moreover, under the Max-Weight scheduling policy, and in order for the for the set of queues  $\{1, 2\}$  to be served at time slot  $T_b^2 - 1$ ,

$$(8) \quad Q_3(T_b^2 - 1) \leq Q_1(T_b^2 - 1) + Q_2(T_b^2 - 1).$$

Finally,

$$(9) \quad Q_1(T_b^2 - 1) + Q_2(T_b^2 - 1) - 2 \leq Q_1(T_b^2) + Q_2(T_b^2).$$

Eq. (6) follows immediately, by combining Eqs. (7)-(9) and using also the fact  $\lambda_3 < 1$ .

**Step 3: growth of queue 2.**

At time  $T_b^2$ ,  $Q_3(T_b^2)$  cannot much larger than  $Q_1(T_b^2) + Q_2(T_b^2)$ . Therefore, queue 3 must have been receiving a certain fraction of the total service between times  $T_b^1$  and  $T_b^2$ . We will show that this results in queue 2 not receiving enough service, and that  $Q_2$  starts growing. In particular, if  $\lambda_2 > (1 + \lambda_1 - \lambda_3)/2$ , then, for the sample paths of interest,  $Q_2(T_b^2) = \Omega_{\hat{H}(b)}(b)$ .

By definition,

$$Q_3(T_b^1) \geq Q_1(T_b^1) + Q_2(T_b^1).$$

By subtracting the two sides of this inequality from Eq. (6), we get

$$(10) \quad Q_3(T_b^2) - Q_3(T_b^1) \leq Q_1(T_b^2) - Q_1(T_b^1) + Q_2(T_b^2) - Q_2(T_b^1) + 2\epsilon(T_b^2 - T_b^1) + 2\delta + 3.$$

For sample paths in  $\hat{H}(b)$ , define the random variables

$$\mu_i = \left( \frac{1}{T_b^2 - T_b^1} \cdot \sum_{t=T_b^1}^{T_b^2-1} S_i(t) \right) \cdot 1_{\hat{H}(b)}, \quad i \in \{1, 2, 3\},$$

which are the average service rates to each queue during the interval  $\{T_b^1, \dots, T_b^2 - 1\}$ . Notice that  $\mu_1 = \mu_2$  and  $\mu_1 + \mu_3 = 1$ .

Since both queues 1 and 3 are nonempty during the interval  $\{T_b^1, \dots, T_b^2 - 1\}$ , we have

$$(11) \quad Q_1(T_b^2) - Q_1(T_b^1) \leq (\lambda_1 + \epsilon - \mu_1) \cdot (T_b^2 - T_b^1) + \delta,$$

$$(12) \quad Q_3(T_b^2) - Q_3(T_b^1) \geq (\lambda_3 - \epsilon - \mu_3) \cdot (T_b^2 - T_b^1) - \delta.$$

Eqs. (10), (11), and (12) imply that

$$\begin{aligned} (\lambda_3 - \epsilon - \mu_3) \cdot (T_b^2 - T_b^1) - \delta &\leq (\lambda_1 + \epsilon - \mu_1) \cdot (T_b^2 - T_b^1) + \delta \\ &+ Q_2(T_b^2) - Q_2(T_b^1) + 2\epsilon(T_b^2 - T_b^1) + 2\delta + 3. \end{aligned}$$

We replace  $\mu_3$  by  $1 - \mu_1$  and collect terms, to obtain

$$\begin{aligned} -\mu_1 \cdot (T_b^2 - T_b^1) &\geq -\left(\frac{1 + \lambda_1 - \lambda_3 + 4\epsilon}{2}\right) \cdot (T_b^2 - T_b^1) + \frac{Q_2(T_b^1) - Q_2(T_b^2)}{2} - \frac{4\delta + 3}{2} \\ &\geq -\left(\frac{1 + \lambda_1 - \lambda_3 + 4\epsilon}{2}\right) \cdot (T_b^2 - T_b^1) - \frac{Q_2(T_b^2)}{2} - \frac{4\delta + 3}{2}. \end{aligned}$$

For sample paths in  $\hat{H}(b)$ , we use the definition of  $\Delta'(b)$  to upper bound the number of arrivals to queue 2. We also use the fact that queue 2 has  $\mu_2(T_b^2 - T_b^1)$  service opportunities, with  $\mu_2 = \mu_1$ , and obtain

$$\begin{aligned} Q_2(T_b^2) &\geq (\lambda_2 - \epsilon - \mu_1) \cdot (T_b^2 - T_b^1) - \delta \\ &\geq \left(\lambda_2 - \frac{1 + \lambda_1 - \lambda_3}{2} - 3\epsilon\right) \cdot (T_b^2 - T_b^1) - \frac{Q_2(T_b^2)}{2} - \frac{6\delta + 3}{2}. \end{aligned}$$

Therefore,

$$Q_2(T_b^2) \geq \frac{2}{3} \cdot \left(\lambda_2 - \frac{1 + \lambda_1 - \lambda_3}{2} - 3\epsilon\right) \cdot (T_b^2 - T_b^1) - 2\delta - 1.$$

