Delay Stability of Back-Pressure Policies in the presence of Heavy-Tailed Traffic

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

Abstract—**We study multi-hop networks with flow-scheduling constraints, no constraints on simultaneous activation of different links, potentially multiple source-destination routes, and a mix of heavy-tailed and light-tailed traffic. In this setting, we analyze the delay performance of the widely studied class of Back-Pressure scheduling policies, known for their throughput optimality property, using as a performance criterion the notion of delay stability, i.e., whether the expected end-to-end delay in steady state is finite. Our analysis highlights the significance of "bottleneck links," i.e., links that are allowed to serve the source queues of heavy-tailed flows. The main insight is that traffic that has to pass through bottleneck links experiences large delays under Back-Pressure. By means of simple examples we provide insights into how the network topology, the routing constraints, and the link capacities may facilitate or hinder the ability of light-tailed flows to avoid bottlenecks. Our delay-stability analysis is greatly simplified by the use of fluid approximations, allowing us to derive analytical results that would have been hard to obtain through purely stochastic arguments. Finally, we show how to achieve the best performance with respect to the delay stability criterion, by using a parameterized version of the Back-Pressure policy.**

I. Introduction

We study scheduling problems arising in multi-hop wireline networks with a mix of heavy-tailed (i.e., arrival processes with infinite variance) and light-tailed traffic and, potentially, multiple source-destination routes for each traffic flow. We analyze the delay performance of the widely studied class of Back-Pressure policies, known for their throughput optimality property. (More concretely, we focus on a particular variant of Back-Pressure policies, known as "Max-Pressure.") Classical results, e.g., the Pollaczek-Khinchin formula, imply that heavy-tailed flows experience large delays, infinite in steady-state expectation. Thus, we focus on the (policy-dependent) impact of heavy-tailed traffic on light-tailed flows, using as a performance criterion the notion of delay stability, i.e., whether the expected end-to-end delay of a traffic flow in steady state is finite.

The class of Back-Pressure policies was introduced in the seminal work of Tassiulas and Ephremides [21], and since then numerous studies have analyzed these policies in a variety of settings; see [6] for an overview. A remarkable property of Back-Pressure policies is their throughput optimality, i.e., their ability to stabilize a queueing network whenever this is possible. Moreover, Back-Pressure policies have been combined with congestion control in "cross-layer control" schemes that are provably stabilizing and utility-optimizing, e.g., see [5], [18].

We are motivated to study networks with a mix of heavy-tailed and light-tailed traffic by empirical evidence of strong correlations and statistical similarity over different time scales in real-world networks. This observation was first made by Leland *et al.* [13] through analysis of Ethernet traffic traces. Subsequent empirical studies have documented this phenomenon in other networks, while accompanying theoretical studies have associated it with bursty/heavy-tailed arrivals.

The impact of heavy tails has been analyzed extensively in relatively simple queueing systems, e.g., single or multi-server queues; for an overview of existing results see [1] and the references therein. Moreover, as alluded to above, there is vast literature on the performance of Back-Pressure policies under light-tailed traffic. However, the delay analysis of Back-Pressure policies in networks with a mix of heavy-tailed and light-tailed traffic has only recently attracted attention. Jagannathan *et al.* [9] consider a system with two parallel queues, receiving heavy-tailed and light-tailed traffic while sharing a single server, and determine the queue-length asymptotics under the Generalized Max-Weight policy. In follow-up work, Jagannathan *et al.* [10] study the case of a server with intermittent connectivity to the queues, and explore the impact of connectivity on queue length asymptotics. In a similar setting, Nair *et al.* [17] analyze the role of intra-queue scheduling, i.e., the way that jobs are served within each queue, on the response time asymptotics.

Closer to the present paper comes our earlier work [15], which studies the delay stability of Max-Weight policies (the single-hop equivalent of Back-Pressure) in networks with a mix of heavy-tailed and light-tailed traffic. In single-hop networks, the decision problem reduces to link-scheduling, i.e., which subset of communication links to activate at any given time slot. This determines directly which traffic flows are to be served, because in single-hop networks there is a one-to-one correspondence between links and flows. However, in multi-hop networks, multiple flows may traverse the same communication link. Thus, apart from link-scheduling, the decision problem in multi-hop networks has an additional dimension, that of flow-scheduling: given the links to be activated, which flow to send through each of them. This can

M.G. Markakis is affiliated with the Department of Economics and Business, Universitat Pompeu Fabra and the Barcelona School of Management (email: mihalis.markakis@upf.edu).

E. Modiano and J.N. Tsitsiklis are affiliated with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology (email: modiano@mit.edu; jnt@mit.edu).

The material in this paper was presented in part at the 2014 Information Theory and Applications Workshop, San Diego, CA, USA.

This work was supported by NSF grants CNS-1217048 and CMMI-1234062, the ONR grant N00014-12-1-0064, the ARO MURI grant W911NF-08-1-0238, the ECO2013-41131-R grant from the Spanish Ministry of Economy and Competitiveness, and a Marie Curie COFUND fellowship.

be interpreted as a joint scheduling and routing decision. This added dimension makes it very difficult to follow the stochastic dynamics of the systems, thus requiring new methodology for delay analysis. In particular, in this paper we utilize fluid approximation techniques that facilitate delay analysis without resorting to tracking the stochastic dynamics. Moreover, since the link-scheduling part of the problem has been analyzed extensively in [15], [16], here we consider a wireline multi-hop network model, where only the flow-scheduling part remains relevant. Thus, in the present paper we focus only on phenomena and insights that originate from the multi-hop nature of the network.

The main *contributions* of the paper can be summarized as follows.

(i) Through simple examples, we provide insights into how the network topology, the routing constraints, the arrival rates, and the link capacities may affect the delay performance of the Back-Pressure policy in the presence of heavy-tailed traffic.

(ii) We illustrate the value of the fluid approximation methodology, developed in the companion paper [16], for the delay analysis of multi-hop networks with heavy-tailed traffic. More specifically, fluid approximations simplify significantly the analysis, and allow us to obtain results that would have been hard to prove solely through stochastic arguments.

(iii) We show how one can achieve optimal performance with respect to the delay stability criterion by using a parameterized version of the Back-Pressure policy, provided the parameters are chosen suitably.

The remainder of the paper is organized as follows. Section II includes a detailed description of a multi-hop wireline network, together with some useful definitions and lemmas. In Section III we present briefly the fluid model of this multi-hop network, and state two essential results, from the companion paper [16], that associate fluid approximations to delay stability. In Section IV we show through simple examples, which "system parameters" may affect the delay performance of the Back-Pressure policy and in what way. Section V contains a delay-stability analysis of the parameterized Back-Pressure-α policy. We conclude with a discussion of our findings in Sections VI and VII.

II. A Multi-Hop Wireline Network under the Back-Pressure Policy

We start with a detailed description of the multi-hop switched queueing network studied in this paper. Subsequently, we present the Back-Pressure policy and we provide some useful definitions and lemmas.

We denote by $\mathbb{R}_+$, $\mathbb{Z}_+$, and $\mathbb{N}$ the sets of nonnegative reals, nonnegative integers, and positive integers, respectively. Also, $[x]^+$ represents $\max\{x, 0\}$, the nonnegative part of scalar x. Finally, 1_E stands for the indicator variable of event E.

The network operates in discrete time slots, which we index by $t \in \mathbb{Z}_+$. The topology of the network is captured by a directed graph $\mathcal{G} = \big(\mathcal{N}, \mathcal{L}\big)$, 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.

Central to our model is the notion of a traffic flow $f \in \mathcal{F}$, where $\mathcal{F} = \{1, \ldots, F\}$, $F \in \mathbb{N}$, which is a long-lived stream of packets that arrives to the network according to a discrete time stochastic arrival process $\{A_f(t);\ t \in \mathbb{Z}_+\}$. 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. The quantity $A_f(t)$ can be interpreted as the random number of packets that flow f brings (exogenously) to s_f at the end of time slot t. We use $A(t)$ to represent the vector $\big(A_f(t);\ f = 1, \ldots, F\big)$. In the remainder of the paper we use the terms "flow" and "traffic flow" interchangeably, while we often use "traffic" to refer to a specific collection of traffic flows. (To what collection we refer will be clear from the context.)

We assume that all arrival processes take values in $\mathbb{Z}_+$, and are independent and identically distributed (IID) over time. Furthermore, different arrival processes are independent. We denote by $\lambda_f = \mathbb{E}[A_f(0)] > 0$ the rate of traffic flow f and by $\lambda = (\lambda_f;\ f = 1, \ldots, F)$ the vector of the rates of all traffic flows.

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

An IID traffic flow is heavy-tailed/light-tailed/exponential-type if the distribution that underlies the corresponding arrival process is heavy-tailed/light-tailed/exponential-type, respectively. 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 exponential-type, and heavy-tailed otherwise. Definition 1 has been used in the literature on data communication networks, e.g., see [19], due to its close connection to long-range dependence.

For technical reasons we assume the existence of some $\gamma \in (0, 1)$ such that $\mathbb{E}\big[A_f^{1+\gamma}(0)\big] < \infty$, for all $f \in \mathcal{F}$.

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 source to destination, we say that flow f has *fixed 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 $\mathcal{L}_f$.

An additional assumption is that there are no "dead-ends," i.e., from every node in $i \in \mathcal{N}_f$ there exists a directed path within the links in $\mathcal{L}_f$ that leads to d_f.

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 each queue is "First Come, First Served." We use the shorthand notation $Q(t)$ for the set of queue lengths

$\{Q_{f,i}(t);\ i \in \mathcal{N}_f,\ f \in \mathcal{F}\}$. We also denote by $\mathcal{F}_t$ the σ-algebra generated by $Q(0), A(0), \ldots, Q(t-1), A(t-1), Q(t)$, which should be distinguished from the set of traffic flows $\mathcal{F}$.

Traffic may arrive to queue (f, i) either exogenously, if i is the source node s_f, 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 use the shorthand notation $S(t)$ for the set of scheduling decisions $\{S_{f,i,j}(t);\ (i, j) \in \mathcal{L}_f,\ f \in \mathcal{F}\},\ t \in \mathbb{Z}_+$. For simplicity, we assume that the capacity of all links is equal to one packet per time slot.

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

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

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 $S(t),\ t \in \mathbb{Z}_+$. For much of the paper 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.,

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

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 [4], is slightly different from the original, and most studied, version of Back-Pressure [21]. 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 our version. However, in the regime of large queue lengths/delays that we are interested in this paper, the two policies are indistinguishable.

With slight abuse of notation, i.e., if we now let $S(t)$ represent the final scheduling decisions made by the Back-Pressure policy at time slot t, the dynamics of the multi-hop switched queueing network can be written in the following form:

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

and

$$Q_{f,i}(t+1) = Q_{f,i}(t) - \sum_{j:(i,j)\in\mathcal{L}_f} S_{f,i,j}(t) + \sum_{j:(j,i)\in\mathcal{L}_f} S_{f,j,i}(t), \quad (2)$$

for all $i \in \mathcal{N}_f \setminus \{s_f, d_f\}$. Finally, by convention,

$$Q_{f,d_f}(t) = 0, \qquad \forall f \in \mathcal{F}. \quad (3)$$

The initial queue lengths are arbitrary nonnegative integers.

