informs.
http://pubsonline.informs.org/journal/moor/

Delay Analysis of the Max-Weight Policy Under Heavy-Tailed Traffic via Fluid Approximations

Mihalis G. Markakis,[a] Eytan Modiano,[b] John N. Tsitsiklis[b]

[a] Department of Economics and Business, Universitat Pompeu Fabra and Barcelona School of Management, 08005 Barcelona, Spain;
[b] Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
Contact: mihalis.markakis@upf.edu, http://orcid.org/0000-0003-1469-7729 (MGM); modiano@mit.edu (EM); jnt@mit.edu (JNT)

Abstract. We consider switched queueing networks with a mix of heavy-tailed (i.e., arrival processes with infinite variance) and exponential-type traffic and study the delay performance of the max-weight policy, known for its throughput optimality and asymptotic delay optimality properties. Our focus is on the impact of heavy-tailed traffic on exponential-type queues/flows, which may manifest itself in the form of subtle rate-dependent phenomena. We introduce a novel class of Lyapunov functions (piecewise linear and nonincreasing in the length of heavy-tailed queues), whose drift analysis provides exponentially decaying upper bounds to queue-length tail asymptotics despite the presence of heavy tails. To facilitate a drift analysis, we employ fluid approximations, proving that if a continuous and piecewise linear function is also a "Lyapunov function" for the fluid model, then the same function is a "Lyapunov function" for the original stochastic system. Furthermore, we use fluid approximations and renewal theory in order to prove delay instability results, i.e., infinite expected delays in steady state. We illustrate the benefits of the proposed approach in two ways: (i) analytically, by studying the delay stability regions of single-hop switched queueing networks with disjoint schedules, providing a precise characterization of these regions for certain queues and inner and outer bounds for the rest. As a side result, we prove monotonicity properties for the service rates of different schedules that, in turn, allow us to identify "critical configurations" toward which the state of the system is driven, and that determine to a large extent delay stability; (ii) computationally, through a bottleneck identification algorithm, which identifies (some) delay unstable queues/flows in complex switched queueing networks by solving the fluid model from certain initial conditions.

1. Introduction

We study resource allocation problems arising in *switched queueing networks*, a class of stochastic models that are often used to capture the dynamics and decisions in data communication networks, e.g., cellular networks, Internet routers, and ad hoc wireless networks; see Neely [49], McKeown et al. [48], and Venkataraman et al. [65], but also in flexible manufacturing systems as in Gans and Van Ryzin [27] and cloud computing clusters as in Maguluri et al. [44]. A switched queueing network can be viewed as a collection of single-class; single-server; first come, first served (FCFS) queues whose service is interdependent, e.g., due to wireless interference constraints, matching constraints in a switch, or flow-scheduling constraints in a wireline network. Thus, only certain subsets of the set of queues, the so-called schedules, can be served simultaneously, giving rise to a fundamental resource allocation problem: which schedule to serve and at which point in time? Clearly, the overall performance of the network depends critically on the policy applied.

The focus of this paper is on a widely studied queue length-based policy, the *max-weight policy*. A remarkable property of the max-weight policy is its throughput optimality, i.e., the ability to stabilize the network whenever this is possible, without explicit information on the arriving traffic; see Tassiulas and Ephremides [62]. Thus, dynamic instability phenomena, such as the ones reported by Kumar and Seidman [40] and Rybko and Stolyar [55], do not arise.[1] Moreover, max-weight-type policies achieve very good delay performance under light-tailed traffic: they achieve optimal or order-optimal average delay for specific network topologies, e.g., Ganti et al. [28] and Neely [49]; optimal large deviations exponent, e.g., Venkataraman et al. [65]; and can be

RIGHTS LINK

asymptotically delay optimal in heavy traffic, e.g., Stolyar [60]. For these reasons, max-weight has become the benchmark for switched queueing networks.

Empirical evidence of high variability phenomena in data communication networks, manufacturing, and cloud computing (e.g., Park and Willinger [52], Fisher and Raman [24], and Ersoz et al. [23]) motivates us to study switched networks with a *mix of heavy-tailed and exponential-type traffic*. For the purposes of this paper, heavy-tailed traffic is defined in terms of arrival processes with infinite variance. Classical results in queueing theory, e.g., the Pollaczek-Khinchin formula and Kingman's bounds, imply that FCFS queues receiving heavy-tailed traffic are delay unstable, i.e., they experience infinite expected delays in steady state. Thus, our focus is on the impact of heavy-tailed traffic on queues that receive exponential-type traffic, using delay instability as a proxy for large delays and exponentially decaying upper bounds on queue-length tail asymptotics as a proxy for low delays. And while there is sizeable literature on the stability properties of max-weight, as well as its delay performance under light-tailed traffic, its delay performance in the presence of heavy-tailed traffic is not equally well understood.

There is a vast literature on the impact of heavy-tailed traffic in a variety of queueing systems, most notably in single-class queueing systems: the $G/G/1$ queue under FCFS, e.g., Borovkov [6], Cohen [18], Pakes [51], and Veraverbeke [66]; the $M/G/1$ queue under processor sharing, e.g., Zwart and Boxma [68], Nunez-Queija [50], and Jelenkovic and Momcilovic [36]; the multiserver $G/G/s$ queue under FCFS, e.g., Foss and Korshunov [25]; and fluid queues, e.g., Borst and Zwart [7], Boxma and Dumas [13], Jelenkovic and Lazar [34], Likhanov and Mazumdar [42], Resnick and Samorodnitsky [54], and Zwart et al. [69]. More recently, several works have studied the impact of heavy tails in multiclass queues: the multiclass $M/G/1$ queue under generalized processor sharing, e.g., Borst et al. [8], Jelenkovic and Momcilovic [35], Kotopoulos et al. [39], and Lelarge [41]; and discriminatory processor sharing, e.g., Avrachenkov et al. [2], Borst et al. [10, 11], and Rege and Sengupta [53]. Finally, there have been some attempts to analyze the impact of heavy tails in network settings, e.g., networks of generalized processor sharing queues in Van Uitert and Borst [64], networks of fluid queues in D'Auria and Samorodnitsky [21], generalized Jackson networks in Baccelli et al. [4], and monotone separable networks in Baccelli and Foss [3] (a class of stochastic systems that includes certain multiserver queues, Jackson networks, and polling systems as special cases). The above works offer a wealth of insights regarding the effect of heavy tails in different queueing systems, and also propose several methodological avenues for analysis. However, the distinctive characteristic of switched queueing networks that the activity of different servers is interdependent, as well as the complex dynamics imposed by the (queue length-based) max-weight policy are absent from these prior works, making the model under study and the methodological approach quite different.

Closer to our work come the papers by Borst et al. [9] and by Jagannathan et al. [33], both of which consider a system with two "parallel" FCFS queues receiving heavy-tailed and exponential-type traffic, while sharing a single server. The authors determine the queue-length tail asymptotics of the generalized processor sharing policy and the generalized max-weight policy, respectively. Also related to our work is the paper by Boxma et al. [14], which analyzes a $M/G/2$ FCFS queue with a heavy-tailed and an exponential-type server, and studies the dependence of the queue-length tail asymptotics on the arrival rate.

The present paper builds on Markakis et al. [46], which considers a single-hop switched queueing network with a mix of heavy-tailed and light-tailed traffic, under the max-weight policy; a brief discussion of the findings of our earlier work and a detailed account of the contributions of the present work is given in the next two subsections. The companion paper Markakis et al. [47] uses technical results derived here (more specifically, Theorems 4 and 5 in Section 4) to study how the network topology, the routing constraints, and the link capacities affect the delay performance of multihop switched queueing networks with heavy-tailed traffic under back-pressure-type policies.

Finally, there is growing literature on fluid models of the max-weight and back-pressure policies in a variety of settings, e.g., single-hop and multihop switched queueing networks, as well as stochastic processing networks; see Bui et al. [17], Dai and Lin [20], Ji et al. [38, 37], Liu et al. [43], and Shah and Wischik [56]. Although the present paper employs very similar fluid models and in that sense builds on these prior works, our objective is quite different: fluid approximations have been used in existing literature in order to prove the stability of the corresponding queueing networks or state-space collapse phenomena under critical loads, while in the present paper to facilitate a delay analysis in the presence heavy-tailed traffic. The work of Baccelli et al. [4] deserves a special mention as it uses fluid models of generalized Jackson networks with subexponential service times to determine the precise tail asymptotics of the steady-state maximal dater, i.e., the time to clear all customers present at time t, assuming arrivals are stopped from that point on, in the limit as t goes to infinity. The tail asymptotics are determined through a sample-path construction of the maximal dater, which preserves crucial monotonicity properties of Jackson networks. In contrast, our approach is based on stochastic Lyapunov theory

Figure 1. (Color online) Delay performance of the max-weight policy under heavy-tailed traffic through a simple example. Queue 1 receives heavy-tailed traffic while queues 2 and 3 receive exponential-type traffic. Queues 1 and 2 can be served simultaneously, whereas queue 3 can only be served alone. Max-weight compares the length of queue 3 to the sum of the lengths of queues 1 and 2, and serves the heavier schedule. Queue 1 is delay unstable under any scheduling policy, and queue 3 is delay unstable under the max-weight policy. Queue 2 may or may not be delay stable under the max-weight policy, depending on the arrival rates.
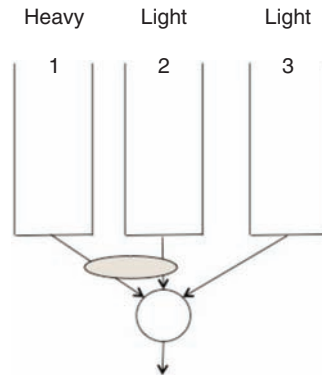
and renewal theory, which on the one hand do not provide as refined results, i.e., moment bounds instead of tail asymptotics, but on the other hand do not rely on any special structure besides Markovianity. Therefore, we are able to obtain results for queueing systems with more complex (nonmonotonic) dynamics, and for more refined steady-state quantities such as queue lengths and delays (cf. maximal dater).

1.1. Motivating Example

We motivate the subsequent development by presenting the main findings of Markakis et al. [46] through a simple example. Consider the queueing system of Figure 1, which includes three queues, indexed 1, 2, and 3, and operates in discrete time. Queues 1 and 2 can be served simultaneously (so, in the terminology of switched queueing networks, they constitute schedule $\{1, 2\}$) whereas queue 3 can only be served alone. Queues are served at unit rate whenever the respective schedules are activated, and the service discipline within each queue is FCFS. Each queue buffers the traffic of a dedicated batch arrival process. Arrivals are independent across queues, and independent, identically distributed (IID) across time slots within each queue. We further assume that arrivals to queue 1 are heavy tailed, whereas arrivals to queues 2 and 3 are exponential type. In this setting, the max-weight policy compares the length of queue 3 to the sum of the lengths of queues 1 and 2 and serves the "heavier" schedule at each time slot.

The delay stability of queue 1 does not depend on the specifics of the queueing system at hand or on the scheduling policy applied, but merely on the fact that it receives heavy-tailed traffic. In the best-case scenario, i.e., when it is served at every time slot, queue 1 is equivalent to a $M/G/1$ queue with infinite second moment of service time. Classical results in queueing theory, e.g., the Pollaczek-Khinchin formula, imply that its expected steady-state delay is infinite. Therefore, queue 1 is delay unstable even in the best case. This observation can be generalized: in a single-hop switched queueing network, a heavy-tailed queue is delay unstable under any scheduling policy.[2]

Coming to queue 3, note that it cannot be served simultaneously with queue 1, so in some sense "conflicts" with it. Thus, under the max-weight policy, queue 3 will not be served unless its length is greater than or equal to the length of queue 1. However, queue 1 is, occasionally, very long due to its heavy-tailed arrivals. On those occasions, queue 3 has to build up to a comparable length, leading to delay instability. This observation can be generalized as well: in a single-hop switched queueing network under the max-weight policy, a queue that conflicts with a heavy-tailed queue is delay unstable.

The most interesting findings though, concern the delay stability of queue 2. One would expect that this queue is delay stable as it is exponential-type itself and it is served together with a heavy-tailed queue, which should result in more service opportunities under the max-weight policy. Surprisingly, if its arrival rate is sufficiently high (but still in the stability region of the system), queue 2 is delay unstable. The key observation is that even though queue 2 does not conflict with a heavy-tailed queue, it does conflict with queue 3, which is delay unstable because it conflicts with queue 1. Conversely, queue 2 is delay stable if its arrival rate is sufficiently low.

Proposition 1 (Rate-Dependent Delay Instability, Markakis et al. [46])**.** *Consider the queueing system of Figure* 1 *under the max-weight policy, with arrival rates in its stability region. If the arrival rates satisfy* $\lambda_2 > (1 + \lambda_1 - \lambda_3)/2$, *then queue* 2 *is delay unstable.*

Proposition 2 (Rate-Dependent Delay Stability, Markakis et al. [46])**.** *Consider the queueing system of Figure* 1 *under the max-weight policy, with arrival rates in its stability region. If the arrival rates satisfy* $\lambda_2 < (1 + \lambda_1 - \lambda_3)/2$, *then queue* 2 *is delay stable and its steady-state queue length is exponential type.*

Propositions 1 and 2 provide a sharp characterization of the *delay stability region* of queue 2, i.e., the set of arrival rates for which queue 2 is delay stable. Earlier proofs of these results are based on purely stochastic arguments, and are somewhat long and tedious. We will show that the use of fluid approximations considerably simplifies the delay analysis, allowing us to extend the findings of Markakis et al. [46].

1.2. Methodological Challenges and Main Contributions

The problem of delay analysis of the max-weight policy in the presence of heavy-tailed traffic poses a number of methodological challenges. Dynamic programming formulations of scheduling problems in queueing systems are analytically intractable and have prohibitive computational requirements in most cases. Monte Carlo methods can be very slow to converge, or may even fail to converge at all due to the very nature of heavy tails (processes with infinite variance). Finally, the complex dynamics imposed by the dependence of server activity on queue lengths hinder the application of "standard" approaches such as stochastic comparisons, transform methods, or sample-path arguments; for an excellent survey of different methods used in the analysis of queueing systems with heavy tails the reader is referred to Borst et al. [12].

The main contribution of this paper is to show how fluid approximations can facilitate a delay analysis of switched queueing networks with heavy-tailed traffic. We use fluid approximations and renewal theory in order to prove delay instability results. Furthermore, we show how fluid approximations can be combined with stochastic Lyapunov theory in order to prove delay stability results. More importantly, we identify a novel class of Lyapunov functions, whose drift analysis can provide exponential upper bounds on queue-length tail asymptotics for exponential-type queues/flows even in the presence of heavy tails at other queues.

More specifically, a standard way of showing that queues exhibit low delays in queueing systems with complex dynamics, e.g., upper bounds on queue-length/delay tail asymptotics or on the corresponding expected values, is drift analysis of suitable Lyapunov functions, since direct stochastic comparisons or coupling arguments are usually helpful only in simpler settings. Unfortunately, popular candidates such as standard piecewise linear functions, e.g., Bertsimas et al. [5] and Down and Meyn [22]; quadratic functions, e.g., Tassiulas and Ephremides [62]; and norms, e.g., Shah et al. [57] and Venkataraman et al. [65], cannot be used under heavy-tailed traffic. This is because they are increasing in all queue lengths, which implies that their steady-state expectation is infinite in the presence of heavy-tailed traffic, rendering their drift analysis uninformative. Our approach to this problem is as follows: we identify a novel class of Lyapunov functions that are *nonincreasing* in the lengths of heavy-tailed queues, *piecewise linear* and, thus, akin to the dynamics imposed by max-weight, and whose drift analysis helps obtain exponential upper bounds on queue-length tail asymptotics; see Equation (36). However, drift analysis of piecewise linear functions can be a challenge on its own, due to the fact that the stochastic descent property is often lost at locations where the function is nondifferentiable. We show how fluid approximations can help overcome this difficulty. Critical to the latter is a connection between fluid approximations and Lyapunov theory: for the class of models considered in this paper, we show that if a function $V(\cdot)$ is continuous, piecewise linear, and a "Lyapunov function" for the fluid model, then $V(\cdot)$ is also a Lyapunov function for the original stochastic system.[3] This connection allows us to carry out the drift analysis in the fluid domain, which is typically much easier. Moreover, if $V(\cdot)$ has exponential type "upward jumps" in the stochastic system, then the results in Hajek [31] imply an exponentially decaying upper bound on the tail of its steady-state distribution.

On the other hand, showing that queues exhibit large delays, e.g., lower bounds on queue-length/delay tail asymptotics or on the corresponding expected values, often relies on sample-path techniques. However, tracking the evolution of sample paths can be hard when the system exhibits complex dynamics. This also hinders the use of transform methods, at least as a way to obtain analytical results. The main idea behind our approach is as follows: even when we are not able to analyze sample paths explicitly, we might still be able to do so approximately in terms of the solution to a fluid model from certain initial conditions of interest. Then, we can use renewal theory to translate sample path analysis to lower bounds on steady-state queue-length moments.

We illustrate the benefits of the proposed methodology in two ways:

(i) analytically, by studying the delay stability regions of single-hop switched queueing networks with disjoint schedules, providing a precise characterization of these regions for certain queues together with inner and outer bounds for the rest (Theorem 3);

(ii) computationally, through a bottleneck identification algorithm, which identifies (some) delay unstable queues by solving the fluid model from certain initial conditions. For all practical purposes the solution to the fluid model can be obtained numerically, allowing the application of the algorithm even to networks with quite complex topologies.

Finally, our analysis of networks with disjoint schedules sheds further light into the behavior of max-weight, a widely studied policy that has become the benchmark in switched queueing networks. We reveal monotonicity properties for the service rates of different schedules under the max-weight policy, which, in turn, allow us to identify "critical configurations" toward which the state of the system is driven (see Lemmas 5–7 in Appendix A), and that determine to a large extent delay stability (see the proof of Theorem 3). These insights could be a starting point toward a better understanding of the behavior of max-weight policies in more complex networks.

1.3. Outline of the Paper

The remainder of the paper is organized as follows. We begin with a detailed description of a single-hop switched queueing network under the max-weight policy, together with its natural fluid model and some useful definitions and notation, in Section 2. Sections 3.1 and 3.2 include the methodological contributions of the paper, i.e., how fluid approximations can facilitate a delay analysis of single-hop switched queueing networks with heavy-tailed traffic under the max-weight policy. Using these results, we study the delay stability regions of networks with disjoint schedules in Section 3.3, introducing, along the way, a novel class of Lyapunov functions that is suitable for the delay analysis of such systems. In Section 3.4 we present the bottleneck identification algorithm, which identifies (some) delay unstable queues by solving the fluid model of the network from specific initial conditions, accompanied by examples showcasing its applicability. Section 4 illustrates how the methodology developed and the results obtained in the context of single-hop networks can be extended to multihop switched queueing networks under the back-pressure policy. Section 5 concludes the paper with a brief discussion of our findings and directions for future research. Appendix A includes the statements and proofs of certain monotonicity properties of the service rates in networks with disjoint schedules under the max-weight policy, which facilitate the delay analysis in Section 3.3. Finally, we have collected in Appendix B technical results that facilitate the delay analysis in Section 4.

2. A Single-Hop Switched Queueing Network Under the Max-Weight Policy

In this section we provide a detailed presentation of the first of the two queueing models to be considered in this paper, together with some necessary definitions and notation.

We denote by $\mathbb{R}_+$, $\mathbb{Z}_+$, and $\mathbb{N}$ the sets of nonnegative reals, nonnegative integers, and positive integers, respectively. The Cartesian products of M copies of $\mathbb{R}_+$ and $\mathbb{Z}_+$ are denoted by $\mathbb{R}_+^M$ and $\mathbb{Z}_+^M$, respectively. With few exceptions, we follow the convention of using lowercase letters to denote real numbers or vectors, and uppercase letters to denote random variables or events. We use $[x]^+$ for $\max\{x, 0\}$, the nonnegative part of $x \in \mathbb{R}$. Similarly, we use $[x]^-$ for $\min\{x, 0\}$, the nonpositive part of $x \in \mathbb{R}$.

The indicator variable of event E is represented by 1_E. The notation $\mathbb{P}(\cdot)$ and $\mathbb{E}[\cdot]$ is used for probabilities and expectations, respectively. We also employ the shorthand notation $\mathbb{P}(X; E \mid \mathcal{H})$ for $\mathbb{P}(X \cdot 1_E \mid \mathcal{H})$, where X is a random variable, E is an event, and $\mathcal{H}$ is a σ-algebra on a given probability space. We define $\mathbb{E}[X; E \mid \mathcal{H}]$ similarly.

We consider a discrete time switched queueing network, where arrivals occur at the end of each time slot. Let $\mathcal{F} = \{1, \ldots, F\}$, $F \in \mathbb{N}$. Central to our model is the notion of a *traffic flow* $f \in \mathcal{F}$, which is a long-lived stream of traffic that arrives to the network according to a discrete time stochastic batch arrival process $\{A_f(t); t \in \mathbb{Z}_+\}$. We assume that all arrival processes take values in $\mathbb{Z}_+$, and are IID over time. Furthermore, different arrival processes are mutually independent. We denote by $\lambda_f = \mathbb{E}[A_f(0)] > 0$ the arrival rate of traffic flow f and by $\lambda = (\lambda_f; f = 1, \ldots, F)$ the vector of arrival rates of all traffic flows.

Definition 1 (Heavy/Light Tails). A nonnegative random variable X is *heavy tailed* if $\mathbb{E}[X^2]$ is infinite, and is *light tailed* otherwise. Moreover, X is *exponential type* if there exists $\theta > 0$ such that $\mathbb{E}[\exp(\theta X)] < \infty$.