If  $\lambda_2 > (1 + \lambda_1 - \lambda_3)/2$ , the constant  $\epsilon$  can be chosen sufficiently small so that

$$\lambda_2 - \frac{1 + \lambda_1 - \lambda_3}{2} - 3\epsilon > 0.$$

A final observation is that the duration of the interval  $\{T_b^1, \dots, T_b^2 - 1\}$  is bounded from below by  $\min\{Q_1(T_b^1), Q_3(T_b^1)\}$ , because both queues are served at unit rate. Therefore,  $T_b^2 - T_b^1 = \Omega_{\hat{H}(b)}(b)$  and

$$(13) \quad Q_2(T_b^2) = \Omega_{\hat{H}(b)}(b).$$



**Step 4: the growth scenario for queue 2 implies large average queue size.**

In Step 3 we showed that for sample paths in  $\hat{H}(b)$ , queue 2 builds up to the order of  $b$ . We use this fact, and renewal theory, to show that the steady-state expected length of queue 2 is infinite. The sequence of times at which  $Q(t) = 0$  are renewal epochs. Denote by  $X_i$  the length of the  $i^{\text{th}}$  inter-renewal period. The random variables  $\{X_i; i \in \mathbb{N}\}$  can be viewed as IID copies of some nonnegative random variable  $X$ , with finite first moment; this is because the empty state is positive recurrent under the Max-Weight policy (see Proposition 2 of [20]).

We define an instantaneous reward on this renewal process:

$$R^M(t) = \min\{Q_2(t), M\}, \quad t \in \mathbb{Z}_+,$$

where  $M$  is a positive integer.

Eq. (13) implies that there exist positive constants  $c$  and  $b_0$ , such that

$$Q_2(T_b^2) \geq cb, \quad \forall b \geq b_0,$$

for all sample paths in  $\hat{H}(b)$ .

Since at most one packet from queue 2 can be served at each time slot, the length of queue 2 is at least  $cb/2$  packets over a time period of length at least  $cb/2$  time slots. Hence, the aggregate reward  $R_{agg}^M$ , i.e., the reward accumulated over a renewal period, satisfies the lower bound

$$R_{agg}^M \cdot 1_{\{b \geq b_0\}} \cdot 1_{\hat{H}(b)} \geq \min \left\{ \left( \frac{cb}{2} \right)^2 \cdot 1_{\{b \geq b_0\}}, M^2 \right\} \cdot 1_{\hat{H}(b)}.$$

Then, the expected aggregate reward is bounded below by

$$\mathbb{E}[R_{agg}^M] \geq (\mathbb{P}(\Delta))^2 \cdot \mathbb{P}(A_2(0) = 0) \cdot \mathbb{P}(A_3(0) = 0) \cdot \sum_{b=1}^{\infty} \min \left\{ \left( \frac{cb}{2} \right)^2 \cdot 1_{\{b \geq b_0\}}, M^2 \right\} \cdot \mathbb{P}(A_1(0) = b).$$

So, there exists a positive constant  $c'$  such that

$$(14) \quad c' \cdot \mathbb{E} \left[ \min \left\{ \left( \frac{cA_1(0)}{2} \right)^2 \cdot 1_{\{A_1(0) \geq b_0\}}, M^2 \right\} \right] \leq \mathbb{E}[R_{agg}^M].$$

We have argued that inter-renewal periods have finite expectation, and, clearly, the expected aggregate reward is finite. Then, the Renewal Reward theorem (e.g., see Section 3.4 of [8]) implies that

$$(15) \quad \frac{\mathbb{E}[R_{agg}^M]}{\mathbb{E}[X]} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} R^M(t), \quad \text{w.p.1.}$$

Also, the fact that the reward is a bounded function of an ergodic Markov chain implies that

$$(16) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} R^M(t) = \mathbb{E}[\min\{Q_2, M\}], \quad \text{w.p.1.}$$

Eqs. (14)-(16) imply that

$$\frac{c'}{\mathbb{E}[X]} \cdot \mathbb{E} \left[ \min \left\{ \left( \frac{cA_1(0)}{2} \right)^2 \cdot 1_{\{A_1(0) \geq b_0\}}, M^2 \right\} \right] \leq \mathbb{E}[\min\{Q_2, M\}].$$

Taking the limit as  $M$  goes to infinity on both sides, and using the Monotone Convergence theorem (e.g., see Section 5.3 of [24]) gives

$$\frac{c' \cdot c^2}{4\mathbb{E}[X]} \cdot \mathbb{E}[A_1^2(0) \cdot 1_{\{A_1(0) \geq b_0\}}] \leq \mathbb{E}[Q_2].$$

Finally, the fact that the random variable  $A_1(0)$  is heavy-tailed, implies that  $\mathbb{E}[Q_2]$  is infinite. Combined with Lemma 2, this gives the desired result.  $\square$

## 4 Delay Stability Region

In this section we establish that when  $\lambda_2 < (1 + \lambda_1 - \lambda_3)/2$ , then queue 2 is delay stable. Notice that this result involves the same threshold as in Proposition 2, and we therefore have an exact characterization of the delay stability region.

The proof of this result relies on drift analysis of a suitable Lyapunov function. The usual analysis of the Max-Weight policy involves a quadratic Lyapunov function [21]. Different types of

Lyapunov functions have also been used for stability and performance analysis of other queueing networks and policies, e.g., piecewise linear functions [2, 7], and norms [18, 23]. However, in the presence of heavy-tailed traffic, the expected value of all these Lyapunov functions in steady-state is infinite, which renders their drift analysis uninformative.

Here we consider a carefully constructed piecewise linear Lyapunov function that is **not** radially unbounded (i.e., the value of the Lyapunov function need not be large when the queue lengths are large), and which has a negative drift only when  $\lambda_2 < (1 + \lambda_1 - \lambda_3)/2$ . A common technical difficulty with piecewise linear Lyapunov functions is that the stochastic descent property is often lost at locations where the Lyapunov function is nondifferentiable. This difficulty can be handled by either smoothing the Lyapunov function (e.g., as in [7]), or by showing that the stochastic descent property still holds if we look ahead a sufficiently large number of time slots (e.g., as in [22]). We follow the second approach.