Coming to the issue of delays, a batch of packets arriving to the network 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 s_f, until the time slot that its last packet reaches d_f. For $k \in \mathbb{N}$, we denote by $D_f(k)$ the end-to-end delay of the k^{th} file of flow f, and use the vector notation $D(k) = \big(D_f(k);\ f = 1, \ldots, F\big)$.

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

Definition 2: (Stability Region) An arrival rate vector $\lambda = \big(\lambda_1, \ldots, \lambda_F\big)$ is in the stability region Λ of the multi-hop switched queueing network described above if there exist $\zeta_{f,i,j} \ge 0,\ f \in \mathcal{F},\ i, j \in \mathcal{N}$, such that the following set of constraints is satisfied:

(i) Flow Efficiency Constraints:

$$\zeta_{f,i,i} = \zeta_{f,i,s_f} = \zeta_{f,d_f,i} = 0, \qquad \forall i \in \mathcal{N}, \qquad \forall f \in \mathcal{F};$$

(ii) Routing Constraints:

$$\zeta_{f,i,j} = 0, \qquad \forall (i, j) \notin \mathcal{L}_f, \qquad \forall f \in \mathcal{F};$$

(iii) Flow Conservation Constraints:

$$\sum_{j\in\mathcal{N}} \zeta_{f,j,i} + \lambda_f \cdot 1_{\{i=s_f\}} = \sum_{j\in\mathcal{N}} \zeta_{f,i,j}, \quad \forall i \ne d_f, \quad \forall f \in \mathcal{F};$$

(iv) Link Capacity Constraints:

$$\sum_{f\in\mathcal{F}} \zeta_{f,i,j} < 1, \qquad \forall (i, j) \in \mathcal{L}.$$

The auxiliary ζ variables via which we define the stability region are often interpreted as "multicommodity flows," e.g., see [6].

If an arrival rate vector λ belongs to the stability region Λ, then there exists a policy that stabilizes the network, in the sense that the sequences $\{Q(t);\ t \in \mathbb{Z}_+\}$ and $\{D(k);\ k \in \mathbb{N}\}$ converge in distribution. [1] This can be shown by arguing similarly to Corollary 3.9 of [6], and by utilizing the independence assumptions that we made regarding the arrival processes, which imply that the underlying Markov chain is aperiodic.

Lemma 1: (Throughput Optimality of Back-Pressure) The multi-hop switched queueing network described above is stable under the Back-Pressure policy, for all $\lambda \in \Lambda$.

[1] Our definition of stability as positive recurrence of the underlying Markov chain of the network (cf. the weaker notion of "rate stability") is precisely the reason that we assume strict inequalities for the link capacity constraints in the definition of the stability region.

Proof: In the case of light-tailed traffic the result is well-known [21]; in the presence of heavy-tailed traffic, the result follows from the findings of [4]. For a formal proof the reader is referred to [16]. ■

All the networks that we analyze in this paper are under the assumption of an arrival rate vector in the respective stability region. We denote by $Q_{f,i}$ the steady-state length of queue (f,i), while we reserve D_f for the steady-state end-to-end delay of traffic flow f. The dependence of these random variables on the scheduling policy that is applied has been suppressed from the notation, but will be clear from the context.

Definition 3: (Delay Stability) Traffic flow f is delay stable under a specific policy if the network is stable under that policy and $\mathbb{E}[D_f]$ is finite; otherwise, f is delay unstable.

Similarly, queue (f,i) is delay stable if $\mathbb{E}\big[Q_{f,i}\big]$ is finite, and delay unstable otherwise. We note that the latter notion of delay stability is related to the delay of packets, whereas the former to the delay of files.

Theorem 1: (Delay Instability of Heavy Tails) Consider the multi-hop switched queueing network described above under any scheduling policy. The source queue of every heavy-tailed flow is delay unstable. Consequently, every heavy-tailed flow is delay unstable.

Proof: Consider the best case for the source queue of a heavy-tailed flow, which is that it is served at each time slot. Then, this queue is a discrete time $M/G/1$ queue with infinite variance of service time (here, a customer is equivalent to a file). The Pollaczek-Khinchin formula [20] implies that this queue is delay unstable. Then, the BASTA property (e.g., see Theorem 5 in [15]) implies that the heavy-tailed flow is delay unstable as well. This argument can be formalized in exactly the same way as in the proof of Theorem 1 in [15]. ■

Since there is little that can be done regarding the delay stability of heavy-tailed flows, we turn our attention to light-tailed traffic. It is well-known that in a multi-hop network with just light-tailed traffic and under the Back-Pressure policy, all traffic flows are delay stable [21]. However, the existence of flow-scheduling constraints couples the evolution of different queues and flows. Below we show that this coupling may cause light-tailed flows to become delay unstable, giving rise to a form of propagation of delay instability.

III. Delay Stability Analysis via Fluid Models

Before we proceed to the findings of this study, we briefly present the *Fluid Model (FM)* of the multi-hop network described above. Fluid models of multi-hop networks with *fixed routing* under the Back-Pressure policy have been employed in previous works in order to show stability, e.g., see [3], [11], [12], [14]. The FM presented below is derived from that in [4], which studies a Stochastic Processing Network (SPN) under the Max-Pressure policy. We note that the SPN in [4] is a quite general model that includes our multi-hop network as a special case, and which allows for multiple source-destination paths, as well as a variety of other capabilities beyond the scope of switched networks. We also state two results, from the companion paper [16], that relate fluid approximations to delay stability. We will make frequent use of these results throughout this paper, since they simplify significantly our delay stability analysis. An in-depth discussion about the derivation of the FM equations and the justification of the fluid approximation (existence of fluid limit, existence and uniqueness of fluid model solution) can be found in [16].

The FM of the multi-hop network of Section II under the Back-Pressure policy is a deterministic dynamical system that aims to capture the evolution of its stochastic counterpart on longer time scales. Fix $T \in \mathbb{R}_+$. The FM is defined by the following relations and differential equations, for every time $t \in [0,T]$ that the derivatives exist (such t is often called a regular time):

$$\dot{q}_{f,i}(t) = -\sum_{j:(i,j)\in\mathcal{L}_f} \dot{s}_{f,i,j}(t) + \sum_{j:(j,i)\in\mathcal{L}_f} \dot{s}_{f,j,i}(t) + \lambda_f \cdot 1_{\{i=s_f\}}, \tag{4}$$

$$q_{f,i}(t) \geq 0, \tag{5}$$

$$s_{f,i,j}(0) = 0 \quad \text{and} \quad \dot{s}_{f,i,j}(t) \geq 0, \tag{6}$$

$$\sum_{f:(i,j)\in\mathcal{L}_f} \dot{s}_{f,i,j}(t) \leq 1, \tag{7}$$

$$\Big[\exists f' : q_{f',i}(t) - q_{f',j}(t) > 0\Big] \Longrightarrow \sum_{f:(i,j)\in\mathcal{L}_f} \dot{s}_{f,i,j}(t) = 1, \tag{8}$$

$$\Big[q_{f',i}(t) - q_{f',j}(t) < \max_{f:(i,j)\in\mathcal{L}_f} \Big\{\big[q_{f,i}(t) - q_{f,j}(t)\big]^+\Big\}\Big] \Longrightarrow \dot{s}_{f',i,j}(t) = 0. \tag{9}$$

In the equations above $i \neq d_f$, $q_{f,i}(t)$ represents the length of queue (f,i) at time t and $s_{f,i,j}(t)$ represents the amount of time that link $(i,j) \in \mathcal{L}_f$ has been serving queue (f,i) up to time t. Eqs. (8)-(9) are the fluid model equations for the variant of Back-Pressure (Max-Pressure) that we use in this paper.

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

$$q_{f,d_f}(t) = 0. \tag{10}$$

Henceforth, we use the shorthand notation $q(t)$ for the set of queue lengths $\big\{q_{f,i}(t);\ i \in \mathcal{N}_f,\ f \in \mathcal{F}\big\}$, and $s(t)$ for the set of scheduling decisions $\big\{s_{f,i,j}(t);\ (i,j) \in \mathcal{L}_f,\ f \in \mathcal{F}\big\}$. A Lipschitz continuous function $\big(q(\cdot), s(\cdot)\big)$ satisfying Eqs. (4)-(10), for all $t \in [0,T]$, is called a *Fluid Model Solution (FMS)*.

The following result illustrates how fluid models can be used for proving delay instability in the presence of heavy-tailed traffic.

Theorem 2 (Delay Instability via Fluid Models [16]): Consider the multi-hop network of Section II under the Back-Pressure policy, and its FM described above. Let $h \in \mathcal{F}$ be a heavy-tailed traffic flow, and $q^*(\cdot)$ be the (necessarily unique) queue-length part of a 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 $(f,i) \neq (h, s_h)$ is delay unstable.

Proof: (Outline) Suppose that there exists $\tau \in [0,T]$ such that $q^*_{f,i}(\tau) > 0$. The results of [4] establish the existence of a fluid limit and, consequently, of a FMS. This, together with the uniqueness of the queue-length part of a FMS (which is established in [16]) imply that after a big arrival to queue (h, s_h), queue (f,i) builds to the order of magnitude of the heavy-tailed queue with high probability. In turn, renewal theory and Little's Law provide the desired delay instability result. For a formal proof the reader is referred to [16]. ■

Finally, the result that follows is helpful in proving delay stability in networks with a mix of heavy-tailed and exponential-type light-tailed traffic.

Theorem 3 (Delay Stability via Fluid Models [16]): Consider the multi-hop switched queueing network of Section II under the Back-Pressure policy, and its FM described above. Consider also a piecewise linear function $V : \mathbb{R}^F_+ \to \mathbb{R}_+$ of the form

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

where $\mathcal{J} = \{1, \ldots, J\}$ is the set of indices of the different pieces of the function, and where $c_{jf} \in \mathbb{R}$, for all $j \in \mathcal{J},\ f \in \mathcal{F}$. 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

$$\mathbb{E}\big[V\big(Q(t+b)\big) - V\big(Q(t)\big) + b\zeta;\ V\big(Q(t)\big) > \alpha b \ \big|\ \mathcal{F}_t\big] \leq 0,$$

for all $b \geq b_0$. This implies that the sequence $\big\{V\big(Q(t)\big);\ t \in \mathbb{Z}_+\big\}$ converges in distribution to the random variable $V(Q)$, where Q is the limiting distribution of $Q(t)$.

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

Proof: (Outline) The first part of the result is established by showing that if a continuous and piecewise linear Lyapunov function can be found for the FM, then the same function is a Lyapunov function for the stochastic system, if the latter is sampled once every T time steps, and T is large enough. The second part is established using results from [7], by showing that if this Lyapunov function has exponential-type "upward-jumps," then its stationary version is also exponential-type, which, in turn, leads to delay stability. For a formal proof the reader is referred to [16]. ■

IV. Delay Stability Analysis of Back-Pressure

We start by analyzing the performance of the Back-Pressure policy with respect to the delay stability criterion. By means of simple examples, we investigate the role of the network topology, the routing constraints, and the arrival rates relative to link capacities on the delay stability of queues and flows. Our analysis highlights the importance of links that are allowed to serve the source queues of heavy-tailed flows, which we call *bottleneck links*. If $h \in \mathcal{F}$ is a heavy-tailed traffic flow, the set of bottleneck links associated with h is defined as follows:

$$\mathcal{B}_h = \big\{(s_h, i) : (s_h, i) \in \mathcal{L}_h\big\}.$$

To illustrate the importance of bottleneck links let us consider the simple system of Figure 1, which includes two traffic flows, the heavy-tailed flow 1 and the light-tailed flow 2. Both flows arrive exogenously at node 1, their packets get buffered in the respective queues, eventually get transmitted through link $(1,0)$, and exit the network as soon as they reach node 0. Link $(1,0)$ is a bottleneck link, since it is allowed to serve the source queue of flow 1. It is not hard to see that this model is equivalent to a single-server system of two parallel queues, where the Back-Pressure policy reduces to Max-Weight scheduling. Theorem 2 of [15] implies that the light-tailed flow 2 is delay unstable. The main idea behind this result is that queue $(1,1)$ is occasionally very long due to the heavy-tailed arrivals that it receives exogenously. During those time periods, flow 1 has very large differential backlog over link $(1,0)$, which implies that under the Back-Pressure policy, queue $(2,1)$ is deprived of service until it builds up a comparable backlog.