We define similarly a heavy-tailed/light-tailed/exponential-type traffic flow. We note that there are several definitions of heavy/light tails in the literature. In fact, a random variable is often defined as light tailed if it

is of exponential type, and heavy tailed otherwise. The definition adopted in this paper has been used in the area of data communication networks, e.g., Park and Willinger [52], due to its close connection to long-range dependence.

For technical reasons, we assume throughout the paper the existence of some $\gamma \in (0,1)$ such that

$$\mathbb{E}[A_f^{1+\gamma}(0)] < \infty, \quad \text{for all } f \in \mathcal{F}. \tag{1}$$

In the first part of the paper we consider a switched queueing network with single-hop traffic flows, i.e., the traffic of flow f is buffered in a dedicated single-server queue (queue f and server f, henceforth), eventually gets served, and then exits the system. Our modeling assumptions imply that the set of traffic flows can be identified with the set of queues and the set of servers of the network. The service discipline within each queue is assumed to be FCFS. The stochastic process $\{Q_f(t); t \in \mathbb{Z}_+\}$ captures the evolution of the length of queue f. Since our main motivation comes from data communication networks, $A_f(t)$ will be interpreted as the number of packets that queue f receives at the end of time slot t, and $Q_f(t)$ as the total number of packets in queue f at the beginning of time slot t. The arrivals and the lengths of the various queues at time slot t are captured by the vectors $A(t) = (A_f(t); f = 1, \ldots, F)$ and $Q(t) = (Q_f(t); f = 1, \ldots, F)$, respectively.

In the context of data communication networks, a batch of packets arriving to a queue at any given time slot can be viewed as a single entity, e.g., as a file that needs to be transmitted. We define the *end-to-end delay of a file* of flow f to be the number of time slots that the file spends in the network, starting from the time slot right after it arrives at queue f, until the time slot that its last packet gets served. For $k \in \mathbb{N}$, we denote by $D_f(k)$ the end-to-end delay of the kth file of flow f, and use the vector notation $D(k) = (D_f(k); f = 1, \ldots, F)$.

The salient feature of a switched queueing network is that not all servers can be simultaneously active, e.g., due to interference in wireless networks or matching constraints in a switch. Consequently, not all traffic flows can be served simultaneously. A set of traffic flows that can be served simultaneously is called a *schedule*. We denote by $\mathcal{S}$ the set of all schedules, which is assumed to be an arbitrary subset of the powerset of $\mathcal{F}$. For simplicity, we assume that all packets have the same size, and that the service rate of all servers is equal to one packet per time slot. We denote by $S_f(t) \in \{0,1\}$ the number of packets that are scheduled for service from queue f at time slot t. Note that this is not necessarily equal to the number of packets that are actually served, because the queue may be empty. We use the vector notation $S(t) = (S_f(t); f = 1, \ldots, F)$. For convenience, we also identify schedules with vectors in $\{0,1\}^F$.

Using the notation above, the *dynamics* of queue f take the form

$$Q_f(t+1) = Q_f(t) + A_f(t) - S_f(t) \cdot 1_{\{Q_f(t)>0\}}.$$

The vector of initial queue lengths $Q(0)$ is assumed to be an arbitrary element of $\mathbb{Z}_+^F$.

The service vector $S(t)$ is determined by the scheduling policy applied to the network. We focus on the *max-weight policy*, where the scheduling vector $S(t)$ satisfies

$$S(t) \in \arg\max_{(S_f) \in \mathcal{S}} \left\{ \sum_{f \in \mathcal{F}} Q_f(t) \cdot S_f \right\},$$

at any given time slot. If the set on the right-hand side includes multiple schedules, then one of them is chosen uniformly at random.

As alluded to in the introduction, a very appealing property of the max-weight policy is throughput optimality, namely, the ability to stabilize (in the sense of the definition that follows) a switched queueing network whenever this is possible.

Definition 2 (Stability). A switched queueing network, operated under a particular policy, is *stable* 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 the initial queue lengths $Q(0)$.

Under a stabilizing scheduling policy, we denote by $Q = (Q_f; f = 1, \ldots, F)$ and $D = (D_f; f = 1, \ldots, F)$ generic random vectors distributed according to the limiting distributions of $\{Q(t); t \in \mathbb{Z}_+\}$ and $\{D(k); k \in \mathbb{N}\}$, respectively. We refer to Q_f as the steady-state length of queue f. Similarly, we refer to D_f as the steady-state delay of a file of traffic flow f. To ease the notation, we have suppressed the dependence of these limiting distributions on the scheduling policy applied.

The ability to stabilize a switched queueing network depends on the arrival rates of the various traffic flows relative to the service rates of the servers, and on the scheduling constraints. This relation is captured by the stability region of the network.

Definition 3 (Stability Region of Single-Hop Network)**.** The *stability region* of the single-hop switched queueing network described earlier is the set of arrival rate vectors

$$\left\{ \lambda \in \mathbb{R}_+^F \mid \exists\, \zeta_s \in \mathbb{R}_+, s \in \mathscr{S} : \lambda \le \sum_{s \in \mathscr{S}} \zeta_s \cdot s, \sum_{s \in \mathscr{S}} \zeta_s < 1 \right\}.$$

Lemma 1 (Throughput Optimality of Max-Weight)**.** *Consider the single-hop switched queueing network described above under the max-weight policy. The network is stable for any arrival rate vector in the stability region.*

Proof. For the case of light-tailed traffic, this result follows from the findings in Tassiulas and Ephremides [62]; in the presence of heavy-tailed traffic, it follows from Proposition 2 of Stolyar [60]. For a formal proof the reader is referred to Lemma 4.1 in Markakis [45]. □

Finally, we define the property that we use to evaluate the delay performance of max-weight.

Definition 4 (Delay Stability)**.** Traffic flow f is *delay stable* if the switched queueing network is stable and $\mathbb{E}[D_f]$ is finite; otherwise, f is *delay unstable*.

2.1. Fluid Model of the Network

In this section we give some background material on the natural fluid model of the single-hop switched queueing network described above under the max-weight policy. The fluid model is a continuous time deterministic dynamical system, which aims to capture the evolution of its stochastic counterpart on longer time scales by taking advantage of laws of large numbers. Initially, we give a brief description and some useful properties of the fluid model. Then, we introduce the notion of fluid scaling, and establish a formal connection between the deterministic and a "fluid-scaled" version of the stochastic system.

The *fluid model* (FM) of a single-hop switched queueing network under the max-weight policy is defined by the set of ordinary differential equations and inequalities in Equations (2)–(7), for every time $t \ge 0$ for which the derivatives exist (such t is often called a regular time):

$$\dot{q}_f(t) = \lambda_f - \sum_{\pi \in \mathscr{S}} \dot{s}_\pi(t)\pi_f + \dot{y}_f(t), \quad \forall\, f \in \mathscr{F}; \tag{2}$$

$$\dot{s}_\pi(t) \ge 0, \quad \forall\, \pi \in \mathscr{S}; \tag{3}$$

$$\sum_{\pi \in \mathscr{S}} \dot{s}_\pi(t) = 1; \tag{4}$$

$$0 \le \dot{y}_f(t) \le \sum_{\pi \in \mathscr{S}} \dot{s}_\pi(t)\pi_f, \quad \forall\, f \in \mathscr{F}; \tag{5}$$

$$q_f(t) > 0 \implies \dot{y}_f(t) = 0, \quad \forall\, f \in \mathscr{F}; \tag{6}$$

$$\sum_{f \in \mathscr{F}} q_f(t)\pi_f < \max_{\sigma \in \mathscr{S}} \left\{ \sum_{f \in \mathscr{F}} q_f(t)\sigma_f \right\} \implies \dot{s}_\pi(t) = 0, \quad \forall\, \pi \in \mathscr{S}. \tag{7}$$

In the equations above, $q(t)$ represents the vector of queue lengths at time t, $y(t)$ represents the vector of cumulative idling/wasted service up to time t, and $s_\pi(t)$ represents the total amount of time that schedule π has been activated up to time t. Equation (4) states that a schedule is to be picked at each time, and Equation (6) that there can be no wasted service when queue lengths are positive. Finally, Equation (7) is the natural analogue of the max-weight policy in the fluid domain: schedules that do not have maximum weight receive no service.

Fix some arbitrary $T > 0$. A *fluid model solution* (FMS) from initial condition $q(0) = q$ is a Lipschitz continuous function $x(\cdot) = (q(\cdot), y(\cdot), s(\cdot))$ that satisfies (i) $x(0) = (q, 0, 0)$; and (ii) Equations (2)–(7) over the subset of $[0, T]$ where $q(\cdot)$ is differentiable.

An FMS is differentiable almost everywhere (equivalently, almost every $t \in [0, T]$ is a regular time), since it is Lipschitz continuous by assumption.

Next, we define the notion of fluid scaling and establish the existence of a fluid limit and of an FMS. Consider a sequence of initial queue lengths $\{Q^b(0); b \in \mathbb{N}\}$ for the queueing system described above, and the corresponding sequence of queue-length processes $\{Q^b(\cdot); b \in \mathbb{N}\}$. While the original processes $Q(\cdot)$ and $Q^b(\cdot)$ are defined for integer times, we extend them to piecewise constant functions of continuous time by setting $Q(t) = Q(\lceil t \rceil)$, and similarly for $Q^b(\cdot)$.

We define the "fluid-scaled" queue-length process as

$$\tilde{q}^b(t) = \frac{Q^b(bt)}{b}, \quad t \in [0, T], b \in \mathbb{N}. \tag{8}$$

We assume the existence of a vector $q \in \mathbb{R}_+^F$ and of a sequence of positive numbers $\{\epsilon_b; b \in \mathbb{N}\}$, converging to zero as b goes to infinity, that satisfy

$$\max_{f \in \mathcal{F}} |\tilde{q}_f^b(0) - q_f| \leq \epsilon_b, \quad \forall\, b \in \mathbb{N}. \tag{9}$$

We recall our standing assumption that there exists $\gamma \in (0,1)$ so that all traffic flows have $(1 + \gamma)$ moments. Fix some $\gamma' \in (0, \gamma)$ and consider the sequence of sets of sample paths of the arrival processes defined by

$$H_b = \left\{ \omega : \max_{f \in \mathcal{F}} \max_{1 \leq t \leq bT} \frac{1}{bT} \left| \sum_{\tau=0}^{t-1} A_f(\tau) - \lambda_f t \right| < (bT)^{-\gamma'/(1+\gamma)} \right\}, \quad b \in \mathbb{N}. \tag{10}$$

Intuitively, H_b contains those sample paths of the arrival processes that stay close to their average behavior over the time interval $[0, bT]$.

Lemma 2 (Existence of Fluid Limit and FMS)**.** *There exists a Lipschitz continuous function $z(t) = (z_1(t), \ldots, z_F(t)), t \in [0,T]$, and for every $\epsilon > 0$ some $b_0(\epsilon)$, so that*

$$\mathbb{P}(H_b) \geq 1 - \epsilon, \ \forall\, b \geq b_0(\epsilon), \quad and \quad \sup_{t \in [0,T]} \max_{f \in \mathcal{F}} |\tilde{q}_f^b(t) - z_f(t)| \leq \epsilon, \quad \forall\, \omega \in H_b, \ \forall\, b \geq b_0(\epsilon).$$

Additionally, there exist Lipschitz continuous functions $v(\cdot)$ and $w(\cdot)$, such that $(z(\cdot), v(\cdot), w(\cdot))$ is an FMS from initial condition $q(0) = q$ over the interval $[0,T]$.

Proof. Let us first establish the convergence of $\mathbb{P}(H_b)$. The Marcinkiewicz-Zygmund strong law of large numbers implies that

$$\frac{\sum_{\tau=0}^{t-1} A_f(\tau) - \lambda_f t}{t^{1/(1+\gamma)}} \xrightarrow{\mathscr{L}_1} 0, \quad \forall\, f \in \mathcal{F};$$

see Theorem 10.3 in Gut [30]. Consequently, for any fixed $c > 0$ there exists $t_0(c)$, such that

$$\mathbb{E}\left[\left| \sum_{\tau=0}^{t-1} A_f(\tau) - \lambda_f t \right| \right] \leq c\, t^{1/(1+\gamma)}, \quad t \geq t_0(c), \forall\, f \in \mathcal{F}.$$

Notice that the sequence $\{\sum_{\tau=0}^{t-1} A_f(\tau) - \lambda_f t; t \in \mathbb{N}\}$ is a martingale, for every $f \in \mathcal{F}$. Thus, the sequence $\{\sum_{f \in \mathcal{F}} |\sum_{\tau=0}^{t-1} A_f(\tau) - \lambda_f t|; t \in \mathbb{N}\}$ is a nonnegative submartingale.

Let $r = bT$ and $\delta_r = r^{-\gamma'/(1+\gamma)}$. If b is sufficiently large, then $r \geq t_0(c)$. Then, Doob's submartingale inequality (e.g., see Section 14.6 of Williams [67]) and the Marcinkiewicz-Zygmund strong law imply that

$$1 - \mathbb{P}(H_b) = \mathbb{P}\left(\sup_{1 \leq t \leq r} \frac{1}{r} \max_{f \in \mathcal{F}} \left| \sum_{\tau=0}^{t-1} A_f(\tau) - \lambda_f t \right| \geq \delta_r \right) \leq \mathbb{P}\left(\sup_{1 \leq t \leq r} \frac{1}{r} \sum_{f \in \mathcal{F}} \left| \sum_{\tau=0}^{t-1} A_f(\tau) - \lambda_f t \right| \geq \delta_r \right)$$

$$\leq \frac{1}{\delta_r r} \sum_{f \in \mathcal{F}} \mathbb{E}\left[\left| \sum_{\tau=0}^{r-1} A_f(\tau) - \lambda_f r \right| \right] \leq cF \cdot \frac{r^{1/(1+\gamma)}}{r^{1-\gamma'/(1+\gamma)}} = cF \cdot r^{-(\gamma-\gamma')/(1+\gamma)}. \tag{11}$$

As b goes to infinity, r goes to infinity. Since $\gamma' < \gamma$, it follows that $\mathbb{P}(H_b)$ converges to one.

The existence of a fluid limit and the fact that a fluid limit is an FMS follow directly from Theorem 4.3 in Shah and Wischik [56] with the following correspondences. Our q_0 corresponds to $\mathbf{q}_0$ in Shah and Wischik [56]. Our FMS from initial condition q_0 corresponds to FMS($\mathbf{q}_0$) in Shah and Wischik [56]. Our b corresponds to both j and z in Shah and Wischik [56]. In particular, ϵ_j in Shah and Wischik [56] is identified with our $(bT)^{-\gamma'/(1+\gamma)}$, ϵ'_j in Shah and Wischik [56] is identified with our ϵ_b, and our set H_b corresponds to the set G_j in Shah and Wischik [56]. Condition (25) in Shah and Wischik [56] is simply the requirement that the arrival sample path belongs to H_b. $\square$

Lemma 3 (Uniqueness and Continuity of FMS)**.** *For any given $q \in \mathbb{R}_+^F$ there exists a (unique) Lipschitz continuous function $z(t) = (z_1(t), \ldots, z_F(t))$, $t \in [0,T]$, such that the queue-length part of every FMS from initial condition q is $z(\cdot)$. Moreover, $z(\cdot)$ depends continuously on both the initial condition q and the arrival rate vector λ.*

Proof. The existence of an FMS was established in Lemma 2. The uniqueness of the queue-length part of the FMS was proved in Subramanian [61], by first showing that the max-weight policy is a maximal monotone map from the space of queue lengths to the space of scheduling vectors, and then invoking known properties of such maps. A more direct proof of this result can be found in Appendix 5.1 of Markakis [45]. $\square$

We note that the above lemma does not guarantee the uniqueness of the FMS as a whole, but only the uniqueness of the queue-length part. Namely, there may be multiple Lipschitz continuous functions for the service and idleness parts of the solution that satisfy the FM equations. In fact, one can construct simple examples where the FMS from zero initial condition is not unique.

3. Delay Analysis via Fluid Approximations

The current section includes the most important findings of the paper. Sections 3.1 and 3.2 present our method-ological contributions, i.e., how fluid approximations can facilitate a delay analysis of single-hop switched queueing networks with heavy-tailed traffic under the max-weight policy. Using these results, we provide an in-depth analysis of the delay stability regions of networks with disjoint schedules in Section 3.3. Finally, in Section 3.4 we introduce the bottleneck identification algorithm, which identifies (some) delay unstable queues by solving the fluid model of the network from specific initial conditions, together with examples illustrating its applicability.

3.1. Delay Instability via Fluid Approximations

In this section we show how fluid approximations can be used for proving delay instability results. Our contri-bution is summarized in the following theorem, which provides a sufficient condition for the delay instability of queues/flows.

Theorem 1. *Consider the single-hop switched queueing network of Section 2 under the max-weight policy, and its natural FM, i.e., Equations (2)–(7). Let $h \in \mathcal{F}$ be a heavy-tailed traffic flow, and $q^*(\cdot)$ be the (unique) queue-length part of an FMS from initial condition $q_h^*(0) = 1$ and $q_f^*(0) = 0$, for all $f \neq h$. If there exists $\tau \in [0, T]$ such that $q_j^*(\tau) > 0$, then traffic flow j is delay unstable.*

Proof. Let us first look at the evolution of the system when it starts from a large initial condition for the heavy-tailed queue h. Specifically, consider a sequence of single-hop switched queueing networks, indexed by $b \in \mathbb{N}$, with initial queue lengths $Q_h^b(0) = b$ and $Q_f^b(0) = 0$, for all $f \neq h$. Let $\{Q^b(\cdot); b \in \mathbb{N}\}$ be the sequence of (unscaled) queue-length processes under the max-weight policy. We define a corresponding sequence of scaled queue-length processes by letting

$$\tilde{q}^b(t) = \frac{Q^b(bt)}{b}.$$

Instead of studying directly the process $Q^b(\cdot)$, we will exploit the fact that its scaled version behaves as a simpler, deterministic fluid model for sufficiently large b.

The initial condition of the scaled processes, and of the corresponding fluid model, is one for queue h and zero for all other queues. Lemma 3 implies that, for the given initial condition, there exists a unique queue-length part for every FMS, which we denote by $q^*(\cdot)$.

Fix $j \in \mathcal{F}$ and suppose that there exist $\epsilon, \tau > 0$, such that

$$q_j(\tau) > \epsilon. \tag{12}$$

Lemma 2 implies that there exists some finite b_0 such that for all $b \geq b_0$,

$$\mathbb{P}(H_b) \geq 1 - \epsilon, \tag{13}$$

and

$$|\tilde{q}_j^b(\tau) - q_j(\tau)| \leq \frac{\epsilon}{2}, \quad \forall \omega \in H_b. \tag{14}$$

(Strictly speaking b_0 is a function of ϵ, but to make the notation simpler we suppress this dependence.) Equa-tions (12) and (14) imply that

$$Q_j^b(b\tau) \geq \frac{\epsilon b}{2}, \quad \forall \omega \in H_b, \forall b \geq b_0. \tag{15}$$

In the remainder of the proof we show that (i) the particular initial condition can be reached with positive probability; (ii) the fact that queue j builds up to order b with positive probability implies the delay instability of traffic flow j. The main idea is that queue j will take order $\Omega(b)$ time to be drained, so that the integral of its length over a busy period is of order $\Omega(b^2)$. Averaging over all possible values of b, and using the assumption that b is drawn from a heavy-tailed distribution, renewal theory implies that the steady-state length of queue j has infinite expectation.

We note that under the max-weight policy, the sequence of time slots that initiate busy periods of the system constitute a renewal process in which the interrenewal intervals have finite expectation, due to the positive recurrence of the original process. We define an instantaneous reward on this renewal process:

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

where M is a positive integer.

Let us focus on a particular busy period of the system, which, without loss of generality, starts at time slot zero. Consider the set of sample paths

$$\hat{H}_b = \{A_h(0) = b\} \cap \{A_f(0) = 0, f \neq h\} \cap H_b.$$

Since

$$1 - \mathbb{P}(A_f(0) = 0) = \mathbb{P}(A_f(0) > 0) \leq \mathbb{E}[A_f(0)] = \lambda_f,$$

and $\lambda_f < 1$ from stability, we have that

$$\mathbb{P}(A_f(0) = 0) > 0, \quad \forall f \neq h.$$

Let $\mathcal{B}_h \subset \mathbb{Z}_+$ be the support of the distribution of $A_h(0)$. Using the independence of the arrival processes, and taking into account Equation (13), we have that

$$\mathbb{P}(\hat{H}_b) \geq (1 - \epsilon)\mathbb{P}(A_h(0) = b) \prod_{f \neq h} \mathbb{P}(A_f(0) = 0) > 0, \quad \forall b \in \{b \in \mathcal{B}_h : b \geq b_0\}. \tag{16}$$

Regarding the unique queue-length part of every FMS from the initial condition of interest, once $q^*(t)$ becomes zero it stays at zero. This fact together with Lemma 2 can be used to conclude that for the sample paths in $\hat{H}_b$ and for b sufficiently large, queue h is nonempty throughout the interval $(0, b\tau]$. Thus, time slot $b\tau$ belongs to the busy period that started at time slot zero.