Furthermore, we also assume that light-tailed traffic has exponentially decaying tails, which allows us to exploit the tools in [10]. Formally, a nonnegative random variable  $X$  is **exponential-type**, if there exists some  $\theta > 0$  such that  $\mathbb{E}[\exp(\theta X)]$  is finite.<sup>2</sup>

**Proposition 3:** For the system under consideration, suppose that the arrival rate vector is positive and within the stability region, and that the Max-Weight policy is used. Suppose furthermore that (i)  $A_2(0)$  and  $A_3(0)$  are exponential-type, and (ii) there exists some  $\gamma > 0$  such that  $\mathbb{E}[A_1^{1+\gamma}(0)] < \infty$ . If the arrival rates satisfy  $\lambda_2 < (1 + \lambda_1 - \lambda_3)/2$ , then queue 2 is delay stable, and the steady-state length of queue 2 is exponential-type.

Before we proceed with the formal proof of Proposition 3, we provide the underlying intuition by arguing (loosely) in terms of a fluid approximation. Our analysis rests on drift analysis of the following piecewise linear Lyapunov function

$$(17) \quad V(t) = 3Q_2(t) + \left[ Q_3(t) - Q_1(t) - Q_2(t) \right]^+, \quad t \in \mathbb{Z}_+.$$

Our goal is to establish a negative drift for  $V(t)$ , which, combined with the special structure of the

---

<sup>2</sup>We conjecture that the statement of Proposition 3 is true for any type of light-tailed traffic. However, the results in [10] cannot be invoked without the additional assumption of exponential-type traffic.

problem (the arrivals to queues 2 and 3 being exponential-type), will imply that

$$\mathbb{E}\left[3Q_2 + [Q_3 - Q_1 - Q_2]^+\right] < \infty,$$

where the expectation is taken with respect to steady-state distributions. Together with Lemma 2, this will imply the delay stability of queue 2.

The suitability of this Lyapunov function can be seen as follows. If  $Q_1(t) + Q_2(t) > Q_3(t)$ , then the Lyapunov function reduces to  $3Q_2(t)$ , which can be easily shown to have negative drift (as long as  $Q_2 > 0$ ), because queue 2 is served under Max-Weight in that region. If, on the other hand,  $Q_1(t) + Q_2(t) < Q_3(t)$ , then the Lyapunov function reduces to  $2Q_2(t) + Q_3(t) - Q_1(t)$ . In that region, queue 3 is served under Max-Weight, so the drift of  $V(t)$  is equal to  $2\lambda_2 + \lambda_3 - \lambda_1 - 1$ , which is strictly negative by our assumption on  $\lambda_2$ . See Figure 2 for a geometric interpretation.

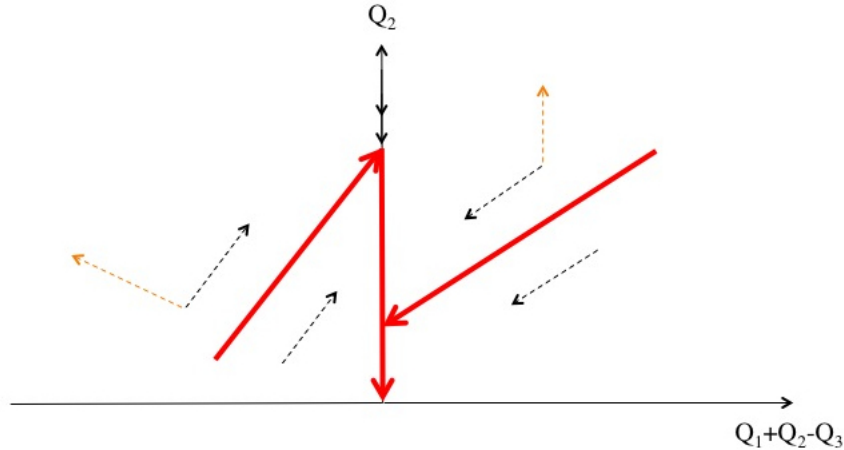


Figure 2: A two-dimensional representation of the dynamics of the system when  $\lambda_2 < (1 + \lambda_1 - \lambda_3)/2$ . The dashed black arrows represent the gradient of the queue length vector under the Max-Weight policy, and the dashed orange arrows represent the gradient of the Lyapunov function, in different regions of the state space. The solid red arrows represent typical trajectories of the associated fluid approximation. Our Lyapunov function has been chosen so that an obtuse angle is formed between the two in all regions of the state space, establishing a uniform negative drift over the entire state space.

The analysis becomes subtler at the boundary between the two regions of the state space, namely when the weights,  $Q_1(t) + Q_2(t)$  and  $Q_3(t)$ , of the two candidate schedules are equal. Once

at the boundary, the Max-Weight policy keeps the state (of the fluid model) there. Similar to our discussion of Proposition 2, the resulting departure rate from queue 2 satisfies

$$\lambda_1 + \lambda_2 - \mu_1 - \mu_2 = \lambda_3 - \mu_3.$$

Using also the properties  $\mu_2 = \mu_1 = 1 - \mu_3$ , and our assumption on  $\lambda_2$ , it follows that the drift at the boundary is negative, driving the state to zero.