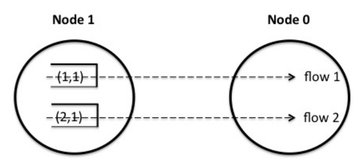

Fig. 1. A single-server system with two parallel queues, cast as a multi-hop network. Traffic flow 1 is heavy-tailed and traffic flow 2 is light-tailed. Since the network has single-hop traffic, the Back-Pressure policy reduces to Max-Weight scheduling. The findings of [15] imply that the light-tailed flow is delay unstable.

In general, *light-tailed flows experience large delays whenever they have to traverse bottleneck links. Consequently, the delay performance of Back-Pressure depends crucially on the ability of light-tailed flows to avoid bottlenecks, in static or dynamic ways*. This ability is dictated by a number of "system parameters," as we show below.

A. The Role of Network Topology

We start by illustrating the role of network topology in the delay stability of light-tailed flows. Consider the "line" network depicted in Figure 2. The heavy-tailed flow 1 arrives exogenously at node 1, eventually gets transmitted through link $(1,0)$, and exits the network as soon as it reaches node 0. The light-tailed flow 2 arrives exogenously at node 2, eventually gets transmitted through link $(2,1)$ first, and through link $(1,0)$ next, and exits the network when it reaches node 0. We are interested in the delay stability of flow 2 under the Back-Pressure policy.

Proposition 1: Consider the network of Figure 2 under the Back-Pressure policy. Traffic flow 2 is delay unstable.

Proof: This result is a special case of Theorem 4, which follows shortly. Here, we sketch the proof for the network

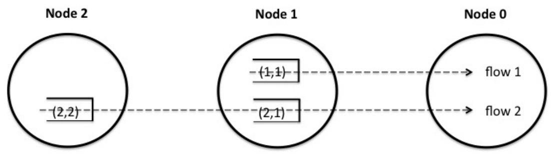

Fig. 2. The heavy-tailed flow 1 enters the network at node 1 and exits at node 0. The light-tailed flow 2 enters the network at node 2, and passing through node 1, it also exits the network at node 0. Traffic flow 2 is delay unstable under the Back-Pressure policy because it has to pass through the bottleneck link $(1, 0)$.

topology of Figure 2. The main idea behind it is that queue $(2, 1)$ becomes very long, occasionally, because it competes with the heavy-tailed queue $(1, 1)$ for link $(1, 0)$. During those occasions, there are no transmission from queue $(2, 2)$ to queue $(2, 1)$ under the Back-Pressure policy, unless queue $(2, 2)$ builds up to the order of magnitude of the heavy-tailed queue. This leads to the delay instability of both queues $(2, 1)$ and $(2, 2)$, and as a result of flow 2 as well. ■

The reason that traffic flow 2 is delay unstable is the topology of the network, and more specifically the fact that the only source-destination path of flow 2 includes a bottleneck link. We will see shortly that this condition leads to delay instability more generally.

B. The Role of Routing Constraints

We continue with the role of routing constraints. Consider the network of Figure 3: the heavy-tailed flow 1 arrives exogenously at node 1, and may reach its destination node 0 through the path $\big((1, 2), (2, 0)\big)$ or through the path $\big((1, 3), (3, 0)\big)$. The same applies to the light-tailed flow 2. In other words, both flows have dynamic routing. We are interested in the delay stability of flow 2 under the Back-Pressure policy.

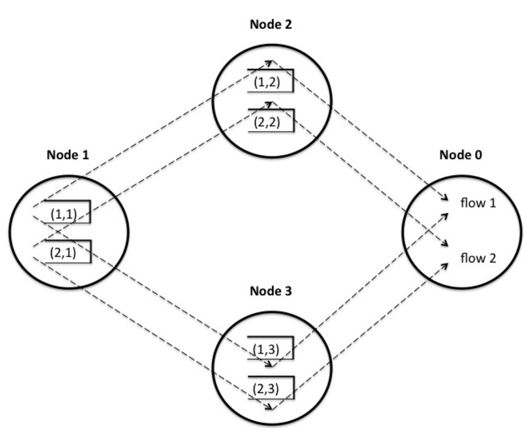

Fig. 3. Both the heavy-tailed flow 1 and the light-tailed flow 2 enter the network at node 1 and exit at node 0. They are both allowed to access all links of the network. Traffic flow 2 is delay unstable under the Back-Pressure policy because it has to pass through, either link $(1, 2)$ or link $(1, 3)$, which are both bottleneck links.

Proposition 2: Consider the network of Figure 3 under the Back-Pressure policy. Traffic flow 2 is delay unstable.

Proof: This result is another special case of Theorem 4, so here we only sketch the proof for the network topology of Figure 3. The main idea behind it is that whenever the heavy-tailed queue $(1, 1)$ receives exogenously a very large batch of packets, it creates simultaneously a very large differential backlog over links $(1, 2)$ and $(1, 3)$. Thus, under the Back-Pressure policy queue $(2, 1)$ will be denied access to both of those links, unless it builds up to a similar length. ■

The reason that traffic flow 2 is delay unstable in Figure 3 lies in the routing constraints of the heavy-tailed flow 1, or, more accurately, the lack of constraints. By not restricting the links that flow 1 is allowed to access, both links $(1, 2)$ and $(1, 3)$ become bottleneck links. In turn, all feasible source-destination paths of flow 2 pass through bottleneck links.

Similar conclusions can be reached if we force both flows 1 and 2 to follow the same fixed route to their destination node.

The insights derived from the simple examples of Figures 1-3 can be unified in a general result. We say that *traffic flow $f \in \mathcal{F}$ has to pass through a set of links $\mathcal{L}' \subset \mathcal{L}$, if every packet arriving at queue (f, s_f) must traverse one of the links in $\mathcal{L}'$ in order to reach d_f.*

Clearly, whether a traffic flow has to pass through a given set of links or not depends on the network topology, the routing constraints, and the routing policy applied.

Theorem 4: Consider the multi-hop switched queueing network of Section II under the Back-Pressure policy. Let $f \in \mathcal{F}$ be a light-tailed traffic flow. If there exists a heavy-tailed flow $h \in \mathcal{F}$ such that f has to pass through the set of bottleneck links $\mathcal{B}_h$, then f is delay unstable.

Proof: The proof of this result is based on Theorem 2, i.e., we study the evolution of the fluid model of the network from initial condition $q_{h,s_h}(0) = 1$ and zero for all other queues.

At time zero, the differential backlog of flow h over every link in $\mathcal{B}_h$ is 1, while the differential backlog of flow f over any of those links is zero. Moreover, the differential backlog of flow h can decrease at rate no more than $2\big|\mathcal{B}_h\big|$ (since the capacity of all links is equal to one), while the differential backlog of flow f can increase at rate no more than λ_f. So, there exists $\tau > 0$ such that

$$q_{h,i}(t) - q_{h,j}(t) > q_{f,i}(t) - q_{f,j}(t),$$

for all $t \in [0, \tau]$, for all $(i, j) \in \mathcal{B}_h$. Eq. (7) implies that flow f receives *none* of the available capacity of links in $\mathcal{B}_h$ during $[0, \tau]$ under the Back-Pressure policy. Therefore, the traffic of flow f that arrives exogenously at queue (f, s_f) cannot move past queue (f, s_h) during the interval $[0, \tau]$.

Now it is useful to view the total traffic of flow f between the source node s_f and the bottleneck node s_h (the source node of flow h) as one fictitious queue, whose length at time t is denoted by $\tilde{q}_f(t)$. The argument above implies that this queue has arrivals at rate $\lambda_f > 0$ and no departures during the interval $[0, \tau]$. Hence, $\tilde{q}_f(\tau) = \lambda_f \tau > 0$, so according to Theorem 2 the fictitious queue is delay unstable. This also implies the delay instability of flow f, since the delay experienced in the fictitious queue bounds from below the end-to-end delay, sample path-wise. ■

C. The Role of Link Capacities

In this section we illustrate the impact of link capacities, relative to the arrival rates, on the delay stability of light-tailed flows. Let us consider a variation of the network of Figure 3, where the heavy-tailed flow 1 has to reach node 0 through the path $\big((1,2),(2,0)\big)$, whereas the light-tailed flow 2 can access all links.

Let us first look at the case where $\lambda_1, \lambda_2 < 1$. The importance of this assumption lies in the fact that, with high probability, it allows flow 2 to route all its traffic through the path $\big((1,3),(3,0)\big)$ whenever the path of the heavy-tailed flow is congested.

Proposition 3: Consider the network of Figure 3 under the Back-Pressure policy, where flow 1 has fixed routing, along the path $\big((1,2),(2,0)\big)$, and flow 2 has dynamic routing. If the arrival rates satisfy $\lambda_1, \lambda_2 < 1$, then traffic flow 2 is delay stable.

Proof: Without loss of generality, we assume that all queues are empty at time slot zero. First, notice that no more than one packet per time slot arrives at nodes 2 and 3 because that is the capacity of links $(1,2)$ and $(1,3)$. Moreover, traffic departs from each of these nodes at rate of one packet per time slot, as long as there are packets waiting for transmission. This is due to the fact that both flows exit the network at node 0, so whenever packets are available, there is positive differential backlog over links $(2,0)$ and $(3,0)$. Therefore, it can be easily verified that

$$Q_{2,i}(t) \le 1, \qquad \forall t \in \mathbb{Z}_+, \qquad \forall i \in \{2,3\}.$$

Furthermore, Lemma 1 implies that the queue-length processes $\big\{Q_{2,2}(t);\ t \in \mathbb{Z}_+\big\}$ and $\big\{Q_{2,3}(t);\ t \in \mathbb{Z}_+\big\}$ converge to some limiting distributions $Q_{2,2}$ and $Q_{2,3}$, respectively. Hence,

$$\mathbb{E}\big[Q_{2,i}\big] \le 1, \qquad \forall i \in \{2,3\}.$$

Little's Law implies that both queues $(2,2)$ and $(2,3)$ are delay stable. In order to show that flow 2 is delay stable, it suffices to show that queue $(2,1)$ is delay stable as well.

