Scheduling Strategies to Mitigate the Impact of Bursty Traffic in Wireless Networks

Krishna Jagannathan[1], Libin Jiang[2], Palthya Lakshma Naik[1], and Eytan Modiano[3]

[1]India-UK Advanced Technology Centre of Excellence in Next Generation Networks, IIT Madras, India
[2] Qualcomm Inc., San Diego, CA
[3] Massachusetts Institute of Technology, Cambridge, MA

***Abstract*—Recent work has shown that certain queue-length based scheduling algorithms, such as max-weight, can lead to poor delays in the presence of bursty traffic. To overcome this phenomenon, we consider the problem of designing scheduling policies that are robust to bursty traffic, while also amenable to practical implementation. Specifically, we discuss two mechanisms, one based on adaptive CSMA, and the second based on maximum-weight scheduling with capped queue lengths. We consider a simple queueing network consisting of two conflicting links. The traffic served by the first link is bursty, and is modeled as being heavy-tailed, while traffic at the second link is modeled using a light-tailed arrival process. In this setting, previous work has shown that even the light-tailed traffic would experience heavy-tailed delays under max-weight scheduling.**

In contrast, we demonstrate a threshold phenomenon in the relationship between the arrival rates and the queue backlog distributions. In particular, we show that with an adaptive CSMA scheme, when the arrival rate of the light-tailed traffic is less than a threshold value, the light-tailed traffic experiences a light-tailed queue backlog at steady state, whereas for arrival rates above the same threshold, the light-tailed traffic experiences a heavy-tailed queue backlog. We also show that a similar threshold behavior for max-weight scheduling with capped queue lengths.

I. Introduction

Since modern data networks support highly heterogeneous traffic sources, it is imperative to design network control policies that are inherently robust to burstiness in traffic. Ideally, it is desirable to design and implement control policies that do not adversely affect the stability and delay properties of one traffic flow, due to erratic or bursty behavior of another traffic flow in the network. In this paper, we study scheduling policies that are amenable to practical implementation, while also exhibiting robustness against bursty traffic.

In the context of communication networks, link scheduling for maximum throughput is a well studied problem. Maximum weight scheduling, first proposed in [1], [2], forms the basis for much of the literature on this topic. Despite being capable of supporting the largest possible set of arrival rates in a constrained queueing network, maximum weight scheduling suffers from some drawbacks when it comes to practical implementation. In particular, finding the maximum weight schedule is generally NP-hard, and requires global exchange of the queue length information.

In networks carrying highly heterogeneous traffic, max-weight scheduling suffers another disadvantage. Specifically, when max-weight scheduling is employed in a queueing network carrying a mix of heavy-tailed and light-tailed traffic, it has the tendency to allocate most of the service to the heavy-tailed traffic flow, following the arrival of a large burst. This in turn leads to inordinately large delays and queue backlogs for the light-tailed flows. This tendency of max-weight scheduling to 'infect' light-tailed traffic flows with heavy-tailed delays was first reported in [3], and more thoroughly characterized in [4]. In fact, [3] showed that under certain assumptions on the arrival process, max-weight scheduling can lead to unbounded expected queue occupancy for the light-tailed links - thus resulting in a "delay instability" phenomenon. To help mitigate this effect, [4] showed that a log-max-weight (LMW) scheduling policy can be used to guarantee light-tailed queue backlog for the light-tailed traffic flow, while also maintaining system-wide stability.

While the results in [3], [4] indicate that generalized max-weight policies with appropriately chosen queue length functions can be robust under heterogeneous and bursty traffic, some problems still remain. Most importantly, the policies proposed in [4] need *a priori* knowledge that a particular flow is heavy-tailed, but this is quite hard to determine in practice. Further, generalized max-weight scheduling also requires the global exchange of queue length information, and the computational complexity involved in computing the optimal schedule is prohibitive in large networks.

In a series of recent papers [5]–[8], adaptive CSMA (carrier sense multiple access) based algorithms have been proposed, and shown to achieve maximum throughput. The key idea of adaptive CSMA scheduling is to adjust the transmission aggressiveness (TA) of each link according to its local queue length. Specifically, when the queue length of a link increases, the link transmits more aggressively by using smaller back-off time or larger transmission time; and the link does the opposite when its queue length decreases. The underlying techniques are inspired by ideas from statistical physics, and exploit the product-form stationary distribution of the transmission states under CSMA. Adaptive CSMA based algorithms are expected to find wide-spread applications in wireless networks, owing to their optimality, simplicity of operation, and inherent scalability.