*Proof.* Consider the piecewise linear Lyapunov function defined by Eq. (17) and let  $\mathcal{F}_t$  be the  $\sigma$ -field that corresponds to the history of the process until just before the arrivals at slot  $t$ ; formally,  $\mathcal{F}_t$  is the  $\sigma$ -algebra generated by  $Q(0), A(0), \dots, Q(t-1), A(t-1), Q(t)$ . Throughout the proof we use  $\mathbb{P}(X; A \mid \mathcal{H})$  and  $\mathbb{E}[X; A \mid \mathcal{H}]$  to denote  $\mathbb{P}(X \cdot 1_A \mid \mathcal{H})$  and  $\mathbb{E}[X \cdot 1_A \mid \mathcal{H}]$ , respectively, where  $X$  is a random variable,  $A$  is an event, and  $\mathcal{H}$  is a  $\sigma$ -algebra on the sample space.

Our goal is to show that for sufficiently large (but fixed)  $T \in \mathbb{N}$ , there exist positive constants  $\alpha$  and  $\epsilon$ , such that

$$\mathbb{E}\left[V(t+T) - V(t) + \epsilon; V(t) > \alpha \mid \mathcal{F}_t\right] \leq 0,$$

and the desired result will follow from [10].

Suppose that  $V(t) > \alpha$ . Then, we must have either  $Q_2(t) > \alpha/6$  or  $Q_3(t) > Q_1(t) + Q_2(t) + \alpha/2$ . We will derive the desired drift inequality by considering separately these two cases. We assume that  $T$  has been fixed to a suitably large value, and we define  $\alpha$  by  $\alpha = 6T$ .

**Case 1:**  $Q_2(t) > \alpha/6 = T$ .

Since at most one packet is removed at each time slot from queue 2, it is immediate that

$$(18) \quad Q_2(\tau) > 0, \quad \forall \tau \in \{t, \dots, t+T-1\}.$$

Moreover, under the Max-Weight scheduling policy,

$$(19) \quad S_3(\tau) \cdot 1_{\{Q_3(\tau) > 0\}} = S_3(\tau), \quad \forall \tau \in \{t, \dots, t+T-1\}.$$

Eq. (19) implies that a service opportunity is never wasted in queue 3 throughout  $\{t, \dots, t+T-1\}$ , assuming  $Q_2(t) > T$ . This is intuitively clear, because, under the Max-Weight policy, queue 3 will not be served unless it becomes at least as large as queue 2, hence positive, and a service opportunity will not be wasted).

Let

$$J(t) = Q_1(t) + Q_2(t) - Q_3(t).$$

Using this notation, and the queue length dynamics from Section 2, it can be verified that

$$(20) \quad \begin{aligned} & \sum_{\tau=t}^{t+T-1} \left( A_1(\tau) + A_2(\tau) - S_1(\tau) \cdot 1_{\{Q_1(\tau)>0\}} - S_2(\tau) \cdot 1_{\{Q_2(\tau)>0\}} \right) \\ &= \sum_{\tau=t}^{t+T-1} \left( A_3(\tau) - S_3(\tau) \cdot 1_{\{Q_3(\tau)>0\}} \right) + J(t+T) - J(t). \end{aligned}$$

Moreover, the scheduling constraints imply that

$$(21) \quad S_1(\tau) = S_2(\tau), \quad \forall \tau \in \{t, \dots, t+T-1\}.$$

Furthermore, under the Max-Weight policy (or, in general, under any non-idling policy),

$$(22) \quad \sum_{\tau=t}^{t+T-1} (S_2(\tau) + S_3(\tau)) = T.$$

Eqs. (19) and (20) imply that

$$(23) \quad \sum_{\tau=t}^{t+T-1} (A_1(\tau) + A_2(\tau) - S_1(\tau) - S_2(\tau)) \leq \sum_{\tau=t}^{t+T-1} (A_3(\tau) - S_3(\tau)) + J(t+T) - J(t).$$

Combining Eqs. (21) and (23), we have

$$(24) \quad \sum_{\tau=t}^{t+T-1} (A_1(\tau) + A_2(\tau) - 2 \cdot S_2(\tau)) \leq \sum_{\tau=t}^{t+T-1} (A_3(\tau) - S_3(\tau)) + J(t+T) - J(t).$$

Then, by taking into account Eq. (22), we get

$$(25) \quad -3 \cdot \sum_{\tau=t}^{t+T-1} S_2(\tau) \leq - \sum_{\tau=t}^{t+T-1} (1 + A_1(\tau) + A_2(\tau) - A_3(\tau)) + J(t+T) - J(t).$$

Let us now examine the implications of this inequality on the evolution of queue 2 throughout  $\{t, \dots, t+T-1\}$ . We have

$$\begin{aligned} 3(Q_2(t+T) - Q_2(t)) &= 3 \cdot \sum_{\tau=t}^{t+T-1} (A_2(\tau) - S_2(\tau) \cdot \mathbf{1}_{\{Q_2(\tau) > 0\}}) \\ &= 3 \cdot \sum_{\tau=t}^{t+T-1} (A_2(\tau) - S_2(\tau)) \\ &\leq - \sum_{\tau=t}^{t+T-1} (1 + A_1(\tau) - A_3(\tau) - 2A_2(\tau)) + J(t+T) - J(t), \end{aligned}$$

where the second equality follows from Eq. (18), and the inequality follows from Eq. (25). This implies that

$$\begin{aligned} V(t+T) - V(t) &= 3(Q_2(t+T) - Q_2(t)) + [-J(t+T)]^+ - [-J(t)]^+ \\ &\leq - \sum_{\tau=t}^{t+T-1} (1 + A_1(\tau) - A_3(\tau) - 2A_2(\tau)) \\ &\quad + J(t+T) - J(t) + [-J(t+T)]^+ - [-J(t)]^+ \\ &= - \sum_{\tau=t}^{t+T-1} (1 + A_1(\tau) - A_3(\tau) - 2A_2(\tau)) + [J(t+T)]^+ - [J(t)]^+. \end{aligned}$$