Link $(1,3)$ is allowed to transmit only packets of flow 2, and, as we showed above, the length of queue $(2,3)$ is never more than one packet. Hence, under the Back-Pressure policy,

$$Q_{2,1}(t) > 1 \implies S_{2,1,3}(t) = 1, \qquad \forall t \in \mathbb{Z}_+.$$

Consider the candidate Lyapunov function $V(t) = Q_{2,1}^2(t)$. Through simple algebra, it can be verified that

$$\begin{aligned}
&\mathbb{E}\big[V(t+1) - V(t);\ V(t) > 1 \,\big|\, \mathcal{F}_t\big] \\
\le& -2\mathbb{E}\Big[\big(S_{2,1,2}(t) + S_{2,1,3}(t) - A_2(t)\big)Q_{2,1}(t); V(t) > 1 \mid \mathcal{F}_t\Big] \\
&+ \mathbb{E}\Big[\big(A_2(t)+2\big)^2;\ V(t) > 1 \,\Big|\, \mathcal{F}_t\Big] \\
\le& -2\mathbb{E}\Big[\big(S_{2,1,3}(t) - A_2(t)\big)Q_{2,1}(t);\ V(t) > 1 \,\Big|\, \mathcal{F}_t\Big] \\
&+ \mathbb{E}\Big[\big(A_2(t)+2\big)^2;\ V(t) > 1 \,\Big|\, \mathcal{F}_t\Big] \\
=& \Big(-2\big(1-\lambda_2\big)Q_{2,1}(t) + \mathbb{E}\Big[\big(A_2(0)+2\big)^2\Big]\Big) \cdot 1_{\{Q_{2,1}(t)>1\}}.
\end{aligned}$$

(We recall that $\mathcal{F}_t$ is the σ-algebra generated by $Q(0), A(0), \ldots, Q(t-1), A(t-1), Q(t)$.)

Notice that $\lambda_2 < 1$, $\mathbb{E}\big[A_2^2(0)\big] < \infty$, and $\big\{Q_{2,1}(t) \le 1\big\}$ is a finite set. Then, the Foster-Lyapunov stability criterion and moment bound (e.g., see Corollary 2.1.5 of [8]) implies that $\mathbb{E}[Q_{2,1}] < \infty$. Thus, all queues of flow 2 are delay stable, implying that traffic flow 2 is delay stable. ■

Now let us consider the case where $\lambda_2 > 1$. It is intuitively clear that irrespective of the specific routing decisions made at each time slot, a nonvanishing fraction of the traffic of flow 2 has to pass through the bottleneck link $(1,2)$. This fraction of the traffic experiences large delays under the Back-Pressure policy, which implies that the delays of flow 2 are, on average, large as well.

Proposition 4: Consider the network of Figure 3 under the Back-Pressure policy, where flow 1 has fixed routing, along the path $\big((1,2),(2,0)\big)$, and flow 2 has dynamic routing. If $\lambda_2 > 1$ then traffic flow 2 is delay unstable.

Proof: We will make use of Theorem 2, i.e., we will consider the FM of the network of Figure 3, with initial conditions $q_{1,1}(0) = 1$ and zero for all other queues. Eq. (7) implies the existence of $\tau > 0$, such that $\dot{s}_{2,1,2}(t) = 0$ for all $t \in [0,\tau]$. In turn, Eqs. (4) and (6) imply that

$$\dot{q}_{2,1}(t) = \lambda_2 - 1 > 0, \qquad \forall t \in [0,\tau].$$

Therefore, $q_{2,1}(\tau) > 0$, which implies that queue $(2,1)$ is delay unstable according to Theorem 2. Consequently, flow 2 is delay unstable since its end-to-end delay is bounded from below by the delay experienced in its source queue. ■

D. The Impact of Heavy Tails on Cross-Traffic

Consider the multi-hop network of Figure 4, which includes three traffic flows: the heavy-tailed flow 1, and the light-tailed flows 2 and 3. The source of flow 1 is node 2, whereas the source of flows 2 and 3 is node 1. The destination of flows 1 and 2 is node 3, whereas the destination of flow 3 is node 4.

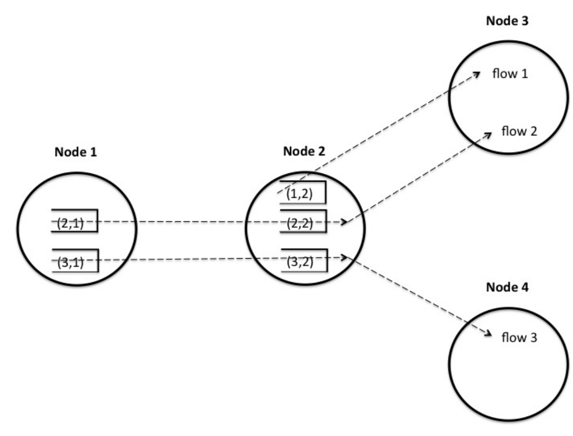

Fig. 4. The heavy-tailed flow 1 enters the network at node 2 and exits at node 3. The light-tailed flow 2 enters the network at node 1 and exits at node 3. The light-tailed flow 3 enters the network at node 1 and exits at node 4. Traffic flow 3 is delay unstable under the Back-Pressure policy if its arrival rate is sufficiently high.

Clearly, traffic flow 2 is delay unstable because it has to pass through the bottleneck link $(2,3)$. So, the remaining question concerns the delay stability of flow 3, which serves as cross-traffic to flow 2. The following results establish that flow 3 has a nontrivial *delay stability region*, and provide a sharp characterization of it.

Proposition 5: Consider the network of Figure 4 under the Back-Pressure policy. If $\lambda_3 > \big(2 + \lambda_1 - 2\lambda_2\big)/3$, then traffic flow 3 is delay unstable.

Proof: (Outline) The proof of this result is based on Theorem 2, i.e., we study the evolution of the fluid model of the network of Figure 4, from initial condition $q_{1,2}(0) = 1$ and zero for all other queues. We show that there exists $\tau > 0$ such that $q_{3,1}(\tau) > 0$, which implies the delay instability of the source queue of flow 3 and, thus, the delay instability of flow 3 itself. A detailed proof can be found in Appendix I. ■

Proposition 6: Consider the network of Figure 4 under the Back-Pressure policy. If $\lambda_3 < \big(2 + \lambda_1 - 2\lambda_2\big)/3$ and flows 2 and 3 are exponential-type, then traffic flow 3 is delay stable.

Proof: (Outline) The proof of this result relies on Theorem 3, i.e., we show that the function:

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

where

$$\begin{aligned} V\big(q(t)\big) &= \max\Big\{\big[q_{3,1}(t) - q_{3,2}(t)\big]^+, \big[q_{2,1}(t) - q_{2,2}(t)\big]^+\Big\} \\ &= \max\Big\{q_{3,1}(t), \big[q_{2,1}(t) - q_{2,2}(t)\big]^+\Big\}, \end{aligned}$$

and

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

is a Lyapunov function for the FM of the network of Figure 4. Note that this function is also continuous, piecewise linear, and has exponential-type “upward-jumps” in the stochastic domain, because flows 2 and 3 are assumed to be exponential-type. Thus, the steady-state length of queue $(3, 1)$ is exponential-type and, consequently, flow 3 is delay stable. A detailed proof can be found in Appendix II. ■

The above example illustrates that under the Back-Pressure policy, heavy-tailed flows may affect light-tailed flows directly, if the latter have to pass through bottleneck links, or indirectly, if they serve as cross-traffic to other light-tailed flows that have become delay unstable.

E. The Role of Intersecting Paths

Finally, consider the network of Figure 5: the heavy-tailed flow 1 enters the network at node 1 and exits the network as soon as it reaches node 5. Flow 1 is allowed to access all links, so packets can get to node 4 either through the path $\big((1, 2), (2, 4)\big)$ or through the path $\big((1, 3), (3, 4)\big)$. After they reach node 4, though, they have to pass through link $(4, 5)$ in order to reach their destination. In that sense, the two paths of flow 1 intersect.

Theorem 1 implies that queue $(1, 1)$ is delay unstable but provides no information regarding the other queues of flow 1, namely queues $(1, 2)$, $(1, 3)$, and $(1, 4)$. Since all links have finite capacities, the endogenous arrivals to those queues are, by definition, light-tailed. So, one might argue that these queues are delay stable. Somewhat surprisingly, we show that these queues are also delay unstable. This is due to the dynamics induced by the Back-Pressure policy, and the fact that multiple paths intersect. In particular, the queue at node 4 builds up, and this effect propagates backwards to cause the buildup of queues 2 and 3.

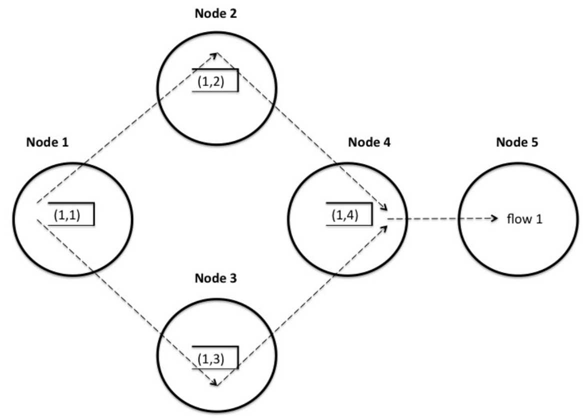

Fig. 5. The heavy-tailed flow 1 enters the network at node 1 and exits the network after it gets transmitted from node 4. Flow 1 is allowed to access all links in the network. Queues $(1, 2)$, $(1, 3)$, and $(1, 4)$ are delay unstable under the Back-Pressure policy because the two alternative paths of flow 1 intersect.

Proposition 7: Consider the network of Figure 5 under the Back-Pressure policy. All queues are delay unstable.

Proof: Again, we make use of Theorem 2 to simplify the proof of this result. More specifically, we consider the fluid model of the network of Figure 5 from initial condition $q_{1,1}(0) = 1$ and zero for all other queues. Since queues $(1, 2)$ and $(1, 3)$ cannot grow at rate higher than one (the capacity of the respective links), there exists $\tau > 0$ such that $q_{1,1}(t) > q_{1,2}(t)$ and $q_{1,1}(t) > q_{1,3}(t)$, for all $t \in [0, \tau]$. Eqs. (6) and (7) imply that

$$\dot{s}_{1,1,2}(t) = \dot{s}_{1,1,3}(t) = 1, \qquad \forall t \in [0, \tau].$$

On the other hand, $\dot{s}_{1,4,5}(t) = 1$, for all $t \in [0, \tau]$, so it is clear that on aggregate traffic is accumulating at rate one between queues $(1, 2)$, $(1, 3)$ and $(1, 4)$ during that interval. As a consequence of Eq. (7), it follows that traffic does not flow from node 2 (or 3) to 4 when the queue length at node 2 (or 3) exceeds that of node 4. As a result, the three queues grow at the same rate, i.e., $\dot{q}_{1,2}(t) = \dot{q}_{1,3}(t) = \dot{q}_{1,4}(t) = 1/3$, for all $t \in [0, \tau]$. Therefore, $q_{1,2}(\tau) = q_{1,3}(\tau) = q_{1,4}(\tau) = \tau/3$, so all three queues are delay unstable according to Theorem 2. ■

In contrast, if node 4 was the destination node of flow 1, then it is easy to show that queues $(1, 2)$ and $(1, 3)$ would have been delay stable. Thus, it is precisely the intersection of paths, combined with the dynamics imposed by Back-Pressure, that causes the delay instability.