The main contribution of this paper is in analyzing the performance of practical scheduling schemes that achieve maximum throughput, yet help mitigate the effect of bursty traffic in the network. First we characterize the performance of an adaptive CSMA algorithm in a queueing network that serves highly heterogeneous traffic. Our study is motivated by the fact that packet switched networks serve a wide variety of bursty as well as benign traffic sources, and otherwise desirable network control policies (such as adaptive CSMA) must be evaluated for robustness under such varied traffic characteristics. To this end, we consider a simple queueing network consisting of two conflicting links that access a server using adaptive CSMA. One of the links serves heavy-tailed traffic, while the other link serves light-tailed traffic. We demonstrate a threshold phenomenon in the relationship between the arrival rates and the queue backlog distributions. Specifically, we identify a threshold arrival rate λ^* such that when the arrival rate of the light-tailed traffic is less than λ^*, the light-tailed link has light-tailed queue backlog at steady-state. When the arrival rate of the light-tailed traffic exceeds λ^*, the light-tailed traffic suffers a heavy-tailed queue backlog. Comparing this result to those in [3], [4], we can conclude that adaptive CSMA exhibits superior robustness to heterogeneous traffic sources compared to max-weight scheduling, in addition to being much simpler to implement.

Second, we develop a variant of max-weight algorithm with capped queue lengths, and again show that the capping of the queue lengths helps mitigate the effects of heavy-tailed traffic. Since the original max-weight algorithm utilizes potentially unbounded queue length information, it is not directly suitable for scenarios such as a wireless uplink, where the mobiles have to quantize their queue lengths and report it to the base station. To address this problem, we propose a simple variant of the max-weight algorithm, in which the nodes report their queue lengths accurately only if the value is less than a predetermined cap value $Q_{\max}$. This variant is clearly amenable to finite bit rate implementation, and is suitable for scenarios such as the wireless uplink. However, this variant is not throughput optimal in general, since the policy has no way of distinguishing between queues that are larger than $Q_{\max}$. On the other hand, we show that the stability region of the capped max-weight policy approaches the system stability region as $Q_{\max}$ tends to infinity. In other words, for any supportable arrival rate vector, there is a large enough cap value $Q_{\max}$ such that the system is stable under the proposed policy. We prove this result using a piecewise quadratic-linear Lyapunov function.

Next, in order to evaluate the robustness of the capped max-weight policy under bursty traffic, we apply the policy to the system with two conflicting links, one of them carrying heavy-tailed traffic. We demonstrate a threshold phenomenon where the light-tailed traffic experiences a heavy-tailed queue backlog if and only if the arrival rate exceeds half the server capacity. Finally, if we have *a priori* information that a particular traffic flow is heavy-tailed, we show that light-tailed link can always be guaranteed a light-tailed backlog, by choosing (slightly) different caps for the links.

The remainder of this paper is organized as follows. In Section II, we describe the system model and the requisite mathematical preliminaries. In Section III, we study the performance of an adaptive CSMA algorithm in the presence of bursty traffic. In Section IV, we study the max-weight scheduling policy with capped queue lengths. In Section V we discuss some numerical results, and we conclude in Section VI. A subset of the results in this paper were presented by invitation at a recent workshop [9].

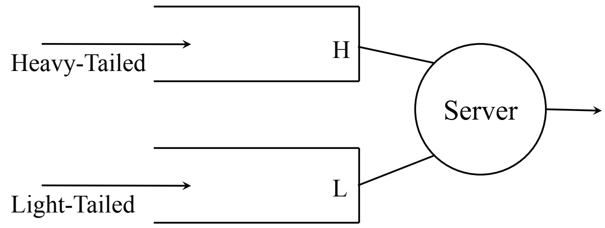

Fig. 1. A queueing system with two conflicting links. One of the links receives heavy-tailed traffic, while the other receives light-tailed traffic.

II. System Description and Preliminaries

Consider a system consisting of two queues, sharing a common server. The queues access the server through conflicting wireless links, i.e., the two queues cannot be served at the same time.