Therefore,

$$(26) \quad \begin{aligned} \mathbb{E} \left[ V(t+T) - V(t) ; Q_2(t) > T \mid \mathcal{F}_t \right] &\leq -\delta \cdot T \cdot \mathbb{P}(Q_2(t) > T \mid \mathcal{F}_t) \\ &\quad + \mathbb{E} \left[ [J(t+T)]^+ - [J(t)]^+ ; Q_2(t) > T \mid \mathcal{F}_t \right]. \end{aligned}$$

where  $\delta = 1 + \lambda_1 - \lambda_3 - 2\lambda_2$  is positive by assumption.

We have also assumed that there exists  $\gamma > 0$ , such that  $\mathbb{E}[A_1^{1+\gamma}(0)]$  is finite. Based on this, we will prove that  $\mathbb{E} \left[ [J(t+T)]^+ - [J(t)]^+ \mid \mathcal{F}_t \right]$  scales sublinearly in  $T$ . This, in turn, will imply that

the right-hand side of Eq. (26) is negative, provided  $T$  is sufficiently large.

By disregarding the contribution of queue 3, we can bound from above  $[J(\tau)]^+$ ,  $\tau \in \{t, \dots, t+T\}$ , by the sum of the lengths of queues 1 and 2 during that interval. Moreover, whenever  $[J(\tau)]^+$  is nonzero, both queues 1 and 2 are served at unit rate. Through Lindley's recursion and simple calculations, it can be verified that

$$(27) \quad [J(t+T)]^+ - [J(t)]^+ \leq \max_{1 \leq s \leq T} \left\{ \sum_{\tau=t+T-s}^{t+T-1} (A_1(\tau) - 1) \right\} + \max_{1 \leq s \leq T} \left\{ \sum_{\tau=t+T-s}^{t+T-1} (A_2(\tau) - 1) \right\}.$$

Also, the following inequality holds for all  $i \in \{1, 2\}$ , and for all  $s \in \{1, \dots, T\}$ :

$$(28) \quad \sum_{\tau=t+T-s}^{t+T-1} (A_i(\tau) - 1) = (\lambda_i - 1)s + \sum_{\tau=t+T-s}^{t+T-1} (A_i(\tau) - \lambda_i) \leq \left| \sum_{\tau=t+T-s}^{t+T-1} (A_i(\tau) - \lambda_i) \right|.$$

Eqs. (27) and (28) imply that for any fixed  $\gamma' \in (0, \gamma)$ ,

$$\begin{aligned} & \mathbb{P}\left([J(t+T)]^+ - [J(t)]^+ \geq c ; Q_2(t) > T \mid \mathcal{F}_t\right) \\ & \leq \mathbb{P}\left(\max_{1 \leq s \leq T} \left\{ \sum_{\tau=t+T-s}^{t+T-1} (A_1(\tau) - 1) \right\} + \max_{1 \leq s \leq T} \left\{ \sum_{\tau=t+T-s}^{t+T-1} (A_2(\tau) - 1) \right\} \geq c ; Q_2(t) > T \mid \mathcal{F}_t\right) \\ & \leq \mathbb{P}\left(\max_{1 \leq s \leq T} \left\{ \left| \sum_{\tau=t+T-s}^{t+T-1} (A_1(\tau) - \lambda_1) \right| \right\} \geq \frac{c}{2} ; Q_2(t) > T \mid \mathcal{F}_t\right) \\ & \quad + \mathbb{P}\left(\max_{1 \leq s \leq T} \left\{ \left| \sum_{\tau=t+T-s}^{t+T-1} (A_2(\tau) - \lambda_2) \right| \right\} \geq \frac{c}{2} ; Q_2(t) > T \mid \mathcal{F}_t\right) \\ & = \mathbb{P}\left(\max_{1 \leq s \leq T} \left\{ \left| \sum_{\tau=t+T-s}^{t+T-1} (A_1(\tau) - \lambda_1) \right|^{1+\gamma'} \right\} \geq \left(\frac{c}{2}\right)^{1+\gamma'} ; Q_2(t) > T \mid \mathcal{F}_t\right) \\ (29) \quad & + \mathbb{P}\left(\max_{1 \leq s \leq T} \left\{ \left| \sum_{\tau=t+T-s}^{t+T-1} (A_2(\tau) - \lambda_2) \right|^{1+\gamma'} \right\} \geq \left(\frac{c}{2}\right)^{1+\gamma'} ; Q_2(t) > T \mid \mathcal{F}_t\right) \end{aligned}$$