Theorem 1 states that the traffic of heavy-tailed flows experiences large end-to-end delays overall, and definitely at the source queues. Whether these large delays are experienced only at the source queues, or at several other queues as well, is not as important from a practical standpoint. What is important, though, is the case of intersecting paths in networks with multiple flows. There, the delay unstable queues that are created by the intersecting paths may cause cross-traffic light-tailed flows to be delay unstable, similarly to the network of Figure 4. We conjecture that, again, the delay stability of cross-traffic flows depends on the exact values of the arrival rates.

V. The Back-Pressure-α Policy

The results and discussion presented above suggest that the Back-Pressure policy may perform poorly in the presence of heavy-tailed traffic. The reason is that by treating heavy-tailed and light-tailed flows "equally," there are long stretches of time during which the source queues of heavy-tailed flows dominate the service. This creates bottleneck links, which, in turn, may affect the delay stability of light-tailed flows directly or indirectly.

Intuitively, by discriminating against heavy-tailed flows, one should be able to eliminate bottlenecks and improve the overall performance of the network. One way to do this would be by giving preemptive priority to light-tailed flows. However, priority policies are undesirable because of fairness considerations, and also because they can be unstable in many network settings [10].

Motivated by the Max-Weight-α scheduling policy, studied in [15] in the context of single-hop networks, here we consider the Back-Pressure-α policy: instead of comparing the differential backlogs of the various flows, we compare the differential backlogs raised to different α-powers, smaller for heavy-tailed flows and larger for light-tailed flows. In that way we give partial priority to light-tailed flows.

More concretely, fix $\alpha_f > 0$, for every traffic flow $f \in \mathcal{F}$. Under the *Back-Pressure-α policy*, $S(t)$ is a feasible scheduling vector that maximizes the aggregate α-weighted Back-Pressure in the network, i.e.,

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

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

Before we state our main result regarding the Back-Pressure-α policy, we make an additional assumption: the set of links that flow f is allowed to access, $\mathcal{L}_f$, together with the associated nodes form a *Directed Acyclic Graph (DAG)* in which nodes s_f and d_f are the only source and sink nodes, respectively. While most of the proof of Theorem 5 goes through without it, the DAG assumption is required in the derivation of Eq. (23), which helps translate the admissibility of the arrivals into negative drift of the considered Lyapunov function.

Theorem 5: Consider the multi-hop switched queueing network of Section II with the additional DAG assumption, under the Back-Pressure-α policy. If $\mathbb{E}\big[A_f^{\alpha_f+1}(0)\big]$ is finite, for all $f \in \mathcal{F}$, then the network is stable and

$$\sum_{f\in\mathcal{F}} \sum_{i\in\mathcal{N}_f} \mathbb{E}\Big[Q_{f,i}^{\alpha_f}\Big] < \infty.$$

Proof: See Appendix III. ■

Corollary 1: (Delay Stability under Back-Pressure-α) Consider the multi-hop network of Section II with the additional DAG assumption, under the Back-Pressure-α policy. If the α-parameters of all light-tailed flows are equal to one, and the α-parameters of heavy-tailed flows are sufficiently small, then all light-tailed flows are delay stable.

Proof: We recall our standing assumption that all traffic flows have $(1+\gamma)$ moments, for some $\gamma > 0$. If the α-parameters of all light-tailed flows are equal to one, and the α-parameters of heavy-tailed flows are less then γ, then Theorem 5 and Little's Law imply that every queue of every light-tailed flow is delay stable. The linearity of expectations implies the delay stability of all light-tailed flows. ■

Combining Corollary 1 with Theorem 1, we conclude that *the Back-Pressure-α policy achieves the best possible performance with respect to the delay stability criterion, provided the α-parameters are suitably chosen.*

A special case of the Back-Pressure-α policy has been considered by Bui *et al.* [2], where all α-parameters take the same value. We note that their setting includes just light-tailed traffic and, additionally, the existence of congestion controllers. Thus, the insight that smaller parameter values should be used for heavy-tailed flows, so that light-tailed flows are given some form of priority, does not arise in their setting.

VI. Relationship to Prior Work

The present paper naturally builds upon and extends previous works that have analyzed single-hop networks with heavy-tailed traffic under Max-Weight-type policies [9], [15]. As a consequence, certain similarities with existing literature can be found at the technical level, e.g., sample path arguments are used to prove delay instability under Max-Weight and Back-Pressure; drift analysis of α-weighted Lyapunov functions is used to prove delay stability under Max-Weight-α and Back-Pressure-α. Moreover, some parallels can be drawn in terms of high-level insights, e.g., a light-tailed flow passing through a bottleneck link resembles, to some extent, the notion of conflict between heavy-tailed and light-tailed flows, cf. Theorem 2 of [15]; the fact that light-tailed flows must be given some form of priority over heavy-tailed traffic is the main reason that both the Back-Pressure-α policy and the Max-Weight-α policy perform well. However, the concrete insights that we derived regarding the impact of network topology, routing constraints, and link capacities on the delay performance of Back-Pressure policies, and the corresponding "network design guidelines," are meaningful only in a multi-hop setting and, thus, novel compared to prior work on single-hop networks.

In terms of methodology, by employing the advanced machinery of fluid approximations for delay stability analysis developed in the companion paper [16], we are able to obtain results that would have been difficult to prove through the direct stochastic analysis adopted in prior works, e.g., Propositions 5 and 6. Moreover, the drift analysis of piecewise linear Lyapunov functions, such as the one we introduce in the proof of Proposition 6, provides a systematic way for the delay analysis of the Back-Pressure policy in networks with a mix of heavy-tailed and exponential-type traffic.

In this work we have attempted to illustrate the behavior of multi-hop networks with heavy-tailed traffic under Back-Pressure policies in the clearest and most concrete way. Thus, we focused on phenomena and insights whose origin is precisely the multi-hop nature of the network and the corresponding flow-scheduling constraints. We believe the behavior of more complex network models that include a combination of flow-scheduling and link-scheduling constraints, can be understood in terms of the insights derived here as well as in previous papers. For example, Theorem 2 of [15] implies that if a link conflicts with a bottleneck link, then it becomes a bottleneck link itself. As another example, we expect that the Back-Pressure-α policy performs well even in multi-hop networks with link-scheduling constraints, at least for arrival rates that can be stably supported by these networks.

VII. Conclusion

The main objective of this paper was to obtain insights on the delay performance of multi-hop networks with heavy-tailed traffic under the widely studied class of Back-Pressure policies. Our analysis highlighted the significance of "bottleneck links," i.e., links that are allowed to serve the source queues of heavy-tailed traffic flows. The fundamental insight was that traffic flows that have to pass through bottleneck links experience large delays under Back-Pressure. We then investigated reasons that may force a light-tailed flow to pass through a bottleneck link, identifying the following: (i) the network topology, i.e., the source-destination paths that the network offers to the given flow; (ii) the routing constraints, i.e., the a priori decisions regarding which links the particular flow is allowed to traverse; (iii) the link capacities relative to the arrival rates, i.e., whether the combined capacity of non-bottleneck paths is sufficient to support the arrival rate of the flow.

These insights can be interpreted as rough "network design guidelines." For example, in a multi-hop network under the Back-Pressure policy, heavy-tailed flows should be relatively constrained in terms of the links that they are allowed to access, whereas the network should provide multiple source-destination paths to light-tailed flows; the latter flows should be left unconstrained to dynamically find their way around heavy-tailed traffic. Moreover, these alternate paths should have enough capacity to support the rates of light-tailed traffic. In contrast, leaving heavy-tailed flows unconstrained while forcing light-tailed flows to compete with them could be detrimental to the overall delay performance of the network.

An alternative way to achieve good delay performance in a multi-hop network with heavy-tailed traffic is through the parameterized Back-Pressure-α policy. We showed that this policy can delay stabilize all light-tailed flows in the network, provided that its α-parameters are chosen suitably. In order to pick appropriate parameter values, though, some knowledge of higher order moments of the different traffic flows is required.

The results of this paper were consistently presented in terms of delay stability, a rather crude performance metric that attempts to capture the notion of large delays in a binary manner. However, many of them can be significantly refined. For instance, if we generalize the notion of a heavy-tailed flow to be one that has infinite $(k+1)$st moment of arrivals, for some $k \in \mathbb{N}$, then any light-tailed flow that has to pass through the bottleneck links of a heavy-tailed flow has infinite k^{th} moment of steady-state aggregate queue length under the Back-Pressure policy; this can be established through a straightforward extension of Theorem 2. Under certain regularity assumptions, this approach could also give lower bounds on queue length asymptotics. Moreover, regarding networks with a mix of heavy-tailed and exponential-type traffic, delay stability can be proved via drift analysis of piecewise linear Lyapunov functions. As Theorem 3 suggests, this type of analysis guarantees not only the delay stability of light-tailed flows, but also exponential upper bounds on the respective steady-state queue-length asymptotics; see Theorem 2.3 in [7].

Finally, while strictly speaking the existence of heavy-tailed traffic, i.e., arrival processes with infinite variance, could be a subject of debate, our results can be directly interpreted in the context of a network with a mix of bursty and non-bursty traffic, a setting that is prevalent in data communication networks.

Appendix I - Proof of Proposition 5

The proof of this result is based on Theorem 2, i.e., we study the evolution of the fluid model of the network of Figure 4, from initial condition $q_{1,2}(0) = 1$ and zero for all other queues. We distinguish between two phases in the evolution of the (fluid) system.

In the first phase, the length of queue $(1,2)$ is greater than the length of queue $(2,2)$ due to the initial conditions, so Eq. (7) implies that the service capacity of link $(2,3)$ is allocated solely to flow 1. Moreover, link $(2,4)$ transmits only traffic of flow 3, which results in queue $(3,2)$ being always empty. Eq. (7) implies that Back-Pressure splits the service capacity of link $(1,2)$ in such a way, so that the differential backlogs of flows 2 and 3 over that link remain the same, i.e., zero. Consequently, queues $(2,1)$ and $(2,2)$ build up together and at a constant rate throughout this phase. In mathematical terms, Eq. (4) implies that $\dot{q}_{1,2}(t) < 0$ whereas $\dot{q}_{2,2} = \dot{q}_{2,1}(t) > 0$. Therefore, there exists $\tau' > 0$ such that $q_{1,2}(\tau') = q_{2,2}(\tau') = q_{2,1}(\tau') > 0$.

In the second phase, the differential backlog of flows 1 and 2 over link $(2,3)$ is the same, so Eq. (7) implies that this link serves from both flows simultaneously until one of the two queues empties. Throughout this interval, the arrival and service rates need to satisfy the following linear system:

$$\big(\lambda_2 - \dot{s}_{2,1,2}(t)\big) - \big(\dot{s}_{2,1,2}(t) - \dot{s}_{2,2,3}(t)\big) = \lambda_3 - \dot{s}_{3,1,2}(t), \quad (11)$$

$$\lambda_1 - \dot{s}_{1,2,3}(t) = \dot{s}_{2,1,2}(t) - \dot{s}_{2,2,3}(t), \quad (12)$$

$$\dot{s}_{2,1,2}(t) + \dot{s}_{3,1,2}(t) = 1, \quad (13)$$

$$\dot{s}_{1,2,3}(t) + \dot{s}_{2,2,3}(t) = 1. \quad (14)$$

Eq. (11) follows from Eq. (7), and is due to the fact that the Back-Pressure policy tries to keep the differential backlogs of flows 2 and 3 over link $(1,2)$ the same. We note that queue $(3,2)$ remains zero throughout both phases, so that the rate of change of its length is also zero. Eq. (12) follows from a similar argument for link $(2,3)$. Eqs. (13) and (14) result from Eq. (6).