Since at most one packet departs from queue j at each time slot, Equation (15) implies that the length of queue j is at least $\epsilon b/4$ packets over a time period of duration $\epsilon b/4$ time slots. Thus, the aggregate reward R_{agg}^M, i.e., the reward accumulated over a renewal period, satisfies the lower bound:

$$R_{agg}^M \cdot 1_{\{b \geq b_0\}} \cdot 1_{\hat{H}_b} \geq \min\left\{ \left(\frac{\epsilon b}{4}\right)^2 \cdot 1_{\{b \geq b_0\}}, M \right\} \cdot 1_{\hat{H}_b}.$$

Then, taking into account Equation (16), the expected aggregate reward is bounded from below by

$$\mathbb{E}[R_{agg}^M] \geq (1 - \epsilon) \prod_{f \neq h} \mathbb{P}(A_f(0) = 0) \sum_{b \in \mathcal{B}_h} \min\left\{ \left(\frac{\epsilon b}{4}\right)^2 \cdot 1_{\{b \geq b_0\}}, M \right\} \mathbb{P}(A_h(0) = b).$$

So, there exists a positive constant ϵ' such that

$$\mathbb{E}[R_{agg}^M] \geq \epsilon' \mathbb{E}\left[\min\left\{ \left(\frac{\epsilon A_h(0)}{4}\right)^2 \cdot 1_{\{A_h(0) \geq b_0\}}, M \right\} \right].$$

Then, the monotone convergence theorem (e.g., see Section 5.3 of Williams [67]), together with a renewal theorem, implies that $\mathbb{E}[Q_j]$ is infinite. Finally, the Bernoulli Arrivals See Time Averages property and Little's Law (see Theorems 2.1 and 2.2 of Markakis [45], respectively, for precise statements of these well-known results in the context of switched queueing networks) imply the desired result, namely, that $\mathbb{E}[D_j]$ is also infinite. $\quad\square$

Remark 1. Theorem 1 holds for any choice of $T > 0$ (the horizon of the FMS). However, the fact that a single-hop switched queueing network is stable under the max-weight policy (Lemma 1) implies the existence of some $T^* > 0$, proportional to the initial condition of the FMS, such that $q(t) = 0$, for all $t > T^*$. Consequently, the most effective application of Theorem 1 is when T is chosen large enough so that the FMS "drains" within this horizon.

3.2. Delay Stability via Fluid Approximations

In this section we shift our attention to delay stability results in networks that receive a mix of heavy-tailed and exponential-type traffic. Typically, proving that queues experience low delays is either based on coupling arguments, if the underlying dynamics are relatively simple or, more often, on drift analysis of suitable Lyapunov functions. We focus on the latter approach. The presence of heavy-tailed traffic, though, introduces an additional complication: popular candidate Lyapunov functions such as standard increasing piecewise linear functions, e.g., Bertsimas et al. [5] and Down and Meyn [22]; quadratic functions, e.g., Tassiulas and Ephremides [62]; and norms, e.g., Shah et al. [57] and Venkataraman et al. [65], cannot be used because the steady-state expectation of these functions is infinite under heavy-tailed traffic, rendering drift analysis uninformative.

We introduce a class of piecewise linear Lyapunov functions that are nonincreasing in the length of the heavy-tailed queues, and that can provide exponential upper bounds on queue-length tail asymptotics even in the presence of heavy-tailed traffic. However, drift analysis of piecewise linear functions is sometimes a challenge by itself, due to the fact that the stochastic descent property is often lost at locations where the function is nondifferentiable. This difficulty can be handled by either smoothing the Lyapunov function, e.g., as in Down and Meyn [22], 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 Tsitsiklis [63]. We follow the second approach, and show how fluid approximations can significantly simplify drift analysis of this class of functions.

Theorem 2. *Consider the single-hop switched queueing network of Section 2 under the max-weight policy, and its natural FM, i.e., Equations (2)–(7), under the standing assumption that for some $\gamma \in (0,1)$, we have that $\mathbb{E}[A_f^{1+\gamma}(0)] < \infty$, for all $f \in \mathscr{F}$ (cf. Equation (1)). Consider a piecewise linear function $V \colon \mathbb{R}_+^F \to \mathbb{R}_+$ of the form*

$$V(x) = \max_{j \in \mathscr{J}} \left\{ \sum_{f \in \mathscr{F}} c_{jf} x_f \right\},$$

where $\mathscr{J} = \{1, \dots, J\}$ is the set of indices of the different pieces of the function, and where $c_{jf} \in \mathbb{R}$, for all $j \in \mathscr{J}$, $f \in \mathscr{F}$. Suppose that there exists $l > 0$ such that, for every initial condition $q(0)$ and regular time $t \geq 0$, the FMS satisfies $\dot{V}(q(t)) \leq -l$, whenever $V(q(t)) > 0$. Then, there exist $\alpha, \zeta > 0$ and $b_0 \in \mathbb{N}$ such that

$$\mathbb{E}[V(Q(t+b)) - V(Q(t)) + b\zeta; V(Q(t)) > \alpha b \mid \mathscr{F}_t] \leq 0, \quad \forall\, b \geq b_0.$$

This implies that the sequence $\{V(Q(t)); t \in \mathbb{Z}_+\}$ converges in distribution to the random variable $V(Q)$, where Q was defined in Section 2 as having the limiting distribution of $Q(t)$.

Moreover, if $c_{jf} > 0$, for some $j \in \mathscr{J}$ only when $f \in \mathscr{F}$ is an exponential-type traffic flow, then there exists $\theta > 0$ such that $\mathbb{E}[\exp(\theta V(Q))] < \infty$.

Proof. Fix $\gamma' \in (0, \gamma)$. For any $b \in \mathbb{N}$ consider the following set of sample paths of the arrival processes:

$$\tilde{H}_b(t) = \left\{ \omega \colon \max_{f \in \mathscr{F}} \max_{1 \leq \kappa \leq b} \frac{1}{b} \left| \sum_{\tau = t}^{t + \kappa - 1} (A_f(\tau) - \lambda_f) \right| < b^{-\gamma'/(1+\gamma)} \right\},$$

and let $\tilde{H}_b^c(t)$ be its complement. For any $b \in \mathbb{N}$ and $\alpha > 0$, we can write

$$\frac{1}{b} \mathbb{E}[V(Q(t+b)) - V(Q(t)); V(Q(t)) > \alpha b \mid \mathscr{F}_t]$$

$$= \frac{1}{b} \mathbb{E}[V(Q(t+b)) - V(Q(t)); V(Q(t)) > \alpha b, \tilde{H}_b(t) \mid \mathscr{F}_t] + \frac{1}{b} \mathbb{E}[V(Q(t+b)) - V(Q(t)); V(Q(t)) > \alpha b, \tilde{H}_b^c(t) \mid \mathscr{F}_t]$$

$$= \mathbb{E}\left[V\left(\frac{Q(t+b)}{b} \right) - V\left(\frac{Q(t)}{b} \right); V\left(\frac{Q(t)}{b} \right) > \alpha, \tilde{H}_b(t) \,\middle|\, \mathscr{F}_t \right] + \frac{1}{b} \mathbb{E}[V(Q(t+b)) - V(Q(t)); V(Q(t)) > \alpha b, \tilde{H}_b^c(t) \mid \mathscr{F}_t], \quad (17)$$

where the last equality follows from the fact that $V(\cdot)$ is homogeneous.

We begin by analyzing the first term on the right-hand side of Equation (17). We can write

$$\mathbb{E}[V(Q(t+b)/b) - V(Q(t)/b); V(Q(t)/b) > \alpha \mid \mathscr{F}_t]$$

$$= \mathbb{E}[V(Q(t+b)/b) - V(Q(t)/b); V(Q(t)/b) > \alpha \mid \tilde{H}_b(t), \mathscr{F}_t] \cdot \mathbb{P}(\tilde{H}_b(t) \mid \mathscr{F}_t)$$

$$= \mathbb{E}[V(Q(b)/b) - V(Q(0)/b); V(Q(0)/b) > \alpha \mid H_b, \mathscr{F}_0] \cdot \mathbb{P}(H_b \mid \mathscr{F}_0), \quad (18)$$

where

$$H_b = \tilde{H}_b(0) = \left\{ \omega: \max_{f \in \mathcal{F}} \max_{1 \leq \kappa \leq b} \frac{1}{b} \left| \sum_{\tau=0}^{\kappa-1} (A_f(\tau) - \lambda_f) \right| < b^{-\gamma'/(1+\gamma)} \right\}$$

is the set introduced in Lemma 2. The last equality follows from the fact that the arrival processes are mutually independent and IID over time slots, and the system is Markovian with respect to the vector of queue lengths.

Lemma 2 implies the existence of constants $\epsilon, b_0 > 0$ such that

$$\mathbb{P}(H_b \mid \mathcal{F}_0) \geq 1 - \epsilon, \quad \forall b \geq b_0. \tag{19}$$

Now consider the sequence of initial conditions $\{Q(0)b; b \in \mathbb{N}\}$, based on which we can construct a sequence of unscaled and scaled queue-length processes, $\{Q^b(\cdot); b \in \mathbb{N}\}$ and $\{\tilde{q}^b(\cdot); b \in \mathbb{N}\}$, respectively. Notice that $\tilde{q}^b(0) = Q(0)$, for all $b \in \mathbb{N}$. So, let $q(\cdot)$ be the queue-length part of the FMS from initial condition $q(0) = Q(0)$. Lemma 3 implies that it is unique, and Lemma 2 shows that, with high probability (for sample paths in H_b), the scaled queue-length process will be arbitrarily close to this FMS as long as the scaling parameter b is chosen sufficiently large. Combined with the fact that $V(\cdot)$ is continuous, it can be seen that there exists a function $g: \mathbb{N} \to \mathbb{R}_+$ that goes to zero as its argument goes to infinity, and that satisfies

$$V(Q(b)/b) - V(Q(0)/b) \leq V(q(1)) - V(q(0)) + g(b), \quad \forall \omega \in H_b, \forall b \geq b_0. \tag{20}$$

By assumption, there exists $l > 0$ such that, for every initial condition $q(0)$ and every regular time t, we have $\dot{V}(q(t)) \leq -l$, whenever $V(q(t)) > 0$. Moreover, almost every $t \in [0,1]$ is a regular time. Finally, if $V(q(0))$ is sufficiently large, then $V(q(t)) > 0$, for all $t \in [0,1]$. These imply that

$$V(q(1)) - V(q(0)) \leq -l, \tag{21}$$

for large enough $V(q(0))$.

Equations (18)–(21) imply that there exist $\alpha > 0$ (sufficiently large), $\delta > 0$, $b_0 \in \mathbb{N}$, and function $g(\cdot)$, such that

$$\mathbb{E}\left[V(Q(t+b)/b) - V(Q(t)/b) + l\delta - g(b); V(Q(t)/b) > \alpha, \tilde{H}_b(t) \mid \mathcal{F}_t \right] \leq 0, \tag{22}$$

for all $b \geq b_0$, and with $g(b) \to 0$ as $b \to \infty$.

Let us now analyze the second term on the right-hand side of Equation (17),

$$\frac{1}{b} \mathbb{E}[V(Q(t+b)) - V(Q(t)); V(Q(t)) > \alpha b, \tilde{H}_b^c(t) \mid \mathcal{F}_t].$$

Let $\bar{j} \in \mathcal{J}$ be a piece of $V(\cdot)$ that "dominates" at time slot $t+b$, i.e.,

$$\max_{j \in \mathcal{J}} \left\{ \sum_{f \in \mathcal{F}} c_{jf} Q_f(t+b) \right\} = \sum_{f \in \mathcal{F}} c_{\bar{j}f} Q_f(t+b).$$

We have that

$$V(Q(t+b)) - V(Q(t)) \leq \sum_{f \in \mathcal{F}} c_{\bar{j}f} \sum_{\tau=t}^{t+b-1} (A_f(\tau) - S_f(\tau) \cdot 1_{\{Q_f(\tau) > 0\}})$$

$$\leq \sum_{f \in \mathcal{F}} (-[c_{\bar{j}f}]^-)b + \sum_{f \in \mathcal{F}} [c_{\bar{j}f}]^+ \left(\sum_{\tau=t}^{t+b-1} A_f(\tau) \right).$$

So, there exists $c > 0$ such that

$$\mathbb{E}[V(Q(t+b)) - V(Q(t)); V(Q(t)) > \alpha b, \tilde{H}_b^c(t) \mid \mathcal{F}_t]$$

$$\leq \sum_{f \in \mathcal{F}} (-[c_{\bar{j}f}]^-)b \cdot \mathbb{P}(\tilde{H}_b^c(t) \mid \mathcal{F}_t) + \sum_{f \in \mathcal{F}} [c_{\bar{j}f}]^+ \mathbb{E}\left[\sum_{\tau=t}^{t+b-1} A_f(\tau); \tilde{H}_b^c(t) \middle| \mathcal{F}_t \right]$$

$$\leq c \sum_{f \in \mathcal{F}} (-[c_{\bar{j}f}]^-)b \cdot b^{-(\gamma-\gamma')/(1+\gamma)} + \sum_{f \in \mathcal{F}} [c_{\bar{j}f}]^+ \mathbb{E}\left[\sum_{\tau=t}^{t+b-1} A_f(\tau); \tilde{H}_b^c(t) \middle| \mathcal{F}_t \right]$$

$$= c \sum_{f \in \mathscr{F}} (-[c_{\bar{j}f}]^-)b^{(1+\gamma')/(1+\gamma)} + \sum_{f \in \mathscr{F}} [c_{\bar{j}f}]^+ \mathbb{E}\left[\sum_{\tau=t}^{t+b-1} A_f(\tau); \tilde{H}_b^c(t) \,\middle|\, \mathscr{F}_t\right], \tag{23}$$

where the last inequality follows from the proof of Lemma 2 whenever b is sufficiently large, c is a constant that does not depend on b, and $0 < \gamma' < \gamma$.

Now, for notational convenience, let

$$X_{k,b} \equiv \max_{1 \le \kappa \le b} \frac{1}{b}\left|\sum_{\tau=0}^{\kappa-1}(A_k(\tau) - \lambda_k)\right|.$$

The fact that the arrival processes are mutually independent, and IID over time with finite $(1+\gamma)$ moments, implies that

$$\mathbb{E}\left[\sum_{\tau=t}^{t+b-1} A_f(\tau); \tilde{H}_b^c(t) \mid \mathscr{F}_t\right] = \mathbb{E}\left[\sum_{\tau=0}^{b-1} A_f(\tau); \max_{k \in \mathscr{F}}\{X_{k,b}\} \ge b^{-\gamma'/(1+\gamma)}\right] \le \sum_{k \in \mathscr{F}} \mathbb{E}\left[\sum_{\tau=0}^{b-1} A_f(\tau); X_{k,b} \ge b^{-\gamma'/(1+\gamma)}\right]$$

$$= \mathbb{E}\left[\sum_{\tau=0}^{b-1} A_f(\tau); X_{f,b} \ge b^{-\gamma'/(1+\gamma)}\right] + \sum_{k \ne f} \lambda_k b \cdot \mathbb{P}(X_{k,b} \ge b^{-\gamma'/(1+\gamma)})$$

$$= \mathbb{E}\left[\sum_{\tau=0}^{b-1}(A_f(\tau) - \lambda_f); X_{f,b} \ge b^{-\gamma'/(1+\gamma)}\right] + \sum_{k \in \mathscr{F}} \lambda_k b \cdot \mathbb{P}(X_{k,b} \ge b^{-\gamma'/(1+\gamma)})$$

$$\le \mathbb{E}\left[\sum_{\tau=0}^{b-1}(A_f(\tau) - \lambda_f)\right] + \sum_{k \in \mathscr{F}} \lambda_k b \cdot \mathbb{P}(X_{k,b} \ge b^{-\gamma'/(1+\gamma)})$$

$$\le \sum_{k \in \mathscr{F}} \lambda_k b(c b^{-(\gamma-\gamma')/(1+\gamma)}), = c\left(\sum_{k \in \mathscr{F}} \lambda_k\right) b^{(1+\gamma')/(1+\gamma)}, \quad \forall f \in \mathscr{F}, \tag{24}$$

where the last inequality follows from the proof of Lemma 2 whenever b is sufficiently large.

Equations (23)–(24) imply the existence of $b_1, d > 0$ such that

$$\mathbb{E}[V(Q(t+b)) - V(Q(t)); V(Q(t)) > \alpha b, \tilde{H}_b^c(t) \mid \mathscr{F}_t] \le db^{(1+\gamma')/(1+\gamma)}, \quad \forall b \ge b_1,$$

which implies that

$$\frac{1}{b}\mathbb{E}[V(Q(t+b)) - V(Q(t)); V(Q(t)) > \alpha b, \tilde{H}_b^c(t) \mid \mathscr{F}_t] \le \frac{d}{b^{(\gamma-\gamma')/(1+\gamma)}}, \quad \forall b \ge b_1. \tag{25}$$

Finally, Equations (17), (22), and (25) imply that there exist $\alpha, \zeta > 0$ such that, for every $t \in \mathbb{Z}_+$ and for sufficiently large $b \in \mathbb{N}$,

$$\mathbb{E}[V(Q(t+b)) - V(Q(t)) + b\zeta; V(Q(t)) > \alpha b \mid \mathscr{F}_t] \le 0.$$

Notice that under our assumption on the coefficients c_{jf}, all queues that have a positive coefficient in any piece of $V(\cdot)$ have exponential-type arrivals. In particular, upward jumps of $V(\cdot)$ are also exponential type. Thus, Foster's criterion and Theorem 2.3 in Hajek [31] apply, and imply that the sequence $\{V(Q(t)); t \in \mathbb{Z}_+\}$ converges in distribution to the random variable $V(Q)$ and $V(Q)$ is exponential type. $\quad\square$

Remark 2. Theorem 2 provides a set of sufficient conditions for the existence of $\theta > 0$, such that $\mathbb{E}[\exp(\theta V(Q))] < \infty$. Depending on the structure of the piecewise linear function $V(\cdot)$, such a result may provide further information about the steady-state tail behavior of individual queues/flows. An example can be found in the following section, where drift analysis of the Lyapunov function in Equation (36), combined with Theorem 2, is used in order to prove that the steady-state queue lengths of certain flows in networks with disjoint schedules are exponential type.

3.3. Delay Stability Regions of a Single-Hop Network with Disjoint Schedules

In this section we consider a single-hop switched queueing network with disjoint schedules, so that each traffic flow belongs to exactly one schedule. Equivalently, the set of queues is partitioned into disjoint subsets, each subset being associated with one of the schedules. (Note that the example introduced in Figure 1 is a special case.) For this class of networks and under the max-weight policy, it turns out that the connections between fluid approximations and delay stability/instability established in Theorems 1 and 2 lead to a sharp characterization of the delay stability regions of certain queues in the presence of heavy-tailed traffic.

First, we describe the model and introduce some notation. We consider a system with $K+1$ schedules, which we denote by $\sigma_0, \sigma_1, \ldots, \sigma_K$. Schedule σ_k, $k = 0, \ldots, K$, includes F_k queues that we denote by (σ_k, f), $f = 1, \ldots, F_k$. Schedule σ_0 plays a special role, by being the one that includes a heavy-tailed flow. Again, since the system only carries single-hop traffic, we use the notions of queue and traffic flow interchangeably.

We denote the arrival rate to queue (σ_k, f) by $\lambda_f^{\sigma_k}$, which we assume to be strictly positive. We assume that within each schedule, the queues are indexed in descending order of arrival rates. We will focus on the generic case where the ordering is strict, as the analysis is more complicated otherwise. Thus, we assume throughout this section that

$$\lambda_{f+1}^{\sigma_k} < \lambda_f^{\sigma_k}, \quad f = 1, \ldots, F_k - 1, \quad k = 0, \ldots, K. \tag{26}$$

At each time slot at most one schedule can be activated. Whenever a schedule is activated then one packet is removed from all nonempty queues of that schedule.

We assume that the arriving traffic is in the stability region of the system, which is easily seen to be equivalent to the condition

$$\sum_{k=0}^{K} \lambda_1^{\sigma_k} < 1. \tag{27}$$

We assume that traffic flow (σ_0, f^*), $f^* \in \{1, \ldots, F_0\}$ is heavy tailed whereas every other traffic flow is exponential type. Theorem 2 in Markakis et al. [46] implies that under the max-weight policy, every traffic flow that does not belong to schedule σ_0 is delay unstable, for any positive arrival rate, because it conflicts with (σ_0, f^*). On the other hand, we expect traffic flows (σ_0, f), $f \neq f^*$, to have nontrivial delay stability regions, since they do not conflict with the heavy-tailed flow (σ_0, f^*); this is, indeed, the case for the three-queue system in Figure 1, where traffic flow 2 has a nontrivial delay stability region.

It turns out that the delay stability of queue (σ_0, f), $f \neq f^*$, is largely determined by the rate at which schedule σ_0 is served at a special configuration, where certain queues are empty and the others are nonempty. To make this more precise, we introduce some terminology and notation.

To every vector q of queues for the fluid model we associate a *configuration* $x \in \{0, 1\}^F$, where $F = F_0 + \cdots + F_K$ is the total number of flows. A typical component $x_f^{\sigma_k}$ is equal to 1 if schedule σ_k has maximum weight and queue (σ_k, f) is nonempty. When necessary, we will also use the notation $x(q)$ to indicate the dependence of x on q.

We now define the special configurations of interest. We define x^i, $i \in \{1, \ldots, F_0\}$, to be the configuration for which
(a) $x_f^{\sigma_0} = 1$, for $f = 1, \ldots, i$;
(b) $x_f^{\sigma_0} = 0$, for $f > i$;
(c) $x_1^{\sigma_k} = 1$, for $k = 1, \ldots, K$;
(d) $x_f^{\sigma_k} = 0$, for $k = 1, \ldots, K$ and $f > 1$.
That is, all schedules have maximum weight. For schedules $\sigma_1, \ldots, \sigma_K$, only the first queue is nonempty; for schedule σ_0, only queues $(\sigma_0, 1), \ldots, (\sigma_0, i)$ are nonempty.

For any configuration x, we define $\mu^{\sigma_k}(x)$ to be the service rate that schedule σ_k receives at any regular time at which the configuration is x. It is not hard to see from the structure of the fluid model that this rate only depends on the configuration and not on the exact value of the vector q.