Notice that the sequence

$$\left\{ \sum_{\tau=t+T-s}^{t+T-1} (A_i(\tau) - \lambda_i) ; s \in \mathbb{N} \right\}, \quad i \in \{1, 2, 3\},$$



is a martingale. Consequently, the sequence

$$\left\{ \left| \sum_{\tau=t+T-s}^{t+T-1} (A_i(\tau) - \lambda_i) \right|^{1+\gamma'} ; s \in \mathbb{N} \right\}$$

is a nonnegative submartingale. Doob's submartingale inequality (see, e.g., Section 14.6 of [24]) and Eq. (29) imply that

$$\begin{aligned} (30) \quad & \mathbb{P}\left([J(t+T)]^+ - [J(t)]^+ \geq c ; Q_2(t) > T \mid \mathcal{F}_t\right) \\ & \leq \left(\frac{2}{c}\right)^{1+\gamma'} \cdot \mathbb{E}\left[\left|\sum_{\tau=t}^{t+T-1} (A_1(\tau) - \lambda_1)\right|^{1+\gamma'}\right] \cdot 1_{Q_2(t) > T} \\ & + \left(\frac{2}{c}\right)^{1+\gamma'} \cdot \mathbb{E}\left[\left|\sum_{\tau=t}^{t+T-1} (A_2(\tau) - \lambda_2)\right|^{1+\gamma'}\right] \cdot 1_{Q_2(t) > T}. \end{aligned}$$

Moreover, the Marcinkiewicz-Zygmund Strong Law of Large Numbers states that

$$(31) \quad \frac{\left|\sum_{\tau=t}^{t+T-1} (A_i(\tau) - \lambda_i)\right|}{T^{1/(1+\gamma)}} \xrightarrow{\mathcal{L}^1} 0, \quad i \in \{1, 2, 3\}$$

(see, e.g., Chapter 6.10 of [9]). Eqs. (30) and (31) imply that if  $T$  is sufficiently large, then there exists  $k > 0$  (independent of  $T$  and  $c$ ), such that

$$\mathbb{P}\left([J(t+T)]^+ - [J(t)]^+ \geq c ; Q_2(t) > T \mid \mathcal{F}_t\right) \leq \frac{k}{c^{1+\gamma'}} \cdot T^{\frac{1+\gamma'}{1+\gamma}} \cdot 1_{Q_2(t) > T},$$

for all  $c \geq 0$ . This gives

$$\mathbb{E}\left[[J(t+T)]^+ - [J(t)]^+ ; Q_2(t) > T \mid \mathcal{F}_t\right] \leq k \cdot T^{\frac{1+\gamma'}{1+\gamma}} \cdot \sum_{c=1}^{\infty} \frac{1}{c^{1+\gamma'}} \cdot 1_{Q_2(t) > T}.$$

Since  $\gamma' > 0$ , the latter sum converges, so there exists  $k' > 0$  (independent of  $T$ ), such that

$$(32) \quad \mathbb{E}\left[[J(t+T)]^+ - [J(t)]^+ ; Q_2(t) > T \mid \mathcal{F}_t\right] \leq k' \cdot T^{\frac{1+\gamma'}{1+\gamma}} \cdot 1_{Q_2(t) > T}.$$

Finally, Eqs. (26) and (32), and the fact that  $\gamma' < \gamma$ , imply that for any fixed  $\epsilon > 0$ , there exists a sufficiently large  $T$ , such that

$$(33) \quad \mathbb{E}\left[V(t+T) - V(t) + \epsilon ; Q_2(t) > T \mid \mathcal{F}_t\right] \leq 0.$$

**Remark:** the upper bound on  $[J(\tau)]^+$  that we analyzed above can be viewed as an upper bound for a discrete-time stable M/GI/1 queue in where customers arrive in a Bernoulli fashion, and their service times are mutually independent and distributed identically to  $A_1(0) + A_2(0)$ . Since service times are heavy-tailed distributed, the expected steady-state workload in this queue is infinite (an immediate corollary of the Pollaczek-Khinchine formula). Combined with Fatou's lemma, this implies that the expected workload at time slot  $t$  goes to infinity as  $t$  increases. Eq. (32) shows that the expected workload goes to infinity at a sublinear rate (essentially, as  $\mathcal{O}(T^{1/(1+\gamma)})$ ), assuming just the existence of the  $(1+\gamma)$  moment of service times. We note that Central-Limit-Theorem-type arguments cannot be used here, since the second moment of arrivals does not exist.

**Case 2:**  $Q_3(t) > Q_1(t) + Q_2(t) + 3T$ .

Since at most one packet is removed at each time slot from queue 3, it is immediate that

$$Q_3(\tau) > 0, \quad \forall \tau \in \{t, \dots, t+T-1\}.$$

Based on this, it can be easily verified that Eq. (25) still holds. This implies that

$$\begin{aligned} 3(Q_2(t+T) - Q_2(t)) &= 3 \cdot \sum_{\tau=t}^{t+T-1} (A_2(\tau) - S_2(\tau) \cdot \mathbf{1}_{\{Q_2(\tau)>0\}}) \\ &= 3 \cdot \sum_{\tau=t}^{t+T-1} (A_2(\tau) - S_2(\tau)) + 3 \cdot \sum_{\tau=t}^{t+T-1} S_2(\tau) \cdot \mathbf{1}_{\{Q_2(\tau)=0\}} \\ &\leq - \sum_{\tau=t}^{t+T-1} (1 + A_1(\tau) - A_3(\tau) - 2A_2(\tau)), \\ &\quad + J(t+T) - J(t) + 3 \cdot \sum_{\tau=t}^{t+T-1} S_2(\tau) \cdot \mathbf{1}_{\{Q_2(\tau)=0\}}. \end{aligned}$$

Letting  $D$  be the event  $Q_3(t) > Q_1(t) + Q_2(t) + 3T$ , this gives

$$(34) \quad \mathbb{E}\left[V(t+T) - V(t) ; D \mid \mathcal{F}_t\right] \leq -\delta \cdot T \cdot 1_D + \mathbb{E}\left[[J(t+T)]^+ - [J(t)]^+ ; D \mid \mathcal{F}_t\right] \\ + 3 \cdot \mathbb{E}\left[\sum_{\tau=t}^{t+T-1} S_2(\tau) \cdot 1_{\{Q_2(\tau)=0\}} ; D \mid \mathcal{F}_t\right].$$