The above equations and some simple algebra imply that

$$\dot{s}_{3,1,2}(t) = \frac{2+\lambda_1 - 2\lambda_2 + 2\lambda_3}{5}.$$

Therefore,

$$\lambda_3 > \dot{s}_{3,1,2}(t) \iff \lambda_3 > \frac{2+\lambda_1 - 2\lambda_2}{3}.$$

Finally, notice that the duration of the second phase is bounded away from zero, since the queue lengths $q_{1,2}(\tau')$ and $q_{2,2}(\tau')$ are also bounded away from zero. Therefore, if $\lambda_3 > \big(2+\lambda_1-2\lambda_2\big)/3$, then there exists $\tau > \tau'$ such that $\dot{q}_{3,1}(t) > 0$, for all $t \in [\tau', \tau]$, because of Eq. (4). Therefore, $q_{3,1}(\tau) > 0$, which implies that queue $(3,1)$ is delay unstable according to Theorem 2. Flow 3 is, thus, delay unstable since the delay experienced in the source queue bounds from below the end-to-end delay.

Appendix II - Proof of Proposition 6

Consider the candidate Lyapunov function for the FM:

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

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

$$\begin{aligned} V\big(q(t)\big) &= \max\Big\{\big[q_{3,1}(t) - q_{3,2}(t)\big]^+, \big[q_{2,1}(t) - q_{2,2}(t)\big]^+\Big\} \\ &= \max\Big\{q_{3,1}(t), \big[q_{2,1}(t) - q_{2,2}(t)\big]^+\Big\}. \end{aligned}$$

Note that this function is continuous, piecewise linear, and has exponential-type "upward-jumps" in the stochastic domain, because flows 2 and 3 are assumed to be exponential-type. Thus, by means of Theorem 3, it suffices to show that $H(\cdot)$ is, indeed, a Lyapunov function for the FM.

Our proof strategy is as follows. First, we distinguish cases regarding the derivatives of the two terms of $H(\cdot)$: cases (a), (b), and (c) pertaining to term $V(\cdot)$, cases (1), (2), and (3) pertaining to term $G(\cdot)$. Then, we combine these cases to compute the derivative of $H(\cdot)$ in the different regions of the state space.

Case (a): if $q_{3,1}(t) > \big[q_{2,1}(t) - q_{2,2}(t)\big]^+$ then $V\big(q(t)\big) = q_{3,1}(t) > 0$, which implies that

$$\dot{V}\big(q(t)\big) = \lambda_3 - 1 < 0;$$

Case (b): if $q_{3,1}(t) < \big[q_{2,1}(t) - q_{2,2}(t)\big]^+$ then $V\big(q(t)\big) = q_{2,1}(t) - q_{2,2}(t) > 0$, which implies that

$$\dot{V}\big(q(t)\big) = (\lambda_2 - 1) - \big(1 - \dot{s}_{2,2,3}(t)\big) = \lambda_2 + \dot{s}_{2,2,3}(t) - 2 < 0;$$

Case (c): if $q_{3,1}(t) = \big[q_{2,1}(t) - q_{2,2}(t)\big]^+$ then $V\big(q(t)\big) = q_{3,1}(t) = q_{2,1}(t) - q_{2,2}(t)$, which implies that

$$\begin{aligned} \dot{V}\big(q(t)\big) &= \lambda_3 - \dot{s}_{3,1,2}(t) \\ &= \big(\lambda_2 - \dot{s}_{2,1,2}(t)\big) - \big(\dot{s}_{2,1,2}(t) - \dot{s}_{2,2,3}(t)\big), \end{aligned}$$

where the service rates satisfy Eq. (11);

Case (1): if $q_{2,2}(t) > q_{1,2}(t)$ then $G\big(q(t)\big) = q_{2,2}(t) - q_{1,2}(t) > 0$, which implies that

$$\dot{G}\big(q(t)\big) = \dot{s}_{2,1,2}(t) - 1 - \lambda_1 < 0;$$

Case (2): if $q_{2,2}(t) < q_{1,2}(t)$ then $G\big(q(t)\big) = 0$, and also

$$\dot{G}\big(q(t)\big) = 0;$$

Case (3): if $q_{2,2}(t) = q_{1,2}(t)$ then $G\big(q(t)\big) = 0$, but now

$$\dot{G}\big(q(t)\big) = \big(\dot{s}_{2,1,2}(t) - \dot{s}_{2,2,3}(t)\big) - \big(\lambda_1 - \dot{s}_{1,2,3}(t)\big) = 0,$$

according to Eq. (12)

Now, in order to show that $H(\cdot)$ is a Lyapunov function for the FM, we have to show that its derivative is negative and bounded away from zero in all nine regions of the state space, whenever $H(\cdot)$ is greater than zero. We note that the stability conditions in this example translate to $\lambda_1 + \lambda_2 < 1$ and $\lambda_2 + \lambda_3 < 1$.

Region (a,1):

$$\dot{G}\big(q(t)\big) = \lambda_3 - 1 + \dot{s}_{2,1,2}(t) - 1 - \lambda_1 = \lambda_3 - \lambda_1 - 2 < 0;$$

Region (a,2): $\dot{G}\big(q(t)\big) = \lambda_3 - 1 < 0;$

Region (a,3): $\dot{G}\big(q(t)\big) = \lambda_3 - 1 < 0;$

Region (b,1):

$$\dot{G}\big(q(t)\big) = \lambda_2 + \dot{s}_{2,2,3}(t) - 2 + \dot{s}_{2,1,2}(t) - 1 - \lambda_1 < 0;$$

Region (b,2): $\dot{G}\big(q(t)\big) = \lambda_2 + \dot{s}_{2,2,3}(t) - 2 < 0;$

Region (b,3): $\dot{G}\big(q(t)\big) = \lambda_2 + \dot{s}_{2,2,3}(t) - 2 < 0;$

Region (c,1): $\dot{G}\big(q(t)\big) = \lambda_2 - \lambda_1 - \dot{s}_{2,1,2}(t).$

We note that in this region we have that $q_{3,1}(t) = \big[q_{2,1}(t) - q_{2,2}(t)\big]^+ > 0$. Thus, $\dot{s}_{2,1,2}(t)$ satisfies Eqs. (11) and (13) with $\dot{s}_{2,2,3}(t) = 1$. The solution to this linear system gives $\dot{s}_{2,1,2}(t) = (2+\lambda_2-\lambda_3)/3$, which implies that

$$\begin{aligned} \dot{G}\big(q(t)\big) &= \lambda_2 - \lambda_1 - \frac{2+\lambda_2-\lambda_3}{3} \\ &= \frac{-2-3\lambda_1+2\lambda_2+\lambda_3}{3} \\ &< 0; \end{aligned}$$

Region (c,2): $\dot{G}\big(q(t)\big) = \lambda_2 - 2\dot{s}_{2,1,2}(t).$

Here, we have to assume that $q_{3,1}(t) = \big[q_{2,1}(t) - q_{2,2}(t)\big]^+ > 0$, otherwise $H(\cdot)$ would have zero value. Thus, $\dot{s}_{2,1,2}(t)$ satisfies Eqs. (11) and (13) with $\dot{s}_{2,2,3}(t) = 0$. The solution to this linear system gives $\dot{s}_{2,1,2}(t) = (1+\lambda_2-\lambda_3)/3$, which implies that

$$\dot{G}\big(q(t)\big) = \lambda_2 - \frac{2+2\lambda_2-2\lambda_3}{3} = \frac{-2+\lambda_2+2\lambda_3}{3} < 0;$$

Region (c,3): $\dot{G}\big(q(t)\big) = \lambda_3 - \dot{s}_{3,1,2}(t)$.

Similarly to the previous case, we have to assume that $q_{3,1}(t) = \big[q_{2,1}(t) - q_{2,2}(t)\big]^+ > 0$, otherwise $H(\cdot)$ would have zero value. Thus, $\dot{s}_{3,1,2}(t)$ satisfies Eqs. (11)-(14). The solution to this linear system, as we saw in Proposition 6, gives $\dot{s}_{3,1,2}(t) = (2+\lambda_1 - 2\lambda_2 + 2\lambda_3)/5$, which implies that

$$\dot{G}\big(q(t)\big) < 0,$$

precisely when $\lambda_3 < (2+\lambda_1-2\lambda_2)/3$.

Appendix III - Proof of Theorem 5

Under the dynamics induced by the Back-Pressure-α policy the sequence $\big\{Q(t);\ t \in \mathbb{Z}_+\big\}$ is a time-homogeneous, irreducible, and aperiodic Markov chain on a countable state space. We will show that this Markov chain is also positive recurrent, and we will obtain moment bounds on the steady-state queue lengths, through drift analysis of the candidate Lyapunov function

$$V\big(Q(t)\big) = \sum_{f\in\mathcal{F}} \sum_{i\in\mathcal{N}_f} \frac{1}{\alpha_f+1} Q_{f,i}^{\alpha_f+1}(t).$$

Throughout the proof we use the shorthand notation

$$T_{f,i}(t) = \sum_{j:(i,j)\in\mathcal{L}_f} S_{f,i,j}(t)$$

for the departures from queue (f,i), and

$$R_{f,i}(t) = \sum_{j:(j,i)\in\mathcal{L}_f} S_{f,j,i}(t) + A_f(t)\cdot 1_{\{i=s_f\}}$$

for the arrivals at queue (f,i), at time slot t.

Moreover, we let

$$V_{f,i}\big(Q(t)\big) = \frac{1}{\alpha_f+1} Q_{f,i}^{\alpha_f+1}(t).$$

The Lyapunov function can be written in the form

$$V\big(Q(t)\big) = \sum_{f\in\mathcal{F}} \sum_{i\in\mathcal{N}_f} V_{f,i}\big(Q(t)\big),$$

which implies that

$$\begin{aligned}&\mathbb{E}\big[V\big(Q(t+1)\big) - V\big(Q(t)\big) \mid \mathcal{F}_t\big]\\ &\quad= \sum_{f\in\mathcal{F}} \sum_{i\in\mathcal{N}_f} \mathbb{E}\big[V_{f,i}\big(Q(t+1)\big) - V_{f,i}\big(Q(t)\big) \mid \mathcal{F}_t\big]. \end{aligned} \quad (15)$$

We will perform the drift analysis of function $V(\cdot)$ by upper-bounding the terms on the right-hand side of Eq. (15).

Using the notation above and the dynamics of the multi-hop network, we have that

$$\mathbb{E}\big[V_{f,i}\big(Q(t+1)\big) \mid \mathcal{F}_t\big] = \frac{1}{\alpha_f+1}\big(Q_{f,i}(t) + \Delta_{f,i}(t)\big)^{\alpha_f+1}, \quad (16)$$

where $\Delta_{f,i}(t) = R_{f,i}(t) - T_{f,i}(t)$. Now we distinguish between three cases:

(i) if $i = d_f$ then Eq. (3) implies that

$$\mathbb{E}\big[V_{f,d_f}\big(Q(t+1)\big) - V_{f,d_f}\big(Q(t)\big) \mid \mathcal{F}_t\big] = 0; \quad (17)$$

(ii) if $i \neq d_f$ **and** $\alpha_f < 1$**,** then we consider the zeroth order Taylor expansion of the right-hand side of Eq. (16) around $Q_{f,i}(t)$:

$$\begin{aligned}&\frac{1}{\alpha_f+1}\big(Q_{f,i}(t)+\Delta_{f,i}(t)\big)^{\alpha_f+1}\\ &\qquad= \frac{1}{\alpha_f+1} Q_{f,i}^{\alpha_f+1}(t) + \Delta_{f,i}(t)\cdot\xi^{\alpha_f},\end{aligned}$$

which implies that

$$\mathbb{E}\big[V_{f,i}\big(Q(t+1)\big) \mid \mathcal{F}_t\big] \le V_{f,i}\big(Q(t)\big) + \mathbb{E}\big[\Delta_{f,i}(t)\cdot\xi^{\alpha_f} \mid \mathcal{F}_t\big],$$

for some $\xi \in \big[Q_{f,i}(t) - T_{f,i}(t), Q_{f,i}(t) + R_{f,i}(t)\big]$.