We are now in a position to state our main result, which provides a tight characterization of the delay stability regions of those light-tailed flows (σ_0, f) of schedule σ_0 whose arrival rates are smaller than the arrival rate of the heavy-tailed flow, i.e., for $f > f^*$. This generalizes considerably the findings of Propositions 1 and 2 in Markakis et al. [46]. Parts (a) and (b) provide the stability conditions. Part (c) elaborates on the nature of these conditions. In particular, it shows that the delay stability regions of light-tailed flows in schedule σ_0 depend on the number of conflicting schedules and their highest arrival rates, $\lambda_1^{\sigma_k}$, but not on how many flows are included in each one of those schedules, or the lower arrival rates. The case of flows (σ_0, f), with $f < f^*$, is more complicated and will be discussed later.

Theorem 3. *Consider the single-hop switched queueing network with disjoint schedules described above under the max-weight policy, and arrival rates satisfying Equations (26) and (27). Fix some $j \in \{f^* + 1, \ldots, F_0\}$.*
(a) If flow (σ_0, j) is delay stable, then $\lambda_i^{\sigma_0} \leq \mu^{\sigma_0}(x^i)$, for all $i \in \{j, \ldots, F_0\}$, where

$$\mu^{\sigma_0}(x^i) = \frac{1}{Ki+1}\left(1 + K\sum_{f=1}^{i}\lambda_f^{\sigma_0} - \sum_{k=1}^{K}\lambda_1^{\sigma_k}\right), \quad i = 1, \ldots, F_0.$$

(b) *If* $\lambda_i^{\sigma_0} < \mu^{\sigma_0}(x^i)$, *for all* $i \in \{j, \ldots, F_0\}$, *then flow* (σ_0, j) *is delay stable and the steady-state length of the associated queue is exponential type.*

(c) *If flow* (σ_0, j) *is delay stable, then every flow* (σ_0, i) *with* $i > j$, *is also delay stable and the steady-state length of the associated queue is exponential type.*

Proof. We start by deriving the formula for $\mu^{\sigma_0}(x^i)$ in part (a) of the theorem. In fact, we will proceed more generally since later in the proof we will also need some information on $\mu^{\sigma_0}(x)$ for other configurations x.

Consider a general configuration x where schedule σ_0 has maximum weight, and let $\mathcal{K}(x)$ be the set of indices $k \geq 1$ for which schedule σ_k also has maximum weight. Clearly, if $\mathcal{K}(x) = \emptyset$ then $\mu^{\sigma_0}(x) = 1$. Otherwise, at any regular time, the max-weight policy splits the total available service rate, which is equal to one according to Equation (4), between the maximum weight schedules so that the weights of those schedules remain the same; this is a direct consequence of Equation (7). This means that for each schedule, the total inflow $\sum_f \lambda_f^{\sigma_k} x_f^{\sigma_k}$ minus the total outflow $\sum_f \mu^{\sigma_k}(x) x_f^{\sigma_k}$ into and out of, respectively, the nonempty queues must be the same for all $k \in \mathcal{K}(x)$. For any configuration for which schedule σ_0 has maximum weight, we have the following system of equations:

$$\sum_{f=1}^{F_0} \lambda_f^{\sigma_0} x_f^{\sigma_0} - |x^{\sigma_0}| \mu^{\sigma_0}(x) = \sum_{f=1}^{F_k} \lambda_f^{\sigma_k} x_f^{\sigma_k} - |x^{\sigma_k}| \mu^{\sigma_k}(x), \quad k \in \mathcal{K}(x), \tag{28}$$

$$\sum_{k \in \mathcal{K}(x) \cup \{0\}} \mu^{\sigma_k}(x) = 1, \tag{29}$$

$$\mu^{\sigma_k}(x) = 0, \quad k \notin \mathcal{K}(x) \cup \{0\}, \tag{30}$$

where $|x^{\sigma_k}| = \sum_{f=1}^{F_k} x_f^{\sigma_k}$, $k = 0, \ldots, K$. This is a system of $K + 1$ equations in the $K + 1$ unknowns $\mu^{\sigma_k}(x)$. It must necessarily have a unique solution, because otherwise we would have a contradiction to the existence and uniqueness of solutions to the FM.

For the special case of configuration x^i, we have $\mathcal{K}(x^i) = \{1, \ldots, K\}$. Furthermore, several of the $x_i^{\sigma_k}$ are equal to zero. By summing both sides of Equation (28) over all $k \in \mathcal{K}(x^i)$, keeping only the nonzero terms, and using Equation (29) to simplify the right-hand side, we obtain

$$K \sum_{f=1}^{i} \lambda_f^{\sigma_0} - Ki\mu^{\sigma_0}(x^i) = \sum_{k=1}^{K} \lambda_1^{\sigma_k} - (1 - \mu^{\sigma_0}(x^i)). \tag{31}$$

By collecting the terms involving $\mu^{\sigma_0}(x^i)$, we have that

$$(Ki + 1)\mu^{\sigma_0}(x^i) = \left(1 + K \sum_{f=1}^{i} \lambda_f^{\sigma_0} - \sum_{k=1}^{K} \lambda_1^{\sigma_k}\right), \tag{32}$$

which is equivalent to the expression in part (a) of the theorem.

We now continue with the remainder of the proof of part (a). We assume that flow (σ_0, j), with $j \in \{f^* + 1, \ldots, F_0\}$, is delay stable. We look into the evolution of the (unique) queue-length part $q(\cdot)$ of an FMS starting with the initial condition $q_{(0, f^*)} = 1$ and $q_{(\sigma_k, f)} = 0$ for all other flows. According to Theorem 1, and since (σ_0, j) is delay stable, we must have $q_{(0, j)}(t) = 0$, for all times. Furthermore, every flow (σ_0, i) with $i > j$ has a smaller arrival rate than flow (σ_0, j), whereas it gets served at the same rate (the rate at which schedule σ_0 is served). This implies that $q_{(0, i)}(t)$ is also zero for such flows. We conclude that the configurations satisfy $x_i^{\sigma_0}(q(t)) = 0$, for all $t \geq 0$ and $i \geq j$.

Starting with the initial condition that we have specified, in the beginning, schedule σ_0 is the only one that gets served. The weight of that schedule decreases, whereas the weights of the other schedules increase. At some point, the weight of some other schedule σ_k, with $k \neq 0$, becomes equal to that of schedule σ_0. Following that time, the weights of these two schedules decrease at the same rate, while the weights of the remaining schedules increase. (The fact that the weights of the maximum-weight schedules keep decreasing is a consequence of the assumption that we are operating within the stability region.) By repeating the same argument, there will be a time at which all schedules have the same weights. After that time, the overall processing rate is shared between the different schedules so that their weights remain equal at all times, and until they all simultaneously reach zero. However, note that all queues of a schedule σ_k receive service at the same rate, but the arrival rates $\lambda_f^{\sigma_k}$, for $f > 1$, are less than the arrival rate $\lambda_1^{\sigma_k}$ of flow $(\sigma_k, 1)$. For this reason, for $k \neq 0$ and $f > 1$, the queue (σ_k, f) will empty before the queue $(\sigma_k, 1)$ empties. The same reasoning holds for schedule σ_0 as well, and establishes that for $i > f^*$, queue (σ_0, i) will empty before queue (σ_0, f^*). We conclude that at some point we will reach a configuration x at which

   (i) $x_1^{\sigma_k} = 1$, for $k = 1, \ldots, K$;

   (ii) $x_f^{\sigma_k} = 0$, for $k = 1, \ldots, K$ and $f > 1$;

   (iii) $x_i^{\sigma_0} = 0$, for $i \geq j$.

The latter property holds because we assumed that $j > f^*$.

As we argue above, the assumption that queue (σ_0, j) is delay stable implies that the length every queue (σ_0, i), with $i \geq j$, remains at zero. Therefore, none of these queues builds up when configuration x is reached, which implies that the arrival rates to these queues are less than or equal to the rate at which schedule σ_0 is served, i.e.,

$$\lambda_i^{\sigma_0} \leq \mu^{\sigma_0}(x), \quad i = j, \ldots, F_0. \tag{33}$$

The proof of part (a) is completed once we establish the following result.

Lemma 4. *If a configuration x with the above properties* (i)–(iii) *satisfies Equation* (33), *then*

$$\lambda_i^{\sigma_0} \leq \mu^{\sigma_0}(x^i), \quad i = j, \ldots, F_0.$$

Proof of Lemma 4. Fix arbitrary $i \in \{j, \ldots, F_0\}$. By repeating the derivation of Equation (32) for such i, we have that

$$(K|x^{\sigma_0}| + 1)\mu^{\sigma_0}(x) = 1 + K\sum_{f=1}^{j-1} \lambda_f^{\sigma_0} x_f^{\sigma_0} - \sum_{k=1}^{K} \lambda_1^{\sigma_k} = 1 + K\sum_{f=1}^{i} \lambda_f^{\sigma_0} x_f^{\sigma_0} - \sum_{k=1}^{K} \lambda_1^{\sigma_k}, \tag{34}$$

where in the last equality we used property (iii) above. By Equation (33), $\lambda_i^{\sigma_0} \leq \mu^{\sigma_0}(x)$, which implies that

$$(K|x^{\sigma_0}| + 1)\lambda_i^{\sigma_0} \leq 1 + K\sum_{f=1}^{i} \lambda_f^{\sigma_0} x_f^{\sigma_0} - \sum_{k=1}^{K} \lambda_1^{\sigma_k}. \tag{35}$$

On the other hand, since queues are indexed in descending order of arrival rates,

$$K\lambda_i^{\sigma_0} \leq K\lambda_f^{\sigma_0}, \quad f = 1, \ldots, i.$$

By adding both sides of this inequality over all $f \in \{1, \ldots, i\}$ such that $x_f^{\sigma_0} = 0$ (there are $i - |x^{\sigma_0}|$ such f) and adding the result to Equation (35), we obtain

$$(Ki + 1)\lambda_i^{\sigma_0} \leq 1 + K\sum_{f=1}^{i} \lambda_f^{\sigma_0} - \sum_{k=1}^{K} \lambda_1^{\sigma_k}.$$

By comparing to Equation (32), we conclude that $\lambda_i^{\sigma_0} \leq \mu^{\sigma_0}(x^i)$. $\square$

We now show that if $\lambda_i^{\sigma_0} < \mu^{\sigma_0}(x^i)$, for $i = j, \ldots, F_0$, then every flow (σ_0, i) with $i \in \{j, \ldots, F_0\}$ is delay stable and the steady-state length of the associated queue is exponential type. This will prove part (b) directly, and in combination with the results above, part (c) as well. Our approach is based on Theorem 2. Consider the candidate Lyapunov function:

$$V(q) = \sum_{i=j}^{F_0} c_i q_i^{\sigma_0} + \max_{k=1,\ldots,K}\left\{\left[\sum_{f=1}^{F_k} q_f^{\sigma_k} - \sum_{f=1}^{F_0} q_f^{\sigma_0}\right]^+\right\}, \tag{36}$$

where $c_i \in (0, 1)$, for all $i \in \{j, \ldots, F_0\}$.

Let $q(0)$ be an arbitrary initial condition for the queue lengths in the FM. We will verify that if the c_i-parameters are properly chosen, then $\dot{V}(q(t))$ is uniformly negative whenever $V(q(t)) > 0$ and the derivative exists. Then, Theorem 2 will directly imply that every queue (σ_0, i), $i = j, \ldots, F_0$, is delay stable, since $V(\cdot)$ is continuous and piecewise linear. Moreover, the associated steady-state length of every such queue is exponential type because the variable $q_{f^*}^{\sigma_0}$ associated with the queue that receives heavy-tailed traffic appears in $V(q)$ with a negative coefficient.

In the analysis that follows, we distinguish between different cases, which correspond to different regions in the space of all possible vectors q.

(i) *Schedule σ_0 does not have maximum weight at time t.* In this case the candidate Lyapunov function reduces to

$$V(q(t)) = \sum_{f=1}^{F_{k^*}} q_f^{\sigma_{k^*}}(t) - \sum_{f=1}^{j-1} q_f^{\sigma_0}(t) - \sum_{i=j}^{F_0}(1-c_i)q_i^{\sigma_0}(t),$$

for some $k^* \in \{1, \ldots, K\}$. Since schedule σ_0 does not have maximum weight, at least one of the queues of schedule σ_{k^*} must be nonempty at time t. If k^* is the unique maximum weight schedule, then

$$\dot{V}(q(t)) = \sum_{f=1}^{F_{k^*}} (\lambda_f^{\sigma_{k^*}} - 1) \cdot 1_{\{q_f^{\sigma_{k^*}}(t)>0\}} - \sum_{f=1}^{j-1} \lambda_f^{\sigma_0} \cdot 1_{\{q_f^{\sigma_0}(t)>0\}} - \sum_{i=j}^{F_0}(1-c_i)\lambda_i^{\sigma_0} \cdot 1_{\{q_i^{\sigma_0}(t)>0\}}.$$

The right-hand side of the expression above is strictly negative. This is due to Equations (26)–(27) and our assumption that $c_i \in (0,1)$, for all $i \in \{j, \ldots, F_0\}$.

The same holds even if k^* is one of multiple schedules with maximum weight; this can be easily derived from the fact that the arriving traffic is in the stability region, and that max-weight drains the weights of all maximum weight schedules at the same rate.

(ii) *Schedule σ_0 has maximum weight at time t.* In this case the candidate Lyapunov function reduces to

$$V(q(t)) = \sum_{i=j}^{F_0} c_i q_i^{\sigma_0}(t).$$

Since $V(q(t)) > 0$, we have that at least one of the queues $(\sigma_0, j), \ldots, (\sigma_0, F_0)$ is nonempty.

We distinguish between two subcases: if schedule σ_0 is the unique maximum weight schedule at time t, then

$$\dot{V}(q(t)) = \sum_{i=j}^{F_0} c_i(\lambda_i^{\sigma_0} - 1) \cdot 1_{\{q_i^{\sigma_0}(t)>0\}} < 0.$$

On the other hand, if schedule σ_0 is one of multiple schedules with maximum weight at time t, then

$$\dot{V}(q(t)) = \sum_{i=j}^{F_0} c_i(\lambda_i^{\sigma_0} - \mu^{\sigma_0}(x(q(t)))) \cdot 1_{\{q_i^{\sigma_0}(t)>0\}}.$$

We now need to further distinguish between two subcases. The details of the argument are quite tedious and are relegated to Appendix A.

(a) *If all queues of schedule σ_0 are nonempty at time t*, then Lemmas 5–7 in Appendix A imply that

$$\mu^{\sigma_0}(x^{F_0}) \le \mu^{\sigma_0}(x(q(t))). \tag{37}$$

Thus, if we chose $c_i \ll c_{F_0}$, for all $i \in \{j, \ldots, F_0-1\}$, then $\dot{V}(q(t)) < 0$ because $\lambda_{F_0}^{\sigma_0} < \mu^{\sigma_0}(x^{F_0})$.

For more details on how we arrive at Equation (37), Lemmas 5 and 6 imply that $\mu^{\sigma_0}(x(q(t)))$ is greater than the service rate that schedule σ_0 receives under a configuration $\bar{x}$, where exactly the same queues are nonempty at σ_0 but only the highest rate queues of the competing schedules that have positive weight in $x(q(t))$ are nonempty. Lemma 5 covers the case where $x(q(t))$ has more than one nonempty queue in a competing schedule, while Lemma 6 covers the case where there is exactly one nonempty queue. Finally, by using iteratively Lemma 7 we establish that $\mu^{\sigma_0}(x^{F_0})$ is less than $\mu^{\sigma_0}(\bar{x})$, since in schedules $x(q(t))$ and $\bar{x}$ some competing schedules may have zero weight.

(b) *If all but one queues of schedule σ_0 are nonempty at time t*, then by arguing similarly to case (a), Lemmas 5–7 imply that

$$\mu^{\sigma_0}(x^{F_0-1}) \le \mu^{\sigma_0}(x(q(t))).$$

Thus, if we chose $c_i \ll c_{F_0-1}, c_{F_0}$, for all $i \in \{j, \ldots, F_0-2\}$, then $\dot{V}(q(t)) < 0$ because at least one of the queues (σ_0, F_0) and (σ_0, F_0-1) is nonempty and $\lambda_{F_0}^{\sigma_0} < \lambda_{F_0-1}^{\sigma_0} < \mu^{\sigma_0}(x^{F_0-1})$.

The other cases are treated similarly. □

Remark 3. The delay stability of traffic flow (σ_0, j), $j > f^*$, when one of the conditions in part (b) of Theorem 3 holds with equality may depend, in general, on higher order moments of the arrivals and not just the rates. To see this, suppose that a large batch of b packets arrives to the heavy-tailed queue (σ_0, f^*). A random walk-type argument can show that queue (σ_0, j) will build up to $\Omega(\sqrt{b})$ during an $\Omega(b)$ time interval, assuming that the configuration corresponding to the equality condition is reached. Thus, the aggregate length of this queue over a busy period will be $\Omega(b^{3/2})$, which implies that the delay stability of traffic flow (σ_0, j) may depend on the 1.5 moment of the arrivals to the heavy-tailed queue.

Theorem 3 provides a tight characterization of the delay stability regions of flows (σ_0, j), with $j > f^*$, but does not address flows (σ_0, j), with $j < f^*$. The delay stability analysis of those flows poses an additional challenge: it is not clear a priori whether the heavy-tailed queue (σ_0, f^*) is empty or not at the point in time where all schedules have maximum weight and only the highest rate queues of schedules σ_k, $k = 1, \ldots, K$, are nonempty. In the terminology of Theorem 3, it is not clear whether the "critical" configurations for delay stability are the configurations x^i or the configurations $\hat{x}^i$, where

(i) $\hat{x}_1^{\sigma_k} = 1$, $k = 1, \ldots, K$;
(ii) $\hat{x}_f^{\sigma_k} = 0$, $f > 1$, $k = 1, \ldots, K$ and ;
(iii) $\hat{x}_i^{\sigma_0} = 0$, $i \geq j$, $i \neq f^*$;
(iv) $\hat{x}_{f^*}^{\sigma_0} = 1$.

Theorem 3 shows that this distinction does not play a role in the delay stability of flows (σ_0, j), with $j > f^*$. This is not surprising because $\mu^{\sigma_0}(\hat{x}^i) = \mu^{\sigma_0}(x^i)$, for $i > f^*$. However, this distinction is expected to play a role in the delay stability regions of flows (σ_0, j), with $j < f^*$. More specifically, Lemma 5 implies that $\mu^{\sigma_0}(\hat{x}^i) < \mu^{\sigma_0}(x^i)$, for $i < f^*$. Thus, if queue (σ_0, f^*) is nonempty when the critical configuration is reached, the result would be a reduced stability region.

Nevertheless, the proof strategy of Theorem 3 can still be followed in order to derive necessary as well as sufficient rate conditions for delay stability, albeit not matching. In particular, through drift analysis of the piecewise linear Lyapunov function:

$$V(q) = \sum_{i \geq j, i \neq f^*} c_i q_i^{\sigma_0} + \max_{k=1,\ldots,K} \left\{ \left[\sum_{f=1}^{F_k} q_f^{\sigma_k} - \sum_{f=1}^{F_0} q_f^{\sigma_0} \right]^+ \right\},$$

combined with Lemmas 5–7, one can prove that if $\lambda_i^{\sigma_0} < \mu^{\sigma_0}(\hat{x}^i)$, for all $i = j, \ldots, F_0$, then flow (σ_0, j) is delay stable and the steady-state length of the associated queue is exponential type.

Conversely, if queue (σ_0, j), with $j < f^*$, is assumed delay stable then the same line of arguments as in the proof of part (a) of Theorem 3 (leading up to Lemma 4) can be followed in order to show that $\lambda_i^{\sigma_0} \leq \mu^{\sigma_0}(\hat{x}^i)$, for all $i > f^*$, are necessary conditions for delay stability. However, it is not clear whether the necessary conditions corresponding to flows i with $j \leq i < f^*$ should be of the form $\lambda_i^{\sigma_0} \leq \mu^{\sigma_0}(\hat{x}^i)$ or $\lambda_i^{\sigma_0} \leq \mu^{\sigma_0}(x^i)$; this depends on whether queue (σ_0, f^*) is empty or not when a critical configuration is reached, which is hard to determine a priori.

3.4. The Bottleneck Identification Algorithm

Theorem 1 provides a sufficient condition for the delay instability of traffic flows, based on the FMS from a specific initial condition. The following algorithmic procedure, which we term the *Bottleneck Identification* (BI) algorithm, tests this for all initial conditions of interest.

BI Algorithm. For every heavy-tailed traffic flow $h \in \mathcal{F}$,
  (i) solve the FM with initial condition 1 for queue h and 0 for all other queues;
  (ii) let U_h be the set of queues that become positive at any point before the FMS drains.

Let U be the set of queues that belong to U_h, for some heavy-tailed traffic flow h. Clearly, all queues/flows included in the set U produced by the algorithm are delay unstable.

The Bottleneck Identification algorithm is consistent with, and perhaps the natural extension of the "single big event/jump principle": the most likely way a queue may become delay unstable is through a single big event, i.e., a single big arrival to exactly one heavy-tailed queue. While this principle has been shown to hold in other (much simpler) single-server FCFS systems with heavy-tailed traffic, e.g., Pakes [51] and Veraverbeke [66], in general it does not hold in our setting, so the above algorithm may not always be "tight." In fact, one can construct simple examples where delay instability is caused by a combination of big events and, thus, would not be identified by the BI algorithm; Sharifnassab [58]. Of course, correspondingly, one can introduce modified versions of the algorithm where the fluid model is solved from more complex initial conditions, in order to account for combinations of big events; in the extreme case, all different combinations of big events. However, it is still not clear whether there exists a BI-type algorithm that provably identifies all delay unstable queues.