Working similar to the previous case, it can be verified that for any fixed  $\gamma' \in (0, \gamma)$ , there exists  $k' > 0$  (independent of  $T$ ) such that

$$(35) \quad \mathbb{E}\left[[J(t+T)]^+ - [J(t)]^+ ; D \mid \mathcal{F}_t\right] \leq k' \cdot T^{\frac{1+\gamma'}{1+\gamma}} \cdot 1_D.$$

In view of Eqs. (34) and (35) and in order to establish the negative drift property of the Lyapunov function, it is sufficient to consider a large time horizon  $T$ , and to show that  $\mathbb{E}[\sum_{\tau=t}^{t+T-1} S_2(\tau) \cdot 1_{\{Q_2(\tau)=0\}} \mid \mathcal{F}_t]$  scales sublinearly in  $T$ .

In order to have a wasted service opportunity at queue 2, the schedule  $\{1, 2\}$  must first claim the server. This can only happen if  $\sum_{\tau=t}^{t+T-1} (A_1(\tau) + A_2(\tau))$ , the aggregate arrivals to queues 1 and 2 during the interval  $\{t, \dots, t+T-1\}$  exceed the initial difference between the weights of the two schedules (which is at least  $3T$ ), minus the departures from queue 3 during the same period (which are at most  $T$ ). It follows that

$$\left\{ \sum_{\tau=t}^{t+T-1} S_2(\tau) \cdot 1_{\{Q_2(\tau)=0\}} > 0 \right\} \subset \left\{ \sum_{\tau=t}^{t+T-1} (A_1(\tau) + A_2(\tau)) > 3T - T \right\} \\ = \left\{ \sum_{\tau=t}^{t+T-1} (A_1(\tau) + A_2(\tau) - \lambda_1 - \lambda_2) > (2 - \lambda_1 - \lambda_2)T \right\} \\ \subset \left\{ \left| \sum_{\tau=t}^{t+T-1} (A_1(\tau) + A_2(\tau) - \lambda_1 - \lambda_2) \right| > (2 - \lambda_1 - \lambda_2)T \right\}.$$

Note that  $\lambda_1 + \lambda_2 < 2$ , since the arrival rate vector was assumed to be in the stability region of the system. Then, using Markov's inequality and Eq. (34), we obtain that there exists  $\xi > 0$

(independent of  $T$ ) such that

$$\begin{aligned}
& \mathbb{P}\left(\sum_{\tau=t}^{t+T-1} S_2(\tau) \cdot 1_{\{Q_2(\tau)=0\}} > 0 ; D \mid \mathcal{F}_t\right) \\
& \leq \mathbb{P}\left(\left|\sum_{\tau=t}^{t+T-1} (A_1(\tau) + A_2(\tau) - \lambda_1 - \lambda_2)\right| > (2 - \lambda_1 - \lambda_2)T ; D \mid \mathcal{F}_t\right) \\
& \leq \frac{\mathbb{E}\left[\left|\sum_{\tau=t}^{t+T-1} (A_1(\tau) + A_2(\tau) - \lambda_1 - \lambda_2)\right|\right]}{(2 - \lambda_1 - \lambda_2)T} \cdot 1_D \\
& \leq \xi \cdot T^{\frac{1}{1+\gamma}-1} \cdot 1_D.
\end{aligned}$$

provided  $T$  is sufficiently large. Since the number of wasted service opportunities at queue 2 during an interval of length  $T$  is bounded by  $T$ , we conclude that

$$(36) \quad \mathbb{E}\left[\sum_{\tau=t}^{t+T-1} S_2(\tau) \cdot 1_{\{Q_2(\tau)=0\}} ; D \mid \mathcal{F}_t\right] \leq T \cdot \xi \cdot T^{\frac{1}{1+\gamma}-1} \cdot 1_D = \xi \cdot T^{\frac{1}{1+\gamma}} \cdot 1_D.$$

Eqs. (34)-(36) imply that for any given  $\epsilon > 0$ , there exists a sufficiently large  $T$ , such that

$$(37) \quad \mathbb{E}\left[V(t+T) - V(t) + \epsilon ; D \mid \mathcal{F}_t\right] \leq 0.$$

This completes the derivation of the drift inequality for the second case.

To summarize, we have shown, in Eqs. (33) and (37), that for any given  $\epsilon > 0$ , we can choose  $T$  sufficiently large and let  $\alpha = 6T$  to guarantee that

$$\mathbb{E}\left[V(t+T) - V(t) + \epsilon ; V(t) > \alpha \mid \mathcal{F}_t\right] \leq 0, \quad t \in \mathbb{Z}_+.$$

Theorem 2.3 of [10], and the fact that the sequence  $\{V(t); t \in \mathbb{Z}_+\}$  converges in distribution, imply that the steady-state length of queue 2,  $Q_2$ , is exponential-type, and, in particular, has finite expectation.  $\square$

An immediate corollary of Proposition 3 and Lemma 2 is that queue 2 is delay stable, provided its arrival rate is sufficiently small. To establish this, we assumed that the stochastic processes

$\{A_2(t); t \in \mathbb{Z}_+\}$  and  $\{A_3(t); t \in \mathbb{Z}_+\}$  are exponential-type. At the same time, this additional assumption enabled us to prove a result that is much stronger than delay stability: combining Proposition 3 with the distributional Little’s Law (e.g., see [25]), we have that **the steady-state delay in queue 2 is exponential-type**.