Consider the event $\Gamma_{f,i}(t) = \big\{\Delta_{f,i}(t) \le 0\big\}$ and its complement. The expression above can be written in the form

$$\begin{aligned}&\mathbb{E}\big[V_{f,i}\big(Q(t+1)\big) \mid \mathcal{F}_t\big]\\ &\le V_{f,i}\big(Q(t)\big) + \mathbb{E}\Big[\Delta_{f,i}(t)\big(Q_{f,i}(t) - T_{f,i}(t)\big)^{\alpha_f};\ \Gamma_{f,i}(t) \ \Big|\ \mathcal{F}_t\Big]\\ &\quad + \mathbb{E}\Big[\Delta_{f,i}(t)\big(Q_{f,i}(t) + R_{f,i}(t)\big)^{\alpha_f};\ \Gamma^c_{f,i}(t) \ \Big|\ \mathcal{F}_t\Big].\end{aligned}$$

Note that $Q_{f,i}(t)$, $R_{f,i}(t)$, and $T_{f,i}(t)$ are nonnegative integers, $T_{f,i}(t) \le Q_{f,i}(t)$, and $T_{f,i}(t) \le d_{\max}$, where $d_{\max}$ is the maximum number of outgoing edges of any node in $\mathcal{G}$. It can be verified that

$$\big(Q_{f,i}(t) + R_{f,i}(t)\big)^{\alpha_f} \le Q_{f,i}^{\alpha_f}(t) + R_{f,i}^{\alpha_f}(t),$$

and

$$\big(Q_{f,i}(t) - T_{f,i}(t)\big)^{\alpha_f} \ge Q_{f,i}^{\alpha_f}(t) - d_{\max}^{\alpha_f}.$$

Using these inequalities we can write

$$\begin{aligned}&\mathbb{E}\big[V_{f,i}\big(Q(t+1)\big) \mid \mathcal{F}_t\big]\\ &\qquad\le V_{f,i}\big(Q(t)\big) + \mathbb{E}\big[\Delta_{f,i}(t) \mid \mathcal{F}_t\big]\cdot Q_{f,i}^{\alpha_f}(t)\\ &\qquad\quad - \mathbb{E}\Big[\Delta_{f,i}(t)\cdot T_{f,i}^{\alpha_f}(t);\ \Gamma_{f,i}(t) \ \Big|\ \mathcal{F}_t\Big]\\ &\qquad\quad + \mathbb{E}\Big[\Delta_{f,i}(t)\cdot R_{f,i}^{\alpha_f}(t);\ \Gamma^c_{f,i}(t) \ \Big|\ \mathcal{F}_t\Big],\end{aligned}$$

which implies that

$$\begin{aligned}&\mathbb{E}\big[V_{f,i}\big(Q(t+1)\big) \mid \mathcal{F}_t\big]\\ &\qquad\le V_{f,i}\big(Q(t)\big) + \mathbb{E}\big[\Delta_{f,i}(t) \mid \mathcal{F}_t\big]\cdot Q_{f,i}^{\alpha_f}(t) + d_{\max}^{\alpha_f+1}\\ &\qquad\quad + \mathbb{E}\Big[R_{f,i}^{\alpha_f+1}(t);\ \Gamma^c_{f,i}(t) \ \Big|\ \mathcal{F}_t\Big].\end{aligned}$$

Since $\mathbb{E}\Big[A_f^{\alpha_f+1}(0)\Big]$ is finite and all arrivals processes are mutually independent and IID over time slots, $\mathbb{E}\Big[R_{f,i}^{\alpha_f+1}(t);\ \Gamma^c_{f,i}(t) \ \Big|\ \mathcal{F}_t\Big]$ is finite. Thus, there exists a finite constant $c_{f,i}$ such that

$$\begin{aligned}&\mathbb{E}\big[V_{f,i}\big(Q(t+1)\big) - V_{f,i}\big(Q(t)\big) \mid \mathcal{F}_t\big]\\ &\qquad\le \mathbb{E}\big[\Delta_{f,i}(t) \mid \mathcal{F}_t\big]\cdot Q_{f,i}^{\alpha_f}(t) + c_{f,i};\end{aligned} \quad (18)$$

(iii) if $i \neq d_f$ **and** $\alpha_f \ge 1$**,** then we consider the first order Taylor expansion of the right-hand side of Eq. (16) around

$Q_{f,i}(t)$:

$$\begin{aligned}&\frac{1}{\alpha_f+1}\big(Q_{f,i}(t)+\Delta_{f,i}(t)\big)^{\alpha_f+1}\\=&\frac{1}{\alpha_f+1}Q_{f,i}^{\alpha_f+1}(t)+\Delta_{f,i}(t)Q_{f,i}^{\alpha_f}(t)+\frac{\Delta_{f,i}^2(t)}{2}\cdot\alpha_f\cdot\xi^{\alpha_f-1}\end{aligned}$$

which implies that

$$\begin{aligned}\mathbb{E}\big[V_{f,i}\big(Q(t+1)\big)\big|\mathcal{F}_t\big]\leq&V_{f,i}\big(Q(t)\big)+\mathbb{E}\big[\Delta_{f,i}(t)\big|\mathcal{F}_t\big]Q_{f,i}^{\alpha_f}(t)\\&+\frac{1}{2}\mathbb{E}\Big[\Delta_{f,i}^2(t)\cdot\alpha_f\cdot\xi^{\alpha_f-1}\ \Big|\ \mathcal{F}_t\Big],\end{aligned}$$

for some $\xi\in\big[Q_{f,i}(t)-T_{f,i}(t),Q_{f,i}(t)+R_{f,i}(t)\big]$.

Since $\alpha_f\geq 1$, the last term can be bounded from above as follows:

$$\begin{aligned}&\frac{1}{2}\mathbb{E}\Big[\Delta_{f,i}^2(t)\cdot\alpha_f\cdot\xi^{\alpha_f-1}\ \Big|\ \mathcal{F}_t\Big]\\&\qquad\leq\frac{1}{2}\mathbb{E}\Big[\Delta_{f,i}^2(t)\alpha_f\big(Q_{f,i}(t)+R_{f,i}(t)\big)^{\alpha_f-1}\ \Big|\ \mathcal{F}_t\Big].\end{aligned}$$

It can be verified that

$$\big(Q_{f,i}(t)+R_{f,i}(t)\big)^{\alpha_f-1}\leq 2^{\alpha_f-1}\Big(Q_{f,i}^{\alpha_f-1}(t)+R_{f,i}^{\alpha_f-1}(t)\Big),$$

and

$$\Delta_{f,i}^2(t)\leq R_{f,i}^2(t)+d_{\max}^2.$$

Using these inequalities we can write

$$\begin{aligned}&\frac{1}{2}\mathbb{E}\Big[\Delta_{f,i}^2(t)\cdot\alpha_f\cdot\xi^{\alpha_f-1}\ \Big|\ \mathcal{F}_t\Big]\\&\qquad\leq 2^{\alpha_f-2}\cdot\alpha_f\cdot\Big(\mathbb{E}\Big[R_{f,i}^2(t)\ \Big|\ \mathcal{F}_t\Big]+d_{\max}^2\Big)Q_{f,i}^{\alpha_f-1}\\&\qquad\quad+2^{\alpha_f-2}\cdot\alpha_f\cdot\Big(\mathbb{E}\Big[R_{f,i}^{\alpha_f+1}(t)\ \Big|\ \mathcal{F}_t\Big]\\&\qquad\quad+d_{\max}^2\cdot\mathbb{E}\Big[R_{f,i}^{\alpha_f-1}(t)\ \Big|\ \mathcal{F}_t\Big]\Big).\end{aligned}$$

Thus, for every $y_{f,i}>0$ there exists a constant $c_{f,i}(y_{f,i})$ such that

$$\frac{1}{2}\mathbb{E}\Big[\Delta_{f,i}^2(t)\cdot\alpha_f\cdot\xi^{\alpha_f-1}\ \Big|\ \mathcal{F}_t\Big]\leq y_{f,n}\cdot Q_{f,i}^{\alpha_f}(t)+c_{f,i}(y_{f,i}).$$

Consequently,

$$\begin{aligned}&\mathbb{E}\big[V_{f,i}\big(Q(t+1)\big)-V_{f,i}\big(Q(t)\big)\ \big|\ \mathcal{F}_t\big]\\\leq&\mathbb{E}\big[\Delta_{f,i}(t)\ \big|\ \mathcal{F}_t\big]\cdot Q_{f,i}^{\alpha_f}(t)+y_{f,n}\cdot Q_{f,i}^{\alpha_f}(t)+c_{f,i}(y_{f,i}).\end{aligned}\tag{19}$$

Eqs. (15), (17), (18), and (19) imply that, for every $\delta>0$ there exist constants $c_{f,i}(\delta)$, $i\in\mathcal{N}_f$, $f\in\mathcal{F}$, such that

$$\begin{aligned}&\mathbb{E}\big[V\big(Q(t{+}1)\big)-V\big(Q(t)\big)\ \big|\ \mathcal{F}_t\big]\\&\quad\leq-\mathbb{E}\Big[\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f}Q_{f,i}^{\alpha_f}(t)\\&\qquad\cdot\Big(\sum_{j:(i,j)\in\mathcal{L}_f}S_{f,i,j}(t)-\sum_{j:(j,i)\in\mathcal{L}_f}S_{f,j,i}(t)\Big)\ \Big|\ \mathcal{F}_t\Big]\\&\qquad+\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f}Q_{f,i}^{\alpha_f}(t)\cdot\mathbb{E}\big[A_f(t);\ i=s_f\ \big|\ \mathcal{F}_t\big]\\&\qquad+\delta\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f}Q_{f,i}^{\alpha_f}(t)+\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f}c_{f,i}(\delta).\end{aligned}\tag{20}$$

Now, for notational convenience, define the quantity

$$W_{i,j}(t)=\max_{f:(i,j)\in\mathcal{L}_f}\Big[Q_{f,i}^{\alpha_f}(t)-Q_{f,j}^{\alpha_f}(t)\Big].$$