Nevertheless, either the basic or a modified version of the BI algorithm can be used to identify (some) delay unstable queues in a mechanical manner. This is particularly important in networks with complex topology, where any form of nonasymptotic analysis becomes quite challenging to apply. Below we present concrete examples that illustrate the use of the proposed algorithm.

3×3 *Switch.* Consider a 3×3 input-queued switch under the max-weight policy. This is a system of nine queues indexed by (i, j), where $i, j \in \{1, 2, 3\}$, with index i representing the input port and index j the output

Figure 2. (Color online) The FMS of a 3×3 switch, with initial condition one for queue $(1, 1)$ and zero for the other queues, and arrival rates $\lambda_{11} = 0.1$, $\lambda_{12} = 0.1$, $\lambda_{13} = 0.1$, $\lambda_{21} = 0.1$, $\lambda_{22} = 0.38$, $\lambda_{23} = 0.4$, $\lambda_{31} = 0.1$, $\lambda_{32} = 0.42$, $\lambda_{33} = 0.44$.
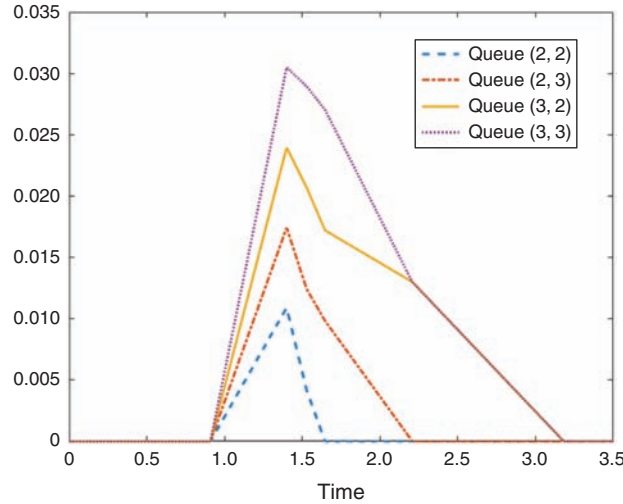
port of the switch. A schedule is a matching between input and output ports, so that the set of all schedules is as follows:

$$\mathcal{S} = \{\{(1, 1), (2, 2), (3, 3)\}, \{(1, 1), (2, 3), (3, 2)\}, \{(1, 2), (2, 1), (3, 3)\},$$
$$\{(1, 2), (2, 3), (3, 1)\}, \{(1, 3), (2, 1), (3, 2)\}, \{(1, 3), (2, 2), (3, 1)\}\}.$$

The 3×3 input-queued switch is a network with nondisjoint schedules, so an explicit characterization of its delay stability regions is not available.

Consider the set of arrival rates $\lambda_{11} = 0.1$, $\lambda_{12} = 0.1$, $\lambda_{13} = 0.1$, $\lambda_{21} = 0.1$, $\lambda_{22} = 0.38$, $\lambda_{23} = 0.4$, $\lambda_{31} = 0.1$, $\lambda_{32} = 0.42$, and $\lambda_{33} = 0.44$. Note that this set of rates satisfies $\sum_i \lambda_{ij} < 1$ for all j, and $\sum_j \lambda_{ij} < 1$ for all i, so that the system is stable under the max-weight policy; see McKeown et al. [48].

We assume that traffic flow $(1, 1)$ is heavy tailed, while all other traffic flows are light tailed. We are interested in the delay stability of flows $(2, 2)$, $(2, 3)$, $(3, 2)$, $(3, 3)$; these are the flows that do not conflict with flow $(1, 1)$. Figure 2 shows the FMS for the considered set of rates, and with initial condition one for queue $(1, 1)$ and zero for all other queues (we present only the queues of interest). The lengths of all queues of interest become positive before the FMS drains, so according to Theorem 1 they are delay unstable.

3×3 *Grid Network.* Consider the 3×3 grid network depicted in Figure 3 under the max-weight policy. This system represents a wireless network with interference constraints. Queues are identified with (directed) links and are indexed by $i = 1, \ldots, 12$. As soon as a packet is transmitted through the respective link, it exits the system. We assume the two-hop interference model, i.e., if a wireless link is transmitting, all links in a two-hop distance must idle. This implies that the set of schedules is as follows:

$$\mathcal{S} = \{\{1, 11\}, \{1, 12\}, \{1, 10\}, \{2, 8\}, \{2, 11\}, \{2, 12\}, \{3, 5\},$$
$$\{3, 10\}, \{3, 12\}, \{4\}, \{5, 8\}, \{5, 11\}, \{6\}, \{7\}, \{8, 10\}, \{9\}\}.$$

Again, this is a network with nondisjoint schedules, so an explicit characterization of its delay stability regions is not available.

Consider the set of arrival rates $\lambda_1 = 0.01$, $\lambda_2 = 0.02$, $\lambda_3 = 0.03$, $\lambda_4 = 0.04$, $\lambda_5 = 0.05$, $\lambda_6 = 0.06$, $\lambda_7 = 0.07$, $\lambda_8 = 0.08$, $\lambda_9 = 0.09$, $\lambda_{10} = 0.1$, $\lambda_{11} = 0.11$, and $\lambda_{12} = 0.12$. It can be verified that this set of rates belongs to the stability region of the system.

We assume that traffic flow 1 is heavy tailed, while all other traffic flows are light tailed. We are interested in the delay stability of traffic flows 10, 11, and 12, since these flows do not conflict with flow 1. Figure 4 shows the FMS for the considered set of rates, and with initial condition one for queue 1 and zero for all other queues (we present only the queues of interest). The lengths of all queues of interest become positive before the FMS drains, so according to Theorem 1 they are delay unstable.

Figure 3. A 3×3 grid wireless network with two-hop interference constraints.
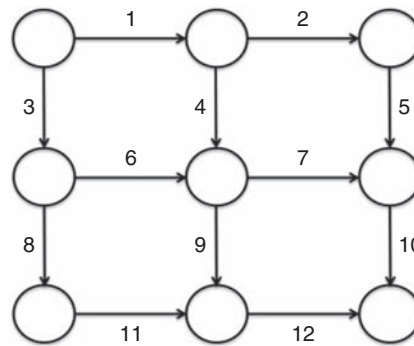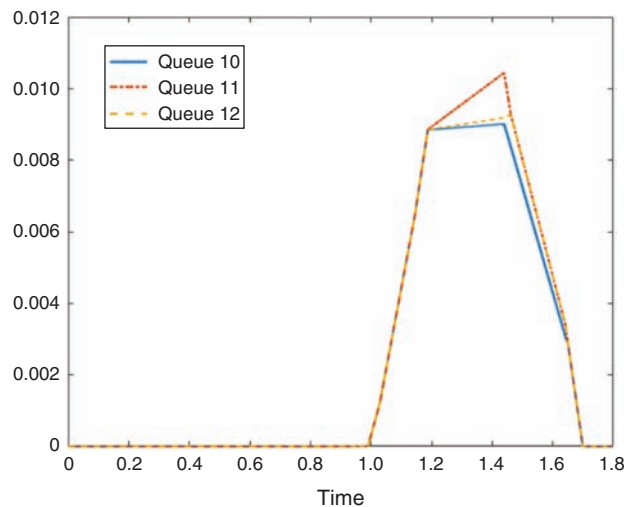

Figure 4. (Color online) The FMS of the 3×3 grid network of Figure 3, with initial condition one for queue 1 and zero for the other queues, and arrival rates $\lambda_1 = 0.01$, $\lambda_2 = 0.02$, $\lambda_3 = 0.03$, $\lambda_4 = 0.04$, $\lambda_5 = 0.05$, $\lambda_6 = 0.06$, $\lambda_7 = 0.07$, $\lambda_8 = 0.08$, $\lambda_9 = 0.09$, $\lambda_{10} = 0.1$, $\lambda_{11} = 0.11$, $\lambda_{12} = 0.12$.

4. Delay Analysis of the Back-Pressure Policy Under Heavy-Tailed Traffic via Fluid Approximations

Here, we show how several of the results and insights in Section 3 generalize naturally to a multihop setting, i.e., a multihop switched queueing network with a mix of heavy-tailed and exponential-type traffic under the back-pressure policy. Even though the extension is relatively straightforward from a mathematical standpoint, it is quite important in practice since most real-world networks are multihop. Moreover, the often complex topology of multihop networks makes a stochastic analysis even more challenging, which provides further motivation for our fluid approximations-based methodology. In the interest of space we provide a brief description of a multihop switched queueing network model, highlighting only the differences from the single-hop case. In particular, unless otherwise stated, most definitions, probabilistic assumptions, and notation remain unchanged. Moreover, several technical results that facilitate our delay analysis, e.g., stability of the network, existence of fluid limit, uniqueness of FMS, are relegated to Appendix B.

We emphasize that the queueing system presented and analyzed below does *not* include as a special case the single-hop network of Section 2, since it does not feature link-scheduling constraints. In other words, in contrast to the network of Section 2, all servers can be active simultaneously. It features though flow-scheduling constraints: any given server serves the traffic of potentially multiple flows, and needs to pick a single one at each time slot. While the extension to a network that incorporates both types of scheduling constraints is possible, we have opted to simplify the exposition and minimize the overlap with Section 2.

As mentioned in the introduction, the companion paper Markakis et al. [47] studies the same setting as this section. However, while the present section, and the paper in general, is methodologically oriented—to introduce analytical tools that facilitate a delay analysis under heavy-tailed traffic—the companion paper is more geared toward applications. In particular, it uses the analytical tools derived here (more specifically, Theorems 4

and 5) to study the impact of the network topology, the routing constraints, and the link capacities on the delay stability of back pressure.

The remainder of Section 4 is organized as follows. We begin by presenting a particular model of a multihop switched queueing network under the back-pressure policy. Then we introduce its natural fluid model, based on which we state the extensions of our earlier results for the single-hop case, accompanied by an example.

4.1. A Multihop Switched Queueing Network Under the Back-Pressure Policy

The topology of a multihop network is captured by a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{L})$, where $\mathcal{N}$ is the set of nodes and $\mathcal{L}$ is the set of directed links. Nodes represent the physical or virtual locations where traffic is buffered before transmission, and edges represent communication links, i.e., the means of transmission. With few exceptions, we use variables i and j to represent nodes, and (i, j) to denote a directed link from node i to node j.

Each traffic flow $f \in \mathcal{F}$ has a unique source node $s_f \in \mathcal{N}$ where it enters the network, and a unique destination node $d_f \in \mathcal{N}$ where it exits the network. Moreover, each traffic flow f has a predetermined set of links $\mathcal{L}_f \subset \mathcal{L}$ that it is allowed to access. We assume that $s_f \neq d_f$ and that there exists at least one directed path from s_f to d_f within the links in $\mathcal{L}_f$. If the set $\mathcal{L}_f$ includes exactly one path from the source to the destination, then we say that flow f has *fixed routing*. On the other hand, if there are multiple source-destination paths, we say that flow f has *dynamic routing*.

Node i belongs to set $\mathcal{N}_f$ if there exists a directed path from s_f to i that includes only links in $\mathcal{L}_f$. Thus, $\mathcal{N}_f \subset \mathcal{N}$ is the set of nodes that traffic flow f can access. Note that the source node s_f is trivially included in $\mathcal{N}_f$, while the destination node d_f is included in $\mathcal{N}_f$ due to our assumptions on $\mathcal{L}_f$.

The network operates in discrete time slots. Traffic flow f maintains a queue at every node $i \in \mathcal{N}_f$. We refer to this queue as queue (f, i) and denote its length at the beginning of time slot $t \in \mathbb{Z}_+$ by $Q_{f,i}(t)$. We emphasize that queue (f, i) buffers only packets of flow f. The service discipline within every queue is FCFS.

Traffic may arrive to queue (f, i) either exogenously if i is the source node s_f (in which case the arrivals are $A_f(t)$), or endogenously through a link in $\mathcal{L}_f$ whose destination node is i. We refer to queue (f, s_f) as the source queue of traffic flow f. We denote by $S_{f,i,j}(t)$ the number of packets that are scheduled for transmission from queue (f, i) through link $(i, j) \in \mathcal{L}_f$. These packets serve as (potential) departures from queue (f, i) and arrivals to queue (f, j), at time slot t.

We assume that *all links can transmit packets simultaneously*, and that all attempted transmissions are successful. Thus, our queueing model is suitable for several wireline applications (although not in the presence of "interference constraints" between links, as for example in switches).

Each link can only serve one traffic flow at any given time slot, giving rise to *flow-scheduling constraints*. The set of decisions regarding which flow is scheduled through each link can be interpreted as joint scheduling and routing. For simplicity, we assume that the capacity of all links is equal to one packet per time slot. We use the shorthand notation $Q(t)$ for the set of queue lengths $\{Q_{f,i}(t); i \in \mathcal{N}_f, f \in \mathcal{F}\}$, and $S(t)$ for the set of scheduling/routing decisions $\{S_{f,i,j}(t); (i, j) \in \mathcal{L}_f, f \in \mathcal{F}\}$. Moreover, we let $D(k) = \{D_{f,i}(t); i \in \mathcal{N}_f, f \in \mathcal{F}\}$, $k \in \mathbb{N}$, be the delays of the kth packet in the various queues of the network. We reserve the notation $D_f(k)$ for the end-to-end delay of the kth packet of flow f.

Similarly to the single-hop case, stability of the multihop network is defined as convergence in distribution of $Q(t)$ and $D(k)$.

In the given context, a queue length-based policy is a sequence of mappings from the history of queue lengths $\{Q(\tau); \tau = 0, \ldots, t\}$ to scheduling decisions $S(t), t \in \mathbb{Z}_+$. Moreover, a scheduling vector $S(t)$ is *feasible* if

    (i) $S_{f,i,j}(t) \in \{0, 1\}$, for all $(i, j) \in \mathcal{L}_f$, $f \in \mathcal{F}$;

    (ii) $\sum_{f \in \mathcal{F}} S_{f,i,j}(t) \leq 1$, for all $(i, j) \in \mathcal{L}$;

    (iii) $\sum_{j:(i,j) \in \mathcal{L}_f} S_{f,i,j}(t) \leq Q_{f,i}(t)$, for all $i \in \mathcal{N}_f$, $f \in \mathcal{F}$.

We focus on a particular stationary and Markovian queue length-based policy, the *back-pressure policy*: at each time slot t, $S(t)$ is a feasible scheduling vector that maximizes the aggregate back-pressure in the network, i.e.,

$$S(t) \in \arg\max \sum_{f \in \mathcal{F}} \sum_{(i,j) \in \mathcal{L}_f} (Q_{f,i}(t) - Q_{f,j}(t)) S_{f,i,j}(t).$$

If the solution is not unique, then each of the maximizing scheduling vectors is chosen with equal probability.

We note that the above description, which is referred to as max-pressure in Dai and Lin [20], is slightly different from the original, and most studied version of back pressure in Tassiulas and Ephremides [62]. The original policy is a greedy one, in the sense that it maximizes the "back pressure" on individual links, one at a time. It is not hard to see that on certain occasions, namely, when queues have few packets to transmit but many outgoing links, the original back-pressure policy may result in different scheduling decisions compared

to the version adopted here. However, in the regime of large queue lengths/delays that we are interested in this paper, the two policies are indistinguishable.

The *dynamics* of the multihop switched queueing network can be written in the following form:

$$Q_{f,s_f}(t+1) = Q_{f,s_f}(t) - \sum_{j:(s_f,j)\in\mathscr{L}_f} S_{f,s_f,j}(t) + A_f(t), \tag{38}$$

and

$$Q_{f,i}(t+1) = Q_{f,i}(t) - \sum_{j:(i,j)\in\mathscr{L}_f} S_{f,i,j}(t) + \sum_{j:(j,i)\in\mathscr{L}_f} S_{f,j,i}(t), \quad f \in \mathscr{N}_f \backslash \{s_f, d_f\}. \tag{39}$$

Finally, by convention,

$$Q_{f,d_f}(t) = 0, \quad \forall f \in \mathscr{F}. \tag{40}$$

The initial queue lengths are arbitrary nonnegative integers

Finally, the amount of traffic that can be stably supported by the network is, again, captured by the notion of stability region.

Definition 5 (Stability Region of Multihop Networks). An arrival rate vector $\lambda = (\lambda_1, \ldots, \lambda_F)$ is in the *stability region* of the multihop switched queueing network described above if there exist $\zeta_{f,i,j} \geq 0$, $f \in \mathscr{F}$, $i, j \in \mathscr{N}$, such that the following set of constraints is satisfied:
(i) Flow efficiency constraints: $\zeta_{f,i,i} = \zeta_{f,i,s_f} = \zeta_{f,d_f,i} = 0$, $\forall i \in \mathscr{N}$, $\forall f \in \mathscr{F}$.
(ii) Routing constraints: $\zeta_{f,i,j} = 0$, $\forall (i,j) \notin \mathscr{L}_f$, $\forall f \in \mathscr{F}$.
(iii) Flow conservation constraints: $\sum_{j\in\mathscr{N}} \zeta_{f,j,i} + \lambda_f \cdot 1_{\{i=s_f\}} = \sum_{j\in\mathscr{N}} \zeta_{f,i,j}$, $\forall i \neq d_f$, $\forall f \in \mathscr{F}$.
(iv) Link capacity constraints: $\sum_{f\in\mathscr{F}} \zeta_{f,i,j} < 1$, $\forall (i,j) \in \mathscr{L}$.

If an arrival rate vector is in the stability region, then there exists a policy that stabilizes the network, in the sense that the sequences of queue lengths and file delays converge in distribution. This can be shown by arguing similarly to Corollary 3.9 in Georgiadis et al. [29], and by utilizing our assumptions on the arrival processes, i.e., independence and finiteness of $(1 + \gamma)$ moments.

Finally, Lemma 8 in Appendix B proves that the multihop switched queueing network is stable under the back-pressure policy for any arrival rate vector in the stability region. This is done by showing that the above queueing system is a special case of a stochastic processing network under the maximum pressure policy, and then using results in Dai and Lin [20].

We denote by $Q_{f,i}$ and $D_{f,i}$ the steady-state length and delay of queue (f,i), respectively, and by D_f the end-to-end delay of traffic flow f in steady state. In a multihop network setting delay stability could be referring to a queue (f,i), depending on whether $\mathbb{E}[D_{f,i}]$ is finite or not, or to a traffic flow f, depending on whether $\mathbb{E}[D_f]$ is finite or not. Note that the source queue of a heavy-tailed flow is delay unstable under the FCFS discipline, which implies that every heavy-tailed flow is delay unstable irrespective of the policy applied.

4.2. Fluid Approximation of the Network

The FM of the multihop switched queueing network under the back-pressure policy is defined by the following set of equations, for every regular time $t \geq 0$:

$$\dot{q}_{f,i}(t) = - \sum_{j:(i,j)\in\mathscr{L}_f} \dot{s}_{f,i,j}(t) + \sum_{j:(j,i)\in\mathscr{L}_f} \dot{s}_{f,j,i}(t) + \lambda_f \cdot 1_{\{i=s_f\}}, \tag{41}$$

$$q_{f,i}(t) \geq 0, \tag{42}$$

$$s_{f,i,j}(0) = 0 \quad \text{and} \quad \dot{s}_{f,i,j}(t) \geq 0, \tag{43}$$

$$\sum_{f:(i,j)\in\mathscr{L}_f} \dot{s}_{f,i,j}(t) \leq 1, \tag{44}$$

$$\exists f': q_{f',i}(t) - q_{f',j}(t) > 0 \implies \sum_{f:(i,j)\in\mathscr{L}_f} \dot{s}_{f,i,j}(t) = 1, \tag{45}$$

$$q_{f',i}(t) - q_{f',j}(t) < \max_{f:(i,j)\in\mathscr{L}_f} \{[q_{f,i}(t) - q_{f,j}(t)]^+\} \implies \dot{s}_{f',i,j}(t) = 0. \tag{46}$$

In the equations above, we assume that $i \neq d_f$. Also, $q_{f,i}(t)$ represents the length of queue (f,i) at time t, while $s_{f,i,j}(t)$ represents the amount of time that link $(i,j) \in \mathscr{L}_f$ has been serving queue (f,i) up to time t. Henceforth, we use the shorthand notation $q(t)$ for the set of queue lengths $\{q_{f,i}(t); i \in \mathscr{N}_f, f \in \mathscr{F}\}$, and $s(t)$ for the set of cumulative scheduling decisions $\{s_{f,i,j}(t); (i,j) \in \mathscr{L}_f, f \in \mathscr{F}\}$.

Equations (41)–(46) are translated from the fluid model of stochastic processing networks in Dai and Lin [20]. More specifically, Equation (41) follows from Equation (14) of Dai and Lin [20], Equation (42) follows from Equation (15) of Dai and Lin [20], Equation (43) follows from Equation (18) of Dai and Lin [20], and Equation (44) follows from Equation (17) of Dai and Lin [20]. Finally, Equation (20) of Dai and Lin [20], at any regular time t, translates into

$$\dot{s}(t) \in \arg\max \sum_{f \in \mathcal{F}} \sum_{(i,j) \in \mathcal{L}_f} (q_{f,i}(t) - q_{f,j}(t)) \dot{s}_{f,i,j}(t). \tag{47}$$

Consider any service rate allocation $s(t)$ that satisfies Equations (41)–(44). It can be verified that if this allocation satisfies Equations (45)–(46), then it satisfies also the above inclusion, and vice versa.

Our convention regarding zero queue lengths at destination nodes provides a final equation for the description of the FM:

$$q_{f,d_f}(t) = 0. \tag{48}$$

Note that unlike the FM in Section 2.1, the lack of link-scheduling constraints in this model implies that we do not need to keep track of the idleness at each queue.