## 5 Discussion

We considered a simple queueing system with heavy-tailed traffic, and showed a rate-dependent delay stability phenomenon under the Max-Weight scheduling policy: there is a part of the stability region where a certain queue is delay stable, and a part of the stability region where the same queue is delay unstable. We note that this phenomenon would not arise if all stochastic primitives were light-tailed.

Despite the simplicity of the underlying queueing system, the proofs of our main results (Propositions 2 and 3) are long, somewhat technical, and, in some sense, “tailored” to the specifics of the system. Hence, generalizing these results using similar methods is not straightforward. What is missing is a systematic and practical methodology for determining which queues are delay stable, which are delay unstable, and over which parts of the stability region, for any given switched queueing system.

On a more technical side, the proof of the delay stability result (Proposition 3) requires the additional assumption that all light-tailed arrival processes are exponential-type. Only then is drift analysis of a piecewise linear Lyapunov functions meaningful (i.e., we can invoke the results in [10]). One possible way to overcome this restriction would be through a suitably defined piecewise quadratic Lyapunov function. However, drift analysis seems cumbersome in that case.

## References

- [1] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, R. Vijayakumar, P. Whiting (2004). Scheduling in a queueing system with asynchronously varying service rates. Probability in the

- Engineering and Informational Sciences, 18, 191-217.
- [2] D. Bertsimas, D. Gamarnik, J.N. Tsitsiklis (2001). Performance of multiclass Markovian queueing networks via piecewise linear Lyapunov functions. *The Annals of Applied Probability*, 11, 1384-1428.
  - [3] S. Borst, O. Boxma, R. Nunez-Queija, B. Zwart (2003). The impact of the service discipline on delay asymptotics. *Performance Evaluation*, 54, 175-206.
  - [4] S. Borst, M. Mandjes, M. van Uitert (2003). Generalized processor sharing with light-tailed and heavy-tailed input. *IEEE/ACM Transactions on Networking*, 11, 821-834.
  - [5] O. Boxma, B. Zwart (2007). Tails in scheduling. *Performance Evaluation Review*, 34, 13-20.
  - [6] J. Dai, W. Lin (2005). Maximum pressure policies in stochastic processing networks. *Operations Research*, 53, 197-218.
  - [7] D. Down, S.P. Meyn (1997). Piecewise linear test functions for stability and instability of queueing networks. *Queueing Systems*, 27, 205-226.
  - [8] R. Gallager (1996). *Discrete stochastic processes*. Kluwer Academic.
  - [9] A. Gut (2005). *Probability: a graduate course*. Springer Texts in Statistics.
  - [10] B. Hajek (1982). Hitting-time and occupation-time bounds implied by drift analysis with applications. *Advances in Applied Probability*, 14, 502-525.
  - [11] K. Jagannathan, M.G. Markakis, E. Modiano, J.N. Tsitsiklis (2010). Throughput optimal scheduling in the presence of heavy-tailed traffic. In: *Proc. Allerton 2010*.
  - [12] P.R. Kumar, T. Seidman (1990). Dynamic instabilities and stabilization methods in distributed real-time scheduling of manufacturing systems. *IEEE Transactions on Automatic Control*, 35, 289-298.
  - [13] W. Leland, M. Taqqu, W. Willinger, D. Wilson (1994). On the self-similar nature of ethernet traffic. *IEEE/ACM Transactions on Networking*, 2, 1-15.

- [14] M.G. Markakis, E. Modiano, J.N. Tsitsiklis (2009). Scheduling policies for single-hop networks with heavy-tailed traffic. In: Proc. Allerton 2009.
- [15] M.G. Markakis, E. Modiano, J.N. Tsitsiklis. Max-Weight scheduling in queueing networks with heavy-tailed traffic. Submitted for publication.
- [16] K. Park, W. Willinger (2000). Self-similar network traffic: an overview. In: Self-Similar Network Traffic and Performance Evaluation, K. Park and W. Willinger, editors, Wiley Inc.
- [17] A. Rybko, A. Stolyar (1992). Ergodicity of stochastic processes describing the operation of open queueing networks. Probl. Peredachi Inf., 3, 3-26.
- [18] D. Shah, J.N. Tsitsiklis, Y. Zhong (2010). Qualitative properties of  $\alpha$ -weighted scheduling policies. In: Proc. Sigmetrics 2010.
- [19] K. Sigman, R. Wolff (1993). A review of regenerative processes. SIAM Review, 35, 269-288.
- [20] A. Stolyar (2004). Maxweight scheduling in a generalized switch: state space collapse and workload minimization in heavy traffic. The Annals of Applied Probability, 14, 1-53.
- [21] L. Tassiulas, A. Ephremides (1992). Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. IEEE Transactions on Automatic Control, 37, 1936-1948.
- [22] J.N. Tsitsiklis (1987). Analysis of a multiaccess control scheme. IEEE Transactions on Automatic Control, 32, 1017-1020.
- [23] V.J. Venkataraman, X. Lin, L. Ying, S. Shakkottai (2010). On scheduling for minimizing end-to-end buffer usage over multihop wireless networks. In: Proc. Infocom 2010.
- [24] D. Williams (1991). Probability with Martingales. Cambridge University Press.
- [25] W. Whitt (1991). A review fo  $L = \lambda W$  and extensions. Queueing Systems, 9, 235-268.