Through simple algebra, and using the fact that under our policy $S_{f,i,j}(t)$ is set to 1 for some f that attains the maximum in the definition of $W_{i,j}(t)$, we have that

$$\begin{aligned}&\mathbb{E}\Big[\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f}Q_{f,i}^{\alpha_f}(t)\\&\qquad\cdot\Big(\sum_{j:(i,j)\in\mathcal{L}_f}S_{f,i,j}(t)-\sum_{j:(j,i)\in\mathcal{L}_f}S_{f,j,i}(t)\Big)\ \Big|\ \mathcal{F}_t\Big]\\&\quad=\mathbb{E}\Big[\sum_{f\in\mathcal{F}}\sum_{(i,j)\in\mathcal{L}_f}\Big(Q_{f,i}^{\alpha_f}(t)-Q_{f,j}^{\alpha_f}(t)\Big)S_{f,i,j}(t)\ \Big|\ \mathcal{F}_t\Big]\\&\quad=\sum_{(i,j)\in\mathcal{L}}W_{i,j}(t).\end{aligned}\tag{21}$$

On the other hand,

$$\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f}Q_{f,i}^{\alpha_f}(t)\cdot\mathbb{E}\big[A_f(t);\ i=s_f\ \big|\ \mathcal{F}_t\big]=\sum_{f\in\mathcal{F}}\lambda_fQ_{f,s_f}^{\alpha_f}(t).\tag{22}$$

Let $\mathcal{P}_f$ be the set of distinct source-destination paths of traffic flow $f\in\mathcal{F}$. The fact that the arrival rate vector λ is in the stability region of the network implies the existence of constants $\epsilon>0$ and $\zeta_{f,p}\geq 0$, for $p\in\mathcal{P}_f$, $f\in\mathcal{F}$, such that

$$\lambda_f=\sum_{p\in\mathcal{P}_f}\zeta_{f,p},\qquad\forall f\in\mathcal{F},$$

$$\zeta_{f,i,j}=\sum_{p:(i,j)\in p}\zeta_{f,p},\qquad\forall(i,j)\in\mathcal{L}_f,\qquad\forall f\in\mathcal{F},$$

and

$$\sum_{f:(i,j)\in\mathcal{L}_f}\zeta_{f,i,j}\leq 1-\epsilon,\qquad\forall(i,j)\in\mathcal{L}.$$

Thus,

$$\begin{aligned}\sum_{f\in\mathcal{F}}\lambda_fQ_{f,s_f}^{\alpha_f}(t)&=\sum_{f\in\mathcal{F}}\sum_{p\in\mathcal{P}_f}\zeta_{f,p}Q_{f,s_f}^{\alpha_f}(t)\\&=\sum_{f\in\mathcal{F}}\sum_{p\in\mathcal{P}_f}\sum_{(i,j)\in p}\zeta_{f,p}\Big(Q_{f,i}^{\alpha_f}(t)-Q_{f,j}^{\alpha_f}(t)\Big),\\&\leq\sum_{f\in\mathcal{F}}\sum_{(i,j)\in\mathcal{L}_f}\zeta_{f,i,j}W_{i,j}(t)\\&=\sum_{(i,j)\in\mathcal{L}}\sum_{f:(i,j)\in\mathcal{L}_f}\zeta_{f,i,j}W_{i,j}(t)\\&\leq(1-\epsilon)\sum_{(i,j)\in\mathcal{L}}W_{i,j}(t).\end{aligned}\tag{23}$$

Eqs. (20)-(23) imply that

$$\begin{aligned}&\mathbb{E}\big[V\big(Q(t+1)\big)-V\big(Q(t)\big)\ \big|\ \mathcal{F}_t\big]\\&\qquad\leq-\epsilon\sum_{(i,j)\in\mathcal{L}}W_{i,j}(t)+\delta\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f}Q_{f,i}^{\alpha_f}(t)\\&\qquad\quad+\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f}c_{f,i}(\delta).\end{aligned}$$

Finally, since $Q_{f,d_f}(t)=0$, the sum of the α-weighted differential backlogs along any (directed and acyclic) source-destination path in $\mathcal{L}_f$ upper bounds the α_f-power of each of the queue lengths of flow f along that path. Hence, it can be verified that there exists a large enough $\epsilon' > 0$ such that

$$\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} Q_{f,i}^{\alpha_f}(t) \le \epsilon' \sum_{(i,j)\in\mathcal{L}} W_{i,j}(t).$$

If δ is chosen sufficiently small, there exist constants $\gamma > 0$ and $\beta < \infty$ such that

$$\mathbb{E}\big[V\big(Q(t+1)\big) - V\big(Q(t)\big) \mid \mathcal{F}_t\big] \le -\gamma \sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f} Q_{f,i}^{\alpha_f}(t) + \beta.$$

Then, the Foster-Lyapunov stability criterion and moment bound (e.g., see Corollary 2.1.5 of [8]) implies that the queueing network is stable and that $\sum_{f\in\mathcal{F}}\sum_{i\in\mathcal{N}_f}\mathbb{E}\Big[Q_{f,i}^{\alpha_f}\Big]$ is finite.

References

[1] O. Boxma, B. Zwart (2007). Tails in scheduling. *Performance Evaluation Review*, 34(4), pp. 13-20.

[2] L. Bui, S. Sanghavi, R. Srikant (2009). Distributed link scheduling with constant overhead. *IEEE/ACM Transactions on Networking*, 17(5), pp. 1467-1480.

[3] L. Bui, R. Srikant, A. Stolyar (2011). A novel architecture for reduction of delay and queueing structure complexity in the Back-Pressure algorithm. *IEEE/ACM Transactions on Networking*, 19(6), pp. 1597-1609.

[4] J.G. Dai, W. Lin (2005). Maximum pressure policies in stochastic processing networks. *Operations Research*, 53(2), pp. 197-218.

[5] A. Eryilmaz, R. Srikant (2006). Joint congestion control, routing and MAC for stability and fairness in wireless networks. *IEEE Journal on Selected Areas in Communications*, 24(8), pp. 1514-1524.

[6] L. Georgiadis, M. Neely, L. Tassiulas (2006). Resource allocation and cross-layer control in wireless nertworks. *Foundations and Trends in Networking*, 1(1), pp. 1-144.

[7] B. Hajek (1982). Hitting-time and occupation-time bounds implied by drift analysis with applications. *Advances in Applied Probability*, 14, pp. 502-525.

[8] B. Hajek (2006). Notes on communication network analysis. Available at: http://www.ifp.illinois.edu/∼hajek.

[9] K. Jagannathan, M.G. Markakis, E. Modiano, J.N. Tsitsiklis (2012). Queue-length asymptotics for generalized Max-Weight scheduling in the presence of heavy-tailed traffic. *IEEE/ACM Transactions on Networking*, 20(4), pp. 1096-1111.

[10] K. Jagannathan, M.G. Markakis, E. Modiano, J.N. Tsitsiklis (2014). Throughput optimal scheduling over time-varying channels in the presence of heavy-tailed traffic. *IEEE Transactions on Information Theory*, 60(5), pp. 2896-2909.

[11] B. Ji, C. Joo, N.B. Shroff (2013). Throughput-optimal scheduling in multihop wireless networks without per-flow information. *IEEE/ACM Transactions on Networking*, 21(2), pp. 634-647.

[12] B. Ji, C. Joo, N.B. Shroff (2013). Delay-based Back-Pressure scheduling in multihop wireless networks. *IEEE/ACM Transactions on Networking*, 21(5), pp. 1539-1552.

[13] W. Leland, M. Taqqu, W. Willinger, D. Wilson (1994). On the self-similar nature of ethernet traffic. *IEEE/ACM Transactions on Networking*, 2(1), pp. 1-15.

[14] J. Liu, A. Stolyar, M. Chiang, H.V. Poor (2009). Queue Back-Pressure random access in multi-hop wireless networks: optimality and stability. *IEEE Transactions on Information Theory*, 55(9), pp. 4087-4098.

[15] M.G. Markakis, E. Modiano, J.N. Tsitsiklis (2014). Max-Weight scheduling in queueing networks with heavy-tailed traffic. *IEEE/ACM Transactions on Networking*, 22(1), pp. 257-270.

[16] M.G. Markakis, E. Modiano, J.N. Tsitsiklis (preprint). Delay analysis of the Max-Weight policy under heavy-tailed traffic via fluid approximations. Available at: http://www.econ.upf.edu/∼mmarkakis.

[17] J. Nair, K. Jagannathan, A. Wierman (2013). When heavy-tailed and light-tailed flows compete: the response time tail under generalized Max-Weight scheduling. *Proceedings of the* 32^{nd} *IEEE Infocom Conference*.

[18] M.J. Neely, E. Modiano, C. Li (2008). Fairness and optimal stochastic control for heterogeneous networks. *IEEE/ACM Transactions on Networking*, 16(2), pp. 396-409.

[19] K. Park, W. Willinger (2000). Self-similar network traffic: an overview. *Self-Similar Network Traffic and Performance Evaluation*, Wiley Inc.

[20] H. Takagi (1993). Queueing analysis, vol. 3: discrete-time systems. North-Holland.

[21] L. Tassiulas, A. Ephremides (1992). Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control*, 37(12), pp. 1936-1948.

Mihalis G. Markakis received the B.S. degree from the National Technical University of Athens in 2005, the M.S. degree from the University of Southern California in 2008, and the Ph.D. degree from the Massachusetts Institute of Technology in 2013, all in Electrical Engineering. He held an internship position with Alcatel-Lucent Bell Labs in the summer of 2010. He is currently an Assistant Professor in the Department of Economics and Business at Universitat Pompeu Fabra, and also affiliated with the Barcelona School of Management and the Barcelona Graduate School of Economics. His research interests are in modeling, analysis, and optimization of stochastic systems and their applications to Operations Research and Management Science.

Eytan Modiano received his B.S. degree in Electrical Engineering and Computer Science from the University of Connecticut at Storrs in 1986 and his M.S. and PhD degrees, both in Electrical Engineering, from the University of Maryland, College Park, MD, in 1989 and 1992 respectively. He was a Naval Research Laboratory Fellow between 1987 and 1992 and a National Research Council Post Doctoral Fellow during 1992-1993. Between 1993 and 1999 he was with MIT Lincoln Laboratory. Since 1999 he has been on the faculty at MIT, where he is a Professor in the Department of Aeronautics and Astronautics and the Laboratory for Information and Decision Systems (LIDS). His research is on communication networks and protocols with emphasis on satellite, wireless, and optical networks. He is currently an Associate Editor for IEEE/ACM Transactions on Networking, and had served as Associate Editor for IEEE Transactions on Information Theory. He was the Technical Program co-chair for IEEE Wiopt 2006, IEEE Infocom 2007, and ACM MobiHoc 2007.

John N. Tsitsiklis received the B.S. degree in mathematics, and the B.S., M.S., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, in 1980, 1980, 1981, and 1984, respectively. He is currently a Clarence J. Lebel Professor with the Department of Electrical Engineering & Computer Science, MIT.

He has served as a Co-Director of the MIT Operations Research Center (2002-2005), in the National Council on Research and Technology in Greece (2005- 2007), and as the Chair of the Council of the Harokopio University in Athens, Greece. His research interests are in systems, optimization, communications, control, and operations research. He has coauthored four books and several journal papers in these areas.

Dr. Tsitsiklis is a member of the National Academy of Engineering. In 2008, he was conferred the title of Doctor honoris causa from the Université catholique de Louvain, Belgium. He is a (co)recipient of an Outstanding Paper Award from the IEEE Control Systems Society in 1986, the MIT Edgerton Faculty Achievement Award in 1989, the Bodossakis Foundation Prize in 1995, two INFORMS Computing Society Prizes (1997, 2012), and an ACM Sigmetrics Best Paper Award (2013).