Fix arbitrary $T > 0$. An FMS from initial condition $q(0) = q$ is a Lipschitz continuous function $x(\cdot) = (q(\cdot), s(\cdot))$ that satisfies (i) $x(0) = (q, 0)$; (ii) Equations (41)–(48) over the subset of $[0, T]$ where $q(\cdot)$ is differentiable. An FMS is differentiable almost everywhere since it is Lipschitz continuous by assumption.

Exactly as in Section 2.1, we introduce a sequence of initial queue lengths $\{Q^b(0); b \in \mathbb{N}\}$; a corresponding sequence of queue-length process $\{Q^b(\cdot); b \in \mathbb{N}\}$; a "fluid-scaled" queue-length process

$$\tilde{q}^b(t) = \frac{Q^b(bt)}{b}, \quad t \in [0, T], \, b \in \mathbb{N}$$

(cf. Equation (8)); a vector q that satisfies

$$\max_{f \in \mathcal{F}} \max_{i \in \mathcal{N}_f} |\tilde{q}^b_{f,i}(0) - q_{f,i}| \leq \epsilon_b, \quad \forall b \in \mathbb{N},$$

where $\epsilon_b \to 0$ (cf. Equation (9)); and the sets H_b of "well-behaved" sample paths, defined in Equation (10).

Lemmas 9 and 10 in Appendix B prove the existence of a fluid limit and the uniqueness of the FMS, respectively.

4.3. Delay Stability Analysis via Fluid Approximations

Fluid approximations of switched queueing networks under the back-pressure policy have been employed in previous studies in order to show stability, e.g., in Dai and Lin [20]. In this section we show how fluid approximations can be used to prove delay stability/instability results in the presence of heavy-tailed traffic, by extending the scope of Theorems 1 and 2 to the multihop setting. A closer look at the proofs of these theorems reveals that they rely mainly on the existence of a fluid limit and the uniqueness of the fluid model solution, without making any further use of the single-hop nature of the network. Thus, having established these properties for the multihop network under consideration (Lemmas 9 and 10 in Appendix B), the extension of Theorems 1 and 2 becomes trivial and their proofs are omitted.

Theorem 4. *Consider the multihop switched queueing network described above under the back-pressure policy, and its natural FM, i.e., Equations (41)–(48). Let $h \in \mathcal{F}$ be a heavy-tailed traffic flow, and $q^*(\cdot)$ be the (unique) queue-length part of an FMS from initial condition $q^*_{h,s_h}(0) = 1$ and zero for every other queue. If there exists $\tau \in [0, T]$ such that $q^*_{f,i}(\tau) > 0$, then queue (f, i) is delay unstable.*

Theorem 5. *Consider the multihop switched queueing network described above under the back-pressure policy, and its natural FM, i.e., Equations (41)–(48). Then, Theorem 2 applies verbatim.*

Similarly to the single-hop case, Theorem 4 can be used in a bottleneck identification algorithm, to systematically test for delay instability in a multihop network. More specifically, for every heavy-tailed traffic flow $h \in \mathcal{F}$, we solve the FM with initial condition one for queue (h, s_h) and zero for all other queues, and find the set, U_h, of queues that become positive at any point before the FMS drains. Then, any queue that belongs to some U_h is delay unstable.

To illustrate the use of the above analytical tools, we borrow an example from the companion paper Markakis et al. [47]: consider the multihop network of Figure 5, which includes the heavy-tailed flow 1 and the exponential-type flows 2 and 3. The source of flow 1 is node 2 and the source of flows 2 and 3 is node 1.

Figure 5. An example from Markakis et al. [47]: the heavy-tailed flow 1 enters the network at node 2 and exits at node 3. The exponential-type flow 2 enters the network at node 1 and exits at node 3. The exponential-type flow 3 enters the network at node 1 and exits at node 4. Under the back-pressure policy, traffic flow 2 is delay unstable for all nonzero arrival rates whereas flow 3 has a nontrivial delay stability region.
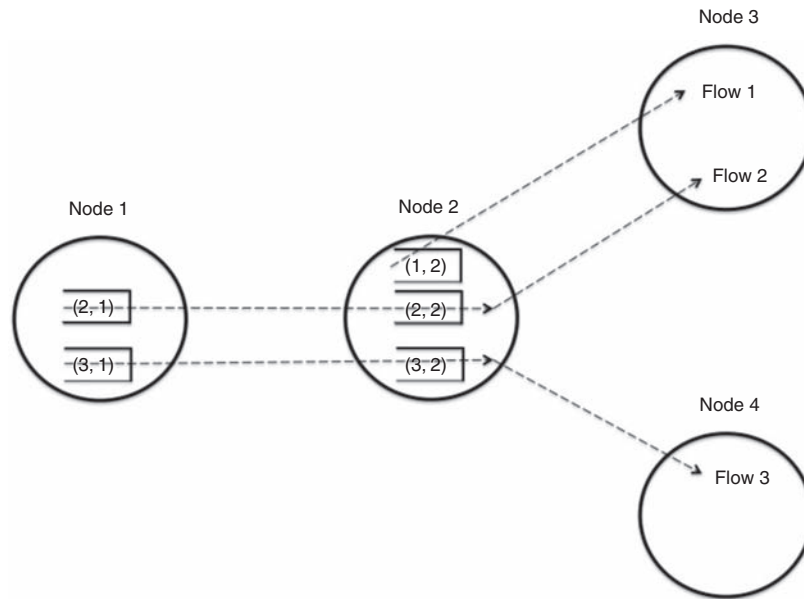
The destination of flows 1 and 2 is node 3 and the destination of flow 3 is node 4. All links can transmit simultaneously at unit rate and the back-pressure policy is applied.

It is not hard to see that traffic flow 2 is delay unstable because it competes for link $(2,3)$ with the source queue of the heavy-tailed flow 1. The more interesting question concerns flow 3, which serves as cross-traffic to flow 2, and that turns out to have a nontrivial delay stability region as the results of Markakis et al. [47] prove. More specifically, if $\lambda_3 < (2 + \lambda_1 - 2\lambda_2)/3$, then traffic flow 3 is delay stable and its aggregate queue length in steady state is exponential type. The proof of this result has two parts: initially, one shows that the function

$$H(q(t)) = V(q(t)) + G(q(t)),$$

where

$$V(q(t)) = \max\{[q_{3,1}(t) - q_{3,2}(t)]^+, [q_{2,1}(t) - q_{2,2}(t)]^+\}$$
$$= \max\{q_{3,1}(t), [q_{2,1}(t) - q_{2,2}(t)]^+\},$$

and

$$G(q(t)) = [q_{2,2}(t) - q_{1,2}(t)]^+,$$

is a Lyapunov function for the FM of the network; then, by applying Theorem 5 one proves directly the delay stability of flow 3; for details see Proposition 6 in Markakis et al. [47]. Conversely, if $\lambda_3 > (2 + \lambda_1 - 2\lambda_2)/3$, then traffic flow 3 is delay unstable. The proof of the latter result is based on a straightforward application of Theorem 4; for details see Proposition 5 in Markakis et al. [47].

Finally, we showcase the application of the BI algorithm on the particular network. Figure 6 shows the FMS for arrival rates $\lambda_1 = 0.2$, $\lambda_2 = 0.1$, and $\lambda_3 = 0.8$, from initial condition one for queue $(1,2)$ and zero for all other queues. The length of queue $(3,1)$ becomes positive before the FMS drains, so Theorem 4 implies that traffic flow 3 is delay unstable for the particular set of rates. We emphasize that we reached this conclusion in a mechanical manner, by solving numerically a set of "well-behaved" ODEs from a certain initial condition, without any need for analysis.

5. Discussion
This paper builds on and extends significantly the results of Markakis et al. [46]. More specifically, we study single-hop switched queueing networks with a mix of heavy-tailed and exponential-type traffic, and carry out a

Figure 6. (Color online) The FMS of the multihop network of Figure 5 from initial condition one for queue $(1, 2)$ and zero for the other queues, and arrival rates $\lambda_1 = 0.2$, $\lambda_2 = 0.1$, $\lambda_3 = 0.8$. (We have zoomed in the figure for clarity—the dashed line, which represents the length of queue $(1, 2)$, continues upward with a negative slope, and intersects the vertical axis at point $(0, 1)$.) Theorem 5 implies that the source queues of all three flows are delay unstable, which implies that the traffic flows themselves are delay unstable as well.
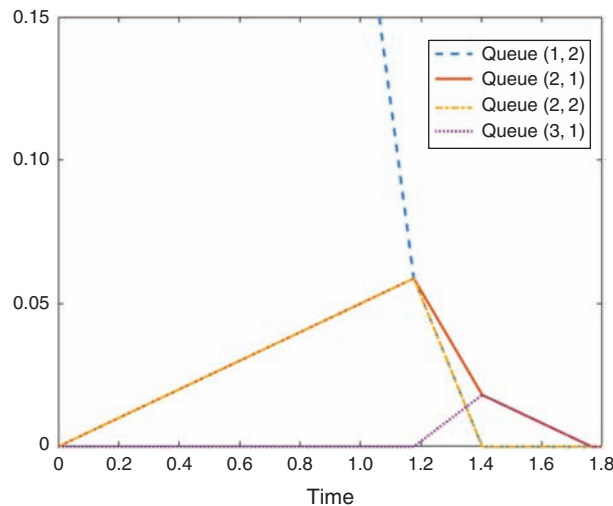
delay analysis of the max-weight policy. Our goal is to showcase the use of fluid approximations in proving both delay instability (using also renewal theory) and delay stability (combined with stochastic Lyapunov theory). Moreover, we apply these results to get a complete characterization of the delay stability regions of certain queues in networks with disjoint schedules.

We conclude the paper with some brief remarks. Theorems 1 and 2 are stated and proved in the context of a single-hop switched queueing network under the max-weight policy. However, the properties that we exploit in the respective proofs are only

  (i) the finiteness of the $(1 + \gamma)$ moment of every arrival process, for some $\gamma > 0$;
  (ii) the existence of a fluid limit;
  (iii) the uniqueness of the fluid model solution.

Thus, Theorems 1 and 2 can be easily extended to any Markovian queueing system for which properties (i)–(iii) hold. We follow exactly this approach in order to extend these results to multihop switched queueing networks under the back-pressure policy in Section 4.

5.1. Open Problems

The application of Theorem 2 rests on the availability of a suitable Lyapunov function for the fluid model. As the proof of Theorem 3 suggests, finding such a Lyapunov function is a nontrivial task and brings up some open problems:

  (a) If a certain queue in a switched queueing network is delay stable under the max-weight policy, does there exist a piecewise linear Lyapunov function (i.e., a function with the properties in Theorem 2) that can demonstrate delay stability?

  (b) Is there a polynomial time algorithm for constructing and certifying such a Lyapunov function, whenever one exists?

Note that if both of the above problems have affirmative answers, then we will have a polynomial time algorithm for deciding delay stability. The undecidability results in Gamarnik and Katz [26] suggest that such results do not hold for certain types of policies. On the other hand, as we are dealing with a different class of policies with special properties/structure, we cannot rule out that the answers will turn out to be positive, either for all switched queueing networks, or at least for some special cases.

A related open problem concerns the BI algorithm. Theorem 1 implies that the BI algorithm identifies *some* delay unstable queues. For the special case of single-hop networks with disjoint schedules, the proof of Theorem 3 essentially establishes that the BI algorithm identifies *all* delay unstable queues whose arrival rate is below the arrival rate of the heavy-tailed queue. However, as already discussed in Section 3.4, one can construct simple examples where delay instability is caused by combinations of big events. For such cases to be identified, one would need a modified BI algorithm that solves the fluid model from more complex initial conditions,

corresponding to the particular combinations of rare events. It is still not clear, though, whether there exists a BI-type algorithm that provably identifies all delay unstable queues in any given switched queueing network, or whether one has to take into account information about higher-order moments of the arrivals in order to determine delay stability.

Acknowledgments

Appendix A. Monotonicity Properties of Networks with Disjoint Schedules
In this appendix we consider the setting of Section 3.3, i.e., a single-hop switched queueing network with disjoint schedules under the max-weight policy. We analyze the FM of this system and prove certain monotonicity properties in the service rates of schedules, which are used in the proof of Theorem 3.

We use e_j to denote the vector whose jth element is equal to one, whereas all other elements are equal to zero. The dimension of this vector will be clear from the context.

Let $x = (x^{\sigma_0}, x^{\sigma_1}, \ldots, x^{\sigma_K})$ be a generic configuration. For reasons that become apparent in Section 3.3, we are interested in the service rate of schedule σ_0 under different configurations, so we assume that the vector x^{σ_0} is nonzero. Moreover, for concreteness, we assume that all vectors x^{σ_k}, $k = 1, \ldots, K$, are nonzero, i.e., all schedules have maximum weight and are drained simultaneously. The proofs of Lemmas 5–7 that follow can be easily modified to accommodate the case where some of the schedules $\sigma_1, \ldots, \sigma_K$ do not have maximum weight.

We also consider four modifications of x:

(i) Configuration $\bar{x}$ differs from x only in the fact that it includes an additional "lower rate" nonempty queue in a schedule different than σ_0. Without loss of generality, suppose that it is schedule σ_1. More precisely, $\bar{x} = (x^{\sigma_0}, \bar{x}^{\sigma_1}, \ldots, x^{\sigma_K})$, where $\bar{x}^{\sigma_1} = x^{\sigma_1} + e_j$, with $\arg\max_{f \in \{1, \ldots, F_1\}} \{x_f^{\sigma_1} > 0\} < j \le F_1$.

(ii) Configuration $\hat{x}$ differs from x only in the fact that it includes an additional "lower rate" nonempty queue in schedule σ_0. Specifically, $\hat{x} = (\hat{x}^{\sigma_0}, x^{\sigma_1}, \ldots, x^{\sigma_K})$, where $\hat{x}^{\sigma_0} = x^{\sigma_0} + e_{j'}$, with $\arg\max_{f \in \{1, \ldots, F_0\}} \{x_f^{\sigma_0} > 0\} < j' \le F_0$.

(iii) Configuration $\tilde{x}$ differs from x only in the fact that one of the nonempty queues of a schedule different than σ_0 has been replaced by another nonempty queue of the same schedule that has lower arrival rate. Without loss of generality, suppose that it is schedule σ_1, and that $x_j^{\sigma_1} = 1$, $x_{j+1}^{\sigma_1} = 0$, for some $j \in \{1, \ldots, F_1 - 1\}$. Then, $\tilde{x}_j^{\sigma_1} = 0$ and $\tilde{x}_{j+1}^{\sigma_1} = 1$, whereas $\tilde{x}_f^{\sigma_k} = x_f^{\sigma_k}$ for every other queue and schedule.

(iv) Configuration $\check{x}$ differs from x only in the fact that one of the nonempty queues of schedule σ_0 has been replaced by another nonempty queue of the same schedule that has lower arrival rate. Without loss of generality, suppose that $x_{j'}^{\sigma_0} = 1$, $x_{j'+1}^{\sigma_1} = 0$, for some $j' \in \{1, \ldots, F_0 - 1\}$. Then, $\check{x}_{j'}^{\sigma_0} = 0$ and $\check{x}_{j'+1}^{\sigma_0} = 1$, whereas $\check{x}_f^{\sigma_k} = x_f^{\sigma_k}$ for every other queue and schedule.

We denote by $\mu^{\sigma_k}(x)$, $\mu^{\sigma_k}(\bar{x})$, $\mu^{\sigma_k}(\hat{x})$, $\mu^{\sigma_k}(\tilde{x})$, and $\mu^{\sigma_k}(\check{x})$ the service rates of schedule σ_k in the FM and under configurations $x, \bar{x}, \hat{x}, \tilde{x}$, and $\check{x}$, respectively.

Lemma 5. *The service rates of schedule σ_0 under configurations x, $\bar{x}$, and $\hat{x}$ are ordered as follows:*

$$\mu^{\sigma_0}(\hat{x}) < \mu^{\sigma_0}(x) < \mu^{\sigma_0}(\bar{x}).$$

Proof. Equations (28)–(29) and some simple algebra imply that the service rate of schedule σ_0 under configuration x satisfies the following equality:

$$\mu^{\sigma_0}(x)\left(1 + \sum_{k=1}^{K} \frac{|x^{\sigma_0}|}{|x^{\sigma_k}|}\right) = 1 + \sum_{k=1}^{K} \frac{1}{|x^{\sigma_k}|}\left(\sum_{f=1}^{F_0} \lambda_f^{\sigma_0} x_f^{\sigma_0} - \sum_{f=1}^{F_k} \lambda_f^{\sigma_k} x_f^{\sigma_k}\right). \tag{A.1}$$

A similar derivation (omitted) shows that the service rates of schedule σ_0 under configurations $\bar{x}$ and $\hat{x}$ satisfy

$$\mu^{\sigma_0}(\bar{x})\left(1 + \frac{|x^{\sigma_0}|}{|x^{\sigma_1}| + 1} + \sum_{k=2}^{K} \frac{|x^{\sigma_0}|}{|x^{\sigma_k}|}\right) = 1 + \frac{1}{|x^{\sigma_1}| + 1}\left(\sum_{f=1}^{F_0} \lambda_f^{\sigma_0} x_f^{\sigma_0} - \sum_{f=1}^{F_1} \lambda_f^{\sigma_1} x_f^{\sigma_1} - \lambda_j^{\sigma_1}\right) + \sum_{k=2}^{K} \frac{1}{|x^{\sigma_k}|}\left(\sum_{f=1}^{F_0} \lambda_f^{\sigma_0} x_f^{\sigma_0} - \sum_{f=1}^{F_k} \lambda_f^{\sigma_k} x_f^{\sigma_k}\right); \tag{A.2}$$

and

$$\mu^{\sigma_0}(\hat{x})\left(1 + \sum_{k=1}^{K} \frac{|x^{\sigma_0}| + 1}{|x^{\sigma_k}|}\right) = 1 + \sum_{k=1}^{K} \frac{1}{|x^{\sigma_k}|}\left(\sum_{f=1}^{F_0} \lambda_f^{\sigma_0} x_f^{\sigma_0} - \sum_{f=1}^{F_k} \lambda_f^{\sigma_k} x_f^{\sigma_k} + \lambda_{j'}^{\sigma_0}\right), \tag{A.3}$$

respectively.

We first show that $\mu^{\sigma_0}(x) < \mu^{\sigma_0}(\bar{x})$. For notational convenience we define

$$A = 1 + \sum_{k=2}^{K} \frac{|x^{\sigma_0}|}{|x^{\sigma_k}|}, \quad B = 1 + \sum_{k=2}^{K} \frac{1}{|x^{\sigma_k}|}\left(\sum_{f=1}^{F_0} \lambda_f^{\sigma_0} x_f^{\sigma_0} - \sum_{f=1}^{F_k} \lambda_f^{\sigma_k} x_f^{\sigma_k}\right), \quad \text{and} \quad C = B|x^{\sigma_1}| + \sum_{f=1}^{F_0} \lambda_f^{\sigma_0} x_f^{\sigma_0} - \sum_{f=1}^{F_1} \lambda_f^{\sigma_1} x_f^{\sigma_1}.$$

Then, Equations (A.1)–(A.2) can be written as follows:

$$\mu^{\sigma_0}(x)(A|x^{\sigma_1}| + |x^{\sigma_0}|) = C \quad \text{and} \quad \mu^{\sigma_0}(\bar{x})(A|x^{\sigma_1}| + A + |x^{\sigma_0}|) = C + B - \lambda_j^{\sigma_1}.$$

The above implies that the inequality $\mu^{\sigma_0}(x) < \mu^{\sigma_0}(\bar{x})$ is equivalent to

$$C/(A|x^{\sigma_1}| + |x^{\sigma_0}|) < (C + B - \lambda_j^{\sigma_1})/(A|x^{\sigma_1}| + A + |x^{\sigma_0}|) \iff AB|x^{\sigma_1}| + B|x^{\sigma_0}| - A|x^{\sigma_1}|\lambda_j^{\sigma_1} - |x^{\sigma_0}|\lambda_j^{\sigma_1} - CA > 0$$

$$\iff B|x^{\sigma_0}| - A|x^{\sigma_1}|\lambda_j^{\sigma_1} - |x^{\sigma_0}|\lambda_j^{\sigma_1} - A\left(\sum_{f=1}^{F_0}\lambda_f^{\sigma_0}x_f^{\sigma_0} - \sum_{f=1}^{F_1}\lambda_f^{\sigma_1}x_f^{\sigma_1}\right) > 0.$$

Substituting the expressions for A and B, we get

$$|x^{\sigma_0}| - \left(\sum_{k=2}^{K}\frac{|x^{\sigma_0}|}{|x^{\sigma_k}|}\right)\sum_{f=1}^{F_k}\lambda_f^{\sigma_k}x_f^{\sigma_k} - \sum_{f=1}^{F_0}\lambda_f^{\sigma_0}x_f^{\sigma_0} + \sum_{f=1}^{F_1}\lambda_f^{\sigma_1}x_f^{\sigma_1} + \left(\sum_{k=2}^{K}\frac{|x^{\sigma_0}|}{|x^{\sigma_k}|}\right)\sum_{f=1}^{F_1}\lambda_f^{\sigma_1}x_f^{\sigma_1} - \lambda_j^{\sigma_1}\left(|x^{\sigma_0}| + |x^{\sigma_1}| + |x^{\sigma_1}|\sum_{k=2}^{K}\frac{|x^{\sigma_0}|}{|x^{\sigma_k}|}\right) > 0.$$

By taking into account Equation (26) and the fact that (σ_1, j) is a "lower rate" queue, in order to prove that $\mu^{\sigma_0}(x) < \mu^{\sigma_0}(\bar{x})$ it suffices to show that

$$|x^{\sigma_0}|\left(1 - \lambda_1^{\sigma_0} - \sum_{k=2}^{K}\lambda_1^{\sigma_k}\right) + |x^{\sigma_1}|\lambda_j^{\sigma_1} + |x^{\sigma_1}|\lambda_j^{\sigma_1}\left(\sum_{k=2}^{K}\frac{|x^{\sigma_0}|}{|x^{\sigma_k}|}\right) - \lambda_j^{\sigma_1}\left(|x^{\sigma_0}| + |x^{\sigma_1}| + |x^{\sigma_1}|\sum_{k=2}^{K}\frac{|x^{\sigma_0}|}{|x^{\sigma_k}|}\right) > 0,$$

or equivalently,

$$|x^{\sigma_0}|\left(1 - \lambda_1^{\sigma_0} - \lambda_j^{\sigma_1} - \sum_{k=2}^{K}\lambda_1^{\sigma_k}\right) > 0.$$

The latter is true because of Equation (27).

Now we show that $\mu^{\sigma_0}(\hat{x}) < \mu^{\sigma_0}(x)$. Again, for notational convenience we define

$$A' = 1 + \sum_{k=1}^{K}\frac{|x^{\sigma_0}|}{|x^{\sigma_k}|} \quad \text{and} \quad B' = 1 + \sum_{k=1}^{K}\frac{1}{|x^{\sigma_k}|}\left(\sum_{f=1}^{F_0}\lambda_f^{\sigma_0}x_f^{\sigma_0} - \sum_{f=1}^{F_k}\lambda_f^{\sigma_k}x_f^{\sigma_k}\right).$$

Then, Equations (A.1) and (A.3) can be written as follows:

$$\mu^{\sigma_0}(x)A' = B' \quad \text{and} \quad \mu^{\sigma_0}(\hat{x})\left(A' + \sum_{k=1}^{K}\frac{1}{|x^{\sigma_k}|}\right) = B' + \sum_{k=1}^{K}\frac{\lambda_{j'}^{\sigma_0}}{|x^{\sigma_k}|}.$$

The above implies that $\mu^{\sigma_0}(\hat{x}) < \mu^{\sigma_0}(x)$ is equivalent to

$$\sum_{k=1}^{K}\frac{1}{|x^{\sigma_k}|}\left(1 + \sum_{k=1}^{K}\frac{1}{|x^{\sigma_k}|}\left(\sum_{f=1}^{F_0}\lambda_f^{\sigma_0}x_f^{\sigma_0} - \sum_{f=1}^{F_k}\lambda_f^{\sigma_k}x_f^{\sigma_k}\right)\right) - \sum_{k=1}^{K}\frac{\lambda_{j'}^{\sigma_0}}{|x^{\sigma_k}|}\left(1 + \sum_{k=1}^{K}\frac{|x^{\sigma_0}|}{|x^{\sigma_k}|}\right) > 0,$$

which, in turn, is equivalent to

$$\sum_{k=1}^{K}\frac{1 - \lambda_{j'}^{\sigma_0}}{|x^{\sigma_k}|} + \sum_{k=1}^{K}\frac{1}{|x^{\sigma_k}|}\left(\sum_{k=1}^{K}\frac{1}{|x^{\sigma_k}|}\left(\sum_{f=1}^{F_0}\lambda_f^{\sigma_0}x_f^{\sigma_0} - \lambda_{j'}^{\sigma_0}|x^{\sigma_0}|\right)\right) - \sum_{k=1}^{K}\frac{1}{|x^{\sigma_k}|}\left(\sum_{k=1}^{K}\frac{1}{|x^{\sigma_k}|}\sum_{f=1}^{F_k}\lambda_f^{\sigma_k}x_f^{\sigma_k}\right) > 0.$$

By taking into account Equation (26), in order to show that $\mu^{\sigma_0}(\hat{x}) < \mu^{\sigma_0}(x)$ it suffices to show that

$$\sum_{k=1}^{K}\frac{1}{|x^{\sigma_k}|}\left(1 - \lambda_{j'}^{\sigma_0} - \sum_{k=1}^{K}\lambda_1^{\sigma_k}\right) + \sum_{k=1}^{K}\frac{1}{|x^{\sigma_k}|}\sum_{k=1}^{K}\frac{1}{|x^{\sigma_k}|}\left(\sum_{f=1}^{F_0}\lambda_f^{\sigma_0}x_f^{\sigma_0} - \lambda_{j'}^{\sigma_0}|x^{\sigma_0}|\right).$$

The latter is true because of Equation (27) and the fact that (σ_0, j') is a "lower rate" queue. $\square$

Lemma 6. *The service rates of schedule σ_0 under configurations x, $\tilde{x}$, and $\check{x}$ are ordered as follows:*

$$\mu^{\sigma_0}(\check{x}) < \mu^{\sigma_0}(x) < \mu^{\sigma_0}(\tilde{x}).$$

Proof. Equation (A.1) regarding the service rate of schedule σ_0 under configuration x, can be rewritten as follows:

$$\mu^{\sigma_0}(x)\left(1 + \sum_{k=1}^{K}\frac{|x^{\sigma_0}|}{|x^{\sigma_k}|}\right) = 1 + \sum_{k=2}^{K}\frac{1}{|x^{\sigma_k}|}\left(\sum_{f=1}^{F_0}\lambda_f^{\sigma_0}x_f^{\sigma_0} - \sum_{f=1}^{F_k}\lambda_f^{\sigma_k}x_f^{\sigma_k}\right) + \frac{1}{|x^{\sigma_1}|}\left(\sum_{f=1}^{F_0}\lambda_f^{\sigma_0}x_f^{\sigma_0} - \sum_{f=1}^{F_1}\lambda_f^{\sigma_1}x_f^{\sigma_1}\right)$$

$$= 1 + \sum_{k=1}^{K}\frac{1}{|x^{\sigma_k}|}\sum_{f=1}^{F_0}\lambda_f^{\sigma_0}x_f^{\sigma_0} - \sum_{k=1}^{K}\frac{1}{|x^{\sigma_k}|}\sum_{f=1}^{F_k}\lambda_f^{\sigma_k}x_f^{\sigma_k}.$$

By arguing similarly, and taking into account the fact that $|x^{\sigma_k}| = |\tilde{x}^{\sigma_k}| = |\check{x}^{\sigma_k}|$, for all $k \in \{0, \ldots, K\}$, we have that the service rates of schedule σ_0 under configurations $\tilde{x}$ and $\check{x}$ satisfy the following equalities:

$$\mu^{\sigma_0}(\tilde{x})\left(1 + \sum_{k=1}^{K} \frac{|x^{\sigma_0}|}{|x^{\sigma_k}|}\right) = 1 + \sum_{k=2}^{K} \frac{1}{|x^{\sigma_k}|}\left(\sum_{f=1}^{F_0} \lambda_f^{\sigma_0} x_f^{\sigma_0} - \sum_{f=1}^{F_k} \lambda_f^{\sigma_k} x_f^{\sigma_k}\right) + \frac{1}{|x^{\sigma_1}|}\left(\sum_{f=1}^{F_0} \lambda_f^{\sigma_0} x_f^{\sigma_0} - \sum_{f=1}^{F_1} \lambda_f^{\sigma_1} \tilde{x}_f^{\sigma_1}\right),$$

and

$$\mu^{\sigma_0}(\check{x})\left(1 + \sum_{k=1}^{K} \frac{|x^{\sigma_0}|}{|x^{\sigma_k}|}\right) = 1 + \sum_{k=1}^{K} \frac{1}{|x^{\sigma_k}|} \sum_{f=1}^{F_0} \lambda_f^{\sigma_0} \check{x}_f^{\sigma_0} - \sum_{k=1}^{K} \frac{1}{|x^{\sigma_k}|} \sum_{f=1}^{F_k} \lambda_f^{\sigma_k} x_f^{\sigma_k}.$$

Let us prove that $\mu^{\sigma_0}(x) < \mu^{\sigma_0}(\tilde{x})$. We use the notation

$$A = 1 + \sum_{k=1}^{K} \frac{|x^{\sigma_0}|}{|x^{\sigma_k}|} \quad \text{and} \quad B = 1 + \sum_{k=2}^{K} \frac{1}{|x^{\sigma_k}|}\left(\sum_{f=1}^{F_0} \lambda_f^{\sigma_0} x_f^{\sigma_0} - \sum_{f=1}^{F_k} \lambda_f^{\sigma_k} x_f^{\sigma_k}\right) + \frac{1}{|x^{\sigma_1}|} \sum_{f=1}^{F_0} \lambda_f^{\sigma_0} x_f^{\sigma_0}.$$

Then, the service rates of schedule σ_0 under configurations x and $\tilde{x}$, respectively, can be written as follows:

$$\mu^{\sigma_0}(x) = \frac{1}{A}\left(B - \frac{1}{|x^{\sigma_1}|} \sum_{f=1}^{F_1} \lambda_f^{\sigma_1} x_f^{\sigma_1}\right) \quad \text{and} \quad \mu^{\sigma_0}(\tilde{x}) = \frac{1}{A}\left(B - \frac{1}{|x^{\sigma_1}|} \sum_{f=1}^{F_1} \lambda_f^{\sigma_1} \tilde{x}_f^{\sigma_1}\right).$$

Since configuration $\tilde{x}$ differs from x only in the fact that one of the nonempty queues of schedule σ_1 has been substituted by another nonempty queue of the same schedule that has lower arrival rate, these equations imply directly that $\mu^{\sigma_0}(x) < \mu^{\sigma_0}(\tilde{x})$.

The fact that $\mu^{\sigma_0}(\tilde{x}) < \mu^{\sigma_0}(x)$ is proved similarly. □

Finally, we consider the configuration $y = (y^{\sigma_0}, e_1, e_1, \ldots, e_1)$, which corresponds to a situation where only the highest rate queue from each of the schedules $\sigma_1, \ldots, \sigma_K$ is nonempty. In contrast, we do not impose any restrictions on which queues are nonempty in schedule σ_0. Let us also consider a modification of this configuration, $\bar{y}$, which differs from y only in the fact that one of the highest rate queues of schedules $\sigma_1, \ldots, \sigma_K$ is empty. Without loss of generality, suppose that it is the highest rate queue of schedule σ_1. In mathematical terms, $\bar{y} = (y^{\sigma_0}, 0, e_1, \ldots, e_1)$.

We denote by $\mu^{\sigma_k}(y)$ and $\mu^{\sigma_k}(\bar{y})$ the service rates of schedule σ_k in the FM and under configurations y and $\bar{y}$, respectively.

Lemma 7. *The service rates of schedule σ_0 under configurations y and $\bar{y}$ are ordered in the following way*:

$$\mu^{\sigma_0}(y) < \mu^{\sigma_0}(\bar{y}).$$

Proof. Equations (28)–(29) imply that the service rates of the different schedules under configuration y satisfy

$$\sum_{f=1}^{F_0} \lambda_f^{\sigma_0} y_f^{\sigma_0} - |y^{\sigma_0}| \mu^{\sigma_0}(y) = \lambda_1^{\sigma_k} - \mu^{\sigma_k}(y), \quad k = 1, \ldots, K;$$

and

$$\sum_{k=0}^{K} \mu^{\sigma_k}(y) = 1.$$

The above equations and some simple algebra imply that

$$\mu^{\sigma_0}(y)(1 + K|y^{\sigma_0}|) = 1 + K\left(\sum_{f=1}^{F_0} \lambda_f^{\sigma_0} y_f^{\sigma_0}\right) - \sum_{k=1}^{K} \lambda_1^{\sigma_k}. \tag{A.4}$$

Now, under configuration $\bar{y}$, we have that $\mu^{\sigma_1}(\bar{y}) = 0$ while the rest of the service rates are split in the max-weight fashion, i.e.,

$$\sum_{f=1}^{F_0} \lambda_f^{\sigma_0} y_f^{\sigma_0} - |y^{\sigma_0}| \mu^{\sigma_0}(\bar{y}) = \lambda_1^{\sigma_k} - \mu^{\sigma_k}(\bar{y}), \quad k = 2, \ldots, K.$$

Then, the work-conserving nature of the policy and some simple algebra imply that

$$\mu^{\sigma_0}(\bar{y})(1 + (K-1)|y^{\sigma_0}|) = 1 + (K-1)\left(\sum_{f=1}^{F_0} \lambda_f^{\sigma_0} y_f^{\sigma_0}\right) - \sum_{k=2}^{K} \lambda_1^{\sigma_k}. \tag{A.5}$$

For notational convenience we define the quantities

$$A = 1 + K|y^{\sigma_0}| \quad \text{and} \quad B = 1 + K\left(\sum_{f=1}^{F_0} \lambda_f^{\sigma_0} y_f^{\sigma_0}\right) - \sum_{k=1}^{K} \lambda_1^{\sigma_k}.$$

Then, Equations (A.4) and (A.5) can be written as follows:

$$\mu^{\sigma_0}(y)A = B \quad \text{and} \quad \mu^{\sigma_0}(\bar{y})(A - |y^{\sigma_0}|) = B - \sum_{f=1}^{F_0} \lambda_f^{\sigma_0} y_f^{\sigma_0} + \lambda_1^{\sigma_1}.$$

The above implies that $\mu^{\sigma_0}(y) < \mu^{\sigma_0}(\bar{y})$ is equivalent to

$$B|y^{\sigma_0}| - A \sum_{f=1}^{F_0} \lambda_f^{\sigma_0} y_f^{\sigma_0} + A \lambda_1^{\sigma_1} > 0,$$

which, in turn, is equivalent to

$$|y^{\sigma_0}| \left(1 - \sum_{k=1}^{K} \lambda_1^{\sigma_k} - \frac{1}{|y^{\sigma_0}|} \sum_{f=1}^{F_0} \lambda_f^{\sigma_0} y_f^{\sigma_0} \right) + (1 + K|y^{\sigma_0}|)\lambda_1^{\sigma_1} > 0.$$

The latter is true because of Equation (27) and the fact that

$$\frac{1}{|y^{\sigma_0}|} \sum_{f=1}^{F_0} \lambda_f^{\sigma_0} y_f^{\sigma_0} \leq \max_{f \in \{1,\dots,F_0\}} \{\lambda_f^{\sigma_0}\}. \quad \square$$

Lemmas 5–7 can be easily modified to accommodate the case where not all schedules have maximum weight: instead of adding over all $k \in \{1, \dots, K\}$ as in the preceding proofs, we add over the set of maximum weight schedules, i.e., over all $k \in \mathcal{K}(x)$, as dictated by Equations (28) and (29).

Appendix B. Technical Results in the Delay Analysis of the Back-Pressure Policy Under Heavy-Tailed Traffic

The Switched Queueing Network in Section 4.1 as a Stochastic Processing Network

Stochastic processing networks (SPNs) are a general class of queueing systems, aiming to capture the dynamics and decisions in a wide range of settings in services and manufacturing. Since their introduction in Harrison [32], several variations and extensions of the original framework have appeared in the literature. In this section we give a brief overview of SPNs, and we show that the multihop switched queueing network and the back-pressure policy described in Section 4.1 are special cases of the SPN model and the maximum pressure policy studied in Dai and Lin [20], respectively. The reason we do this is twofold: (i) all fluid models of multihop switched queueing networks under the back-pressure policy that have appeared thus far assume fixed routing, e.g., Bui et al. [17], Ji et al. [38, 37] and Liu et al. [43]. By appealing to the very broad modeling class of SPNs we are able to extract a concrete fluid model for back-pressure that allows for multiple source-destination paths and loops, both quite common characteristics of real-world networks; (ii) through this mapping, certain technical lemmas that are needed for the delay analysis of back-pressure will follow directly from Dai and Lin [20].

An SPN can be described in terms of four entities: buffers, jobs, processors, and activities. In our multihop network context, buffers correspond to queues, jobs correspond to packets, and processors correspond to links. SPNs also include a special buffer, termed buffer 0, where all jobs waiting to enter the network are queued. However, what makes the comparison between the two models a nontrivial task is the notion of activity, an equivalent of which does not exist in switched networks. An activity can simultaneously process jobs from a set of buffers. To do this, it requires the simultaneous occupation of a set of processors. Each activity has a certain processing time, upon the completion of which jobs depart from the associated buffers and may arrive at other buffers. Depending on the availability of processors, multiple activities may be undertaken at the same time. In general, there are two types of activities: input activities that process jobs only from buffer 0, and service activities that process jobs only from the other buffers. Upon the completion of an input activity, jobs depart from buffer 0 and arrive to certain buffers. Upon the completion of a service activity, jobs depart from some buffers (but not buffer 0) and arrive to other buffers. In the context of the multihop network described above an input activity is, essentially, an exogenous arrival process, while a service activity is a queue-link allocation that satisfies the routing constraints imposed by the sets $\mathcal{L}_f$.

Dai and Lin [20] study two variations of SPNs. The first assumes that the capacities of processors are infinitely divisible, so that multiple activities can be undertaken at the same time, at utilization level less than 100% at each one. The second variation assumes that the capacities of servers are nondivisible, so that activities can be undertaken at utilization level 100%, or not at all. Since a link can serve packets from only one queue at any given time slot, the multihop network described above clearly falls within the class of SPNs with nondivisible server capacities.

In the SPNs considered in Dai and Lin [20], activities have general processing requirements and can be preempted by other activities before their completion. The in-service jobs of a preempted activity are "frozen," and their service is resumed only when that activity is undertaken again. In our discrete time model described above, the processing requirement of all activities is equal to one time slot. Moreover, the decision of which activities to undertake is made at the beginning of each time slot. Thus, in our model activities are never preempted and there are no "frozen" jobs.

An important characteristic of the SPN model in Dai and Lin [20] is that, for an activity to be undertaken at any given point in time, there have to be jobs available for processing at each of the constituent buffers. In other words, if a certain buffer is empty then activities that process jobs from that buffer cannot be undertaken. In the language of multihop networks, a queue is served only if it has packets available for transmission, which implies that there are no wasted service opportunities.

With these correspondences, it can be verified that the multihop network described in Section 4.1 is a special case of the SPN analyzed in Dai and Lin [20]. Moreover, our version of back-pressure does not waste service opportunities, and it is a maximum pressure policy, i.e., it satisfies Equation (7) of Dai and Lin [20].

As a final remark, we note that Assumption 1 of Dai and Lin [20] holds in switched networks, while the static planning problem defined by Equations (24)–(27) of Dai and Lin [20] is, essentially, the stability region given in Definition 5.

As alluded to earlier, our main motivation for viewing multihop networks as SPNs is to take advantage of known results from the SPN literature, which will serve as intermediate lemmas for the purposes of this section. One such result is the throughput optimality of the back-pressure policy. This property was first proved in Tassiulas and Ephremides [62] for the original version of back-pressure assuming light-tailed traffic. The following lemma establishes the throughput optimality of the slightly modified version of back pressure introduced in Section 4.1, and in the presence of heavy-tailed traffic.

Lemma 8 (Throughput Optimality of Back Pressure). *The multihop switched queueing network described in Section 4.1 is stable under the back-pressure policy for any arrival rate vector in the stability region.*

Proof (Outline). Under the back-pressure policy and our independence assumptions on the arrival processes, the sequence $\{Q(t); t \in \mathbb{Z}_+\}$ is a time-homogeneous, irreducible, and aperiodic Markov chain on a countable state space. The fact that this Markov chain is also positive recurrent is implied by results in Dai and Lin [20]. More specifically, Theorem 8 of Dai and Lin [20] implies that the fluid model of the multihop network under the back-pressure policy is weakly stable, i.e., if we consider the fluid model solution with the queues being initially empty, then the queue-length part of the solution remains zero. Then, Theorem 3 of Dai and Lin [20] implies that the stochastic system is pathwise stable, i.e., the long-term departure rates are equal to the respective long-term arrival rates, for all queues. Of course, this is a weaker form of stability compared to the one adopted in this paper. However, we have defined the stability region of the multihop network in terms of strict inequalities for all link capacity constraints. As a consequence, we can strengthen this result in a straightforward manner, very similarly to the way Theorem 5 in Dai and Lin [20] strengthens their Theorem 4. In particular, it can be verified that the fluid model is stable, i.e., if we consider the fluid model solution from a finite initial condition, then the queue-length part of the solution becomes zero in finite time, and remains zero thereafter. Consequently, according to Theorem 3.1 in Dai [19], the Markov chain that describes the stochastic system is positive recurrent. Hence, $\{Q(t); t \in \mathbb{Z}_+\}$ converges in distribution, and its limiting distribution does not depend on $Q(0)$. Because of the latter fact, 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 Sigman and Wolff [59]. □

Existence and Uniqueness Results for the Fluid Model in Section 4.2

In this section we state and prove the existence of a fluid limit, and the existence and uniqueness of the solution to the fluid model in Section 4.2. Both results are necessary for the fluid approximations-based delay analysis proposed in Section 4.3.

Lemma 9 (Existence of Fluid Limit and FMS). *Consider the FM in Section 4.2. There exists a Lipschitz continuous function $z(t) = \{z_{f,i}(t); i \in \mathcal{N}_f, f \in \mathcal{F}\}$, $t \in [0, T]$, such that for every $\epsilon > 0$ there exists $b_0(\epsilon)$ so that*

$$\mathbb{P}(H_b) \geq 1 - \epsilon \quad and \quad \sup_{t \in [0,T]} \max_{f \in \mathcal{F}} \max_{i \in \mathcal{N}_f} |\bar{q}^b_{f,i}(t) - z_{f,i}(t)| \leq \epsilon, \quad \forall \omega \in H_b,$$

for all $b \geq b_0(\epsilon)$. Additionally, there exists a Lipschitz continuous function $w(\cdot)$, such that $(z(\cdot), w(\cdot))$ is an FMS from initial condition $q(0) = q$ over the interval $[0, T]$.

Proof. The fact that $\mathbb{P}(H_b)$ converges to one as b goes to infinity was proved in Lemma 2. The existence of a fluid limit, and that a fluid limit is an FMS, follows directly from Appendix A of Dai and Lin [20]. □

As discussed in Section 2.1, the uniqueness of the FMS of a single-hop network under the max-weight policy has been established in Subramanian [61]. However, we are not aware of any similar uniqueness results in a multihop context. For this reason, a complete proof is provided using a strategy similar to that in Subramanian [61].

Lemma 10 (Uniqueness and Continuity of FMS). *Consider the FM in Section 4.2. For any given $q = \{q_{f,i} \in \mathbb{R}_+, i \in \mathcal{N}_f, f \in \mathcal{F}\}$ there exists a (unique) Lipschitz continuous function $z(t) = \{z_{f,i}(t); i \in \mathcal{N}_f, f \in \mathcal{F}\}$, $t \in [0, T]$, such that the queue-length part of every FMS from initial condition q is $z(\cdot)$. Moreover, $z(\cdot)$ depends continuously on both the initial condition q and the arrival rate vector λ.*

Proof. The existence of an FMS was established in Lemma 9. Regarding the uniqueness of the FMS, we proceed as follows: fix time $T > 0$, initial condition $v(0) = (v_{f,i}(0); i \in \mathcal{N}_f, f \in \mathcal{F})$, arrival rate vector λ^v, and let $v(\cdot)$ be the queue-length part of the FMS from $v(0)$ on the interval $[0, T]$, in vector form. Equation (41) implies that, at any regular time $t \in [0, T]$, this solution satisfies

$$\dot{v}_{f,i}(t) = -\sum_{j: (i,j) \in \mathcal{L}_f} \dot{s}^v_{f,i,j}(t) + \sum_{j: (j,i) \in \mathcal{L}_f} \dot{s}^v_{f,j,i}(t) + \lambda^v_f \cdot 1_{\{i = s_f\}},$$

$i \in \mathcal{N}_f$, $f \in \mathcal{F}$. Also, let $w(\cdot)$ be the queue-length part of the FMS from initial condition $w(0)$ on the interval $[0, T]$, under arrival rate vector λ^w. Similarly, this solution satisfies

$$\dot{w}_{f,i}(t) = -\sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^w_{f,i,j}(t) + \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^w_{f,j,i}(t) + \lambda^w_f \cdot 1_{\{i=s_f\}},$$

$i \in \mathcal{N}_f$, $f \in \mathcal{F}$. We measure the distance between the queue-length parts of the two solutions with the square of the Euclidean norm of their difference:

$$\|v(t) - w(t)\|_2^2 = \sum_{f\in\mathcal{F}} \sum_{i\in\mathcal{N}_f} (v_{f,i}(t) - w_{f,i}(t))^2.$$

At any regular time $t \in [0, T]$,

$$\frac{d}{dt}\|v(t) - w(t)\|_2^2 = 2\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} v_{f,i}(t)\dot{v}_{f,i}(t) + 2\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} w_{f,i}(t)\dot{w}_{f,i}(t) - 2\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} v_{f,i}(t)\dot{w}_{f,i}(t) - 2\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} w_{f,i}(t)\dot{v}_{f,i}(t).$$

We have that

$$v_{f,i}(t) \cdot \dot{v}_{f,i}(t) = -v_{f,i}(t) \cdot \sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^v_{f,i,j}(t) + v_{f,i}(t) \cdot \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^v_{f,j,i}(t) + v_{f,i}(t) \cdot \lambda^v_f \cdot 1_{\{i=s_f\}},$$

$$w_{f,i}(t) \cdot \dot{w}_{f,i}(t) = -w_{f,i}(t) \cdot \sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^w_{f,i,j}(t) + w_{f,i}(t) \cdot \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^w_{f,j,i}(t) + w_{f,i}(t) \cdot \lambda^w_f \cdot 1_{\{i=s_f\}},$$

$$v_{f,i}(t) \cdot \dot{w}_{f,i}(t) = -v_{f,i}(t) \cdot \sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^w_{f,i,j}(t) + v_{f,i}(t) \cdot \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^w_{f,j,i}(t) + v_{f,i}(t) \cdot \lambda^w_f \cdot 1_{\{i=s_f\}},$$

$$w_{f,i}(t) \cdot \dot{v}_{f,i}(t) = -w_{f,i}(t) \cdot \sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^v_{f,i,j}(t) + w_{f,i}(t) \cdot \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^v_{f,j,i}(t) + w_{f,i}(t) \cdot \lambda^v_f \cdot 1_{\{i=s_f\}},$$

for all $i \in \mathcal{N}_f$, and for all $f \in \mathcal{F}$. Therefore,

$$\frac{d}{dt}\|v(t) - w(t)\|_2^2 = 2\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} (\lambda^v_f - \lambda^w_f)(v_{f,i}(t) - w_{f,i}(t))1_{\{i=s_f\}} + A + B,$$

where the term A is equal to

$$-2\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} v_{f,i}(t)\left(\sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^v_{f,i,j}(t) - \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^v_{f,j,i}(t)\right) + 2\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} v_{f,i}(t)\left(\sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^w_{f,i,j}(t) - \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^w_{f,j,i}(t)\right),$$

and the term B is equal to

$$-2\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} w_{f,i}(t)\left(\sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^w_{f,i,j}(t) - \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^w_{f,j,i}(t)\right) + 2\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} w_{f,i}(t)\left(\sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^v_{f,i,j}(t) - \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^v_{f,j,i}(t)\right).$$

Now notice that, by rearranging the terms, we have that

$$\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} v_{f,i}(t)\left(\sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^v_{f,i,j}(t) - \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^v_{f,j,i}(t)\right) = \sum_{f\in\mathcal{F}}\sum_{(i,j)\in\mathcal{L}_f} \dot{s}^v_{f,i,j}(t)\left(v_{f,i}(t) - v_{f,j}(t)\right),$$

and

$$\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} w_{f,i}(t)\left(\sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^w_{f,i,j}(t) - \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^w_{f,j,i}(t)\right) = \sum_{f\in\mathcal{F}}\sum_{(i,j)\in\mathcal{L}_f} \dot{s}^w_{f,i,j}(t)\left(w_{f,i}(t) - w_{f,j}(t)\right).$$

The identities above, together with Equation (47), imply that

$$\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} v_{f,i}(t)\left(\sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^v_{f,i,j}(t) - \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^v_{f,j,i}(t)\right) \geq \sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} v_{f,i}(t)\left(\sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^w_{f,i,j}(t) - \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^w_{f,j,i}(t)\right),$$

and

$$\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} w_{f,i}(t)\left(\sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^w_{f,i,j}(t) - \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^w_{f,j,i}(t)\right) \geq \sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} w_{f,i}(t)\left(\sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}^v_{f,i,j}(t) - \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}^v_{f,j,i}(t)\right),$$

so that both terms A and B are nonpositive.

Consequently,

$$\frac{d}{dt}\|v(t) - w(t)\|_2^2 \leq 2\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} (\lambda^v_f - \lambda^w_f)(v_{f,i}(t) - w_{f,i}(t)) \cdot 1_{\{i=s_f\}} \leq 2\|\lambda^v - \lambda^w\|_\infty \sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} |v_{f,i}(t) - w_{f,i}(t)|$$

$$\leq 2\|\lambda^v - \lambda^w\|_\infty (\|v(t) - w(t)\|_2^2 + 1).$$

Finally, Gronwall's inequality and the fact that $v(\cdot)$ and $w(\cdot)$ are differentiable almost everywhere imply that

$$\|v(t) - w(t)\|_2^2 \leq \|v(0) - w(0)\|_2^2 \exp(2t\|\lambda^v - \lambda^w\|_\infty) + 2t\|\lambda^v - \lambda^w\|_\infty, \quad \forall\, t \in [0, T]. \tag{B.1}$$

If $v(0) = w(0)$ and $\lambda^v = \lambda^w$, so that $v(\cdot)$ and $w(\cdot)$ represent two solutions to the FM for a given initial condition and arrival rate vector, then Equation (B.1) implies that

$$\|v(t) - w(t)\|_2^2 = 0, \quad \forall\, t \in [0, T],$$

resulting in the uniqueness of the queue-length part of the FMS. The continuity with respect to the initial condition and arrival rate vector follows directly from Equation (B.1). □

Similarly to the single-hop case (cf. Lemma 3), the above lemma guarantees only the uniqueness of the queue-length part of the FMS. One can, again, construct simple examples where the service part of the FMS from zero initial condition is not unique.

Endnotes

[1] We note that while in single-hop switched queueing networks (i.e., the main focus of the present paper) this is always the case, in a multihop setting max-weight scheduling needs to be combined with back-pressure routing for throughput optimality to be guaranteed; see Dai and Lin [20] and Tassiulas and Ephremides [62]. Otherwise, dynamic instability phenomena may, again, arise; see Andrews and Zhang [1] and Bramson et al. [16].

[2] This statement depends critically on the assumption that the service discipline within each queue is FCFS. Under different disciplines, e.g., processor sharing or last come, last served that are known to be insensitive to the tails of the arriving traffic, this may not be the case. However, FCFS is prevalent in the applications that motivate the study of switched queueing networks.

[3] This is to be contrasted to the more common use of fluid approximations for stability analysis, where a Lyapunov function $V(\cdot)$ for the fluid model implies the existence of *another* Lyapunov function $G(\cdot)$ for the stochastic model; see, e.g., Bramson [15].

References

[1] Andrews M, Zhang L (2003) Achieving stability in networks of input-queued switches. *IEEE/ACM Trans. Networking* 11(5):848–857.
[2] Avrachenkov K, Ayesta U, Brown P, Nunez-Queija R (2005) Discriminatory processor sharing revisited. *Proc. 24th Annual Joint Conf. IEEE Comput. Comm. Societies, INFOCOM '05* (IEEE, Piscataway, NJ), 784–795.
[3] Baccelli F, Foss S (2004) Moments and tails in monotone-separable stochastic networks. *Ann. Appl. Probab.* 14(2):612–650.
[4] Baccelli F, Foss S, Lelarge M (2005) Tails in generalized Jackson networks with subexponential service-time distributions. *J. Appl. Probab.* 42(2):513–530.
[5] Bertsimas D, Gamarnik D, Tsitsiklis JN (2001) Performance of multiclass Markovian queueing networks via piecewise linear Lyapunov functions. *Ann. Appl. Probab.* 11(4):1384–1428.
[6] Borovkov AA (1970) Factorization identities and properties of the distribution of the supremum of sequential sums. *Theory Probab. Appl.* 15(3):359–402.
[7] Borst SC, Zwart AP (2005) Fluid queues with heavy-tailed $M/G/\infty$ input. *Math. Oper. Res.* 30(4):852–879.
[8] Borst SC, Boxma OJ, Jelenkovic PR (2003a) Reduced-load equivalence and induced burstiness in GPS queues with long-tailed traffic flows. *Queueing Systems* 43(4):273–306.
[9] Borst SC, Mandjes M, Van Uitert MJG (2003c) Generalized processor sharing with light-tailed and heavy-tailed input. *IEEE/ACM Trans. Networking* 11(5):821–834.
[10] Borst SC, Nunez-Queija R, Zwart AP (2006) Sojourn time asymptotics in processor-sharing queues. *Queueing Systems* 53(1):31–51.
[11] Borst SC, Van Ooteghem DTMB, Zwart AP (2005) Tail asymptotics for discriminatory processor-sharing queues with heavy-tailed service requirements. *Performance Evaluation* 61(2–3):281–298.
[12] Borst SC, Boxma OJ, Nunez-Queija R, Zwart AP (2003b) The impact of the service discipline on delay asymptotics. *Performance Evaluation* 54(2):175–206.
[13] Boxma OJ, Dumas V (1998) Fluid queues with long-tailed activity period distributions. *Comput. Comm.* 21(17):1509–1529.
[14] Boxma OJ, Deng Q, Zwart AP (2002) Waiting-time asymptotics for the $M/G/2$ queue with heterogeneous servers. *Queueing Systems* 40(1):5–31.
[15] Bramson M (2008) Stability of queueing networks. *Probab. Surveys* 5:169–345.
[16] Bramson M, D'Auria B, Walton N (2016) Instability of max-weight in multi-hop networks. Working paper, University of Minnesota.
[17] Bui L, Srikant R, Stolyar A (2011) A novel architecture for reduction of delay and queueing structure complexity in the back-pressure algorithm. *IEEE/ACM Trans. Networking* 19(6):1597–1609.
[18] Cohen JW (1973) Some results on regular variation for distributions in queueing and fluctuation theory. *J. Appl. Probab.* 10(2):343–353.
[19] Dai JG (1995) On positive Harris recurrence of multiclass queueing networks: A unified approach via fluid limit models. *Ann. Appl. Probab.* 5(1):49–77.
[20] Dai JG, Lin W (2005) Maximum pressure policies in stochastic processing networks. *Oper. Res.* 53(2):197–218.
[21] D'Auria B, Samorodnitsky G (2005) Limit behavior of fluid queues and networks. *Oper. Res.* 53(6):933–945.
[22] Down D, Meyn SP (1997) Piecewise linear test functions for stability and instability of queueing networks. *Queueing Systems* 27(3–4):205–226.
[23] Ersoz D, Yousif M, Das C (2007) Characterizing network traffic in a cluster-based, multi-tier data center. *Proc. 27th IEEE Internat. Conf. Distributed Comput. Systems, ICDCS '07* (IEEE Computer Society, Washington, DC) .

[24] Fisher M, Raman A (1996) Reducing the cost of demand uncertainty through accurate response to early sales. *Oper. Res.* 44(1):87–99.

[25] Foss S, Korshunov D (2012) On large delays in multi-server queues with heavy tails. *Math. Oper. Res.* 37(2):201–218.

[26] Gamarnik D, Katz D (2009) On deciding stability of multiclass queueing networks under buffer priority scheduling policies. *Ann. Appl. Probab.* 19(5):2008–2037.

[27] Gans N, Van Ryzin G (1997) Optimal control of a multi-class, flexible queueing system. *Oper. Res.* 45(5):677–693.

[28] Ganti A, Modiano E, Tsitsiklis JN (2007) Optimal transmission scheduling in symmetric communication models with intermittent connectivity. *IEEE Trans. Inform. Theory* 53(3):998–1008.

[29] Georgiadis L, Neely M, Tassiulas L (2006) Resource allocation and cross-layer control in wireless networks. *Foundations and Trends in Networking* 1(1):1–144.

[30] Gut A (2005) *Probability: A Graduate Course.* (Springer, New York).

[31] Hajek B (1982) Hitting-time and occupation-time bounds implied by drift analysis with applications. *Adv. Appl. Probab.* 14(3):502–525.

[32] Harrison JM (2000) Brownian models of open processing networks: Canonical representation of workload. *Ann. Appl. Probab.* 10(1): 75–103.

[33] Jagannathan K, Markakis MG, Modiano E, Tsitsiklis JN (2012) Queue-length asymptotics for generalized max-weight scheduling in the presence of heavy-tailed traffic. *IEEE/ACM Trans. Networking* 20(4):1096–1111.

[34] Jelenkovic PR, Lazar AA (1999) Asymptotic results for multiplexing subexponential on-off processes. *Adv. Appl. Probab.* 31(2):394–421.

[35] Jelenkovic PR, Momcilovic P (2002) Finite buffer queue with generalized processor sharing and heavy-tailed input processes. *Comput. Networks* 40(3):433–443.

[36] Jelenkovic PR, Momcilovic P (2003) Large deviation analysis of subexponential waiting times in a processor sharing queue. *Math. Oper. Res.* 28(3):587–608.

[37] Ji B, Joo C, Shroff NB (2013a) Delay-based back-pressure scheduling in multihop wireless networks. *IEEE/ACM Trans. Networking* 21(5):1539–1552.

[38] Ji B, Joo C, Shroff NB (2013b) Throughput-optimal scheduling in multihop wireless networks without per-flow information. *IEEE/ACM Trans. Networking* 21(2):634–647.

[39] Kotopoulos C, Likhanov N, Mazumdar RR (2001) Asymptotic analysis of the GPS system fed by heterogeneous long-tailed sources. *Proc. 20th Annual Joint Conf. IEEE Comput. Comm. Societies, INFOCOM '05* (IEEE, Piscataway, NJ), 299–308.

[40] Kumar PR, Seidman T (1990) Dynamic instabilities and stabilization methods in distributed real-time scheduling of manufacturing systems. *IEEE Trans. Automatic Control* 35(3):289–298.

[41] Lelarge M (2009) Asymptotic behavior of generalized processor sharing queues under subexponential assumptions. *Queueing Systems* 62(1–2):51–73.

[42] Likhanov N, Mazumdar RR (2000) Loss asymptotics in large buffers fed by heterogeneous long-tailed sources. *Adv. Appl. Probab.* 32(4):1168–1189.

[43] Liu J, Stolyar A, Chiang M, Poor HV (2009) Queue back-pressure random access in multi-hop wireless networks: Optimality and stability. *IEEE Trans. Inform. Theory* 55(9):4087–4098.

[44] Maguluri ST, Srikant R, Ying L (2012) Stochastic models of load balancing and scheduling in cloud computing clusters. Greenberg AG, Sohraby K, eds. *Proc. 31st Annual IEEE Internat. Conf. Comput. Comm., INFOCOM '12* (IEEE, Piscataway, NJ), 702–710.

[45] Markakis MG (2013) Scheduling in switched queueing networks with heavy-tailed traffic. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

[46] Markakis MG, Modiano E, Tsitsiklis JN (2014) Max-weight scheduling in queueing networks with heavy-tailed traffic. *IEEE/ACM Trans. Networking* 22(1):257–270.

[47] Markakis MG, Modiano E, Tsitsiklis JN (2016) Delay stability of back-pressure policies in the presence of heavy-tailed traffic. *IEEE/ACM Trans. Networking* 24(4):2046–2059.

[48] McKeown N, Mekkittikul A, Anantharam V, Walrand J (1999) Achieving 100% throughput in an input-queued switch. *IEEE Trans. Comm.* 47(8):1260–1267.

[49] Neely MJ (2008) Order optimal delay for opportunistic scheduling in multi-user wireless uplinks and downlinks. *IEEE/ACM Trans. Networking* 16(5):1188–1199.

[50] Nunez-Queija R (2002) Queues with equally heavy sojourn time and service requirement distributions. *Ann. Oper. Res.* 113(1):101–117.

[51] Pakes AG (1975) On the tails of waiting-time distributions. *J. Appl. Probab.* 12(3):555–564.

[52] Park K, Willinger W (2000) Self-similar network traffic: An overview. *Self-Similar Network Traffic and Performance Evaluation* (John Wiley & Sons, New York), 1–38.

[53] Rege KM, Sengupta B (1994) A decomposition theorem and related results for the discriminatory processor sharing queue. *Queueing Systems* 18(3):333–351.

[54] Resnick S, Samorodnitsky G (1999) Activity periods of an infinite server queue and performance of certain heavy tailed fluid queues. *Queueing Systems* 33(1):43–71.

[55] Rybko A, Stolyar A (1992) Ergodicity of stochastic processes describing the operation of open queueing networks. *Problems Inform. Transmission* 28(3):199–220.

[56] Shah D, Wischik D (2012) Switched networks with maximum weight policies: Fluid approximation and multiplicative state space collapse. *Ann. Appl. Probab.* 22(1):70–127.

[57] Shah D, Tsitsiklis JN, Zhong Y (2010) Qualitative properties of α-weighted scheduling policies. Misra V, Barford P, Squillante MS, eds. *Proc. ACM SIGMETRICS Internat. Conf. Measurement Modeling Comput. Systems, SIGMETRICS '10* (ACM, New York), 239–250.

[58] Sharifnassab A (2016) Sharif University, Iran. Personal communication with the authors, October 12, 2016.

[59] Sigman K, Wolff R (1993) A review of regenerative processes. *SIAM Rev.* 35(2):269–288.

[60] Stolyar A (2004) Maxweight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic. *Ann. Appl. Probab.* 14(1):1–53.

[61] Subramanian V (2010) Large deviations of max-weight scheduling policies on convex rate regions. *Math. Oper. Res.* 35(4):881–910.

[62] Tassiulas L, Ephremides A (1992) Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Trans. Automatic Control* 37(12):1936–1948.

[63] Tsitsiklis JN (1987) Analysis of a multiaccess control scheme. *IEEE Trans. Automatic Control* 32(11):1017–1020.

[64] Van Uitert MJG, Borst SC (2002) A reduced-load equivalence for generalised processor sharing networks with long-tailed input flows. *Queueing Systems* 41(1):123–163.

[65] Venkataraman VJ, Lin X, Ying L, Shakkottai S (2010) On scheduling for minimizing end-to-end buffer usage over multihop wireless networks. *Proc. 29th IEEE Internat. Conf. Comput. Comm., INFOCOM '10* (IEEE, Piscataway, NJ), 2847–2855.

[66] Veraverbeke N (1977) Asymptotic behaviour of Wiener-Hopf factors of a random walk. *Stochastic Processes and Their Appl.* 5(1):27–37.

[67] Williams D (1991) *Probability with martingales* (Cambridge University Press, New York). .

[68] Zwart AP, Boxma OJ (2000) Sojourn time asymptotics in the $M/G/1$ processor sharing queue. *Queueing Systems* 35(1):141–166.

[69] Zwart AP, Borst SC, Mandjes M (2004) Exact asymptotics for fluid queues fed by multiple heavy-tailed on-off flows. *Ann. Appl. Probab.* 14(2):903–957.