Time is slotted, with bursts of packets arriving at random to each queue at the beginning of each slot. The server is capable of serving one packet per time slot. Although we postpone the precise assumptions on the traffic statistics, let us loosely say that one of the queues receives light-tailed traffic, while the other receives heavy-tailed traffic (Fig. 1). We will refer to the two queues as the light queue and the heavy queue, respectively. Before we specify the precise assumptions on the arrival processes, we pause to present some relevant definitions.

A. Heavy-tailed and light-tailed random variables

Definition 1: A non-negative random variable X is said to be *light-tailed* if there exists $\theta > 0$ for which $\mathbb{E}\left[\exp(\theta X)\right] < \infty$. A random variable is *heavy-tailed* if it is not light-tailed. In other words, a light-tailed random variable is one that has a well defined moment generating function in a neighborhood of the origin. The complementary distribution function of a light-tailed random variable decays at least exponentially fast. Heavy-tailed random variables are those which have complementary distribution functions that decay slower than any exponential. We now define the tail-coefficient of a random variable.

Definition 2: The *tail coefficient* of a random variable X is defined by

$$C_X = \sup\{c \geq 0 \mid \mathbb{E}\left[X^c\right] < \infty\}.$$

In words, the tail coefficient is the threshold where the power moment of a random variable starts to blow up. Note that the tail coefficient of a light-tailed random variable is infinite. On the other hand, the tail coefficient of a heavy-tailed random variable may be infinite (e.g., log-normal distribution) or finite (e.g., Pareto distribution). In this paper, we restrict our attention to the class of heavy-tailed random variables which have a finite tail coefficient.

Next, we state the precise assumptions on the arrival processes.

B. Assumptions on the arrival processes

We assume that the arrival processes to the two queues are independent of each other, and that the arrival process to each queue is independent and identically distributed (i.i.d.) from slot-to-slot. The number of packet arrivals to the light queue during any slot is a light-tailed random variable, with mean λ_L. The number of packet arrivals to the heavy queue during any slot is a heavy-tailed random variable, with tail coefficient C_H ($2 < C_H < \infty$), and mean λ_H.

III. Adaptive CSMA with Heavy-Tailed Traffic

In this section, we characterize the robustness of an adaptive CSMA algorithm to bursty traffic. Specifically, we employ the CSMA algorithm to schedule the links in Fig. 1 and characterize the steady-state queue lengths. Our main result in this section identifies a threshold arrival rate λ^* such that when $\lambda_L < \lambda^*$, the backlog at the light-tailed link is light-tailed, and when $\lambda_L > \lambda^*$, the light-tailed traffic experiences heavy-tailed queue backlog.

A. CSMA scheduling algorithm

We first describe the operation of the adaptive CSMA algorithm. Without loss of generality, we assume that a discrete-time version of the adaptive CSMA algorithm is utilized, where time is slotted and one unit of data can be transmitted in one slot. We note, however, that our main results are not affected if the continuous-time CSMA in [5] are adopted. As mentioned before, the key idea of adaptive CSMA scheduling is to adjust the transmission aggressiveness (TA) of each link according to its queue length. In this paper, we assume that the TA is adjusted at every frame boundary, where each frame has T time slots. Specifically, at the beginning of time slot jT ($j = 0, 1, 2, \dots$)[1], the TA of link k ($k = H, L$) is updated according to

$$r_k[j] = \min\left\{\frac{\alpha}{T}Q_k[j], r_{\max}\right\} \tag{1}$$

where $Q_k[j]$ is the queue length of link k at the beginning of time slot jT, $\alpha < 1$ is a constant, and the constant parameter $r_{\max}$ is the maximum permissible TA.

With TA $r_k[j]$, the CSMA operation of link k is essentially as follows (see [6] for more details). If link k is active (i.e., transmitting) in slot t, then it becomes inactive in slot $t+1$ with probability $1/(1+\exp(r_k[j]))$. If link k and all its conflicting links are inactive in slot t, then link k becomes active in slot $t+1$ with probability $\exp(r_k[j])/(1+\exp(r_k[j]))$. Clearly, link k transmits more aggressively with a higher $r_k[j]$.

[1]Here, the indices of time slots start with 0.

Suppose that the TA's of the two links are fixed at r_H and r_L respectively. Then, when the underlying CSMA Markov chain reaches steady-state, and it can be shown following [5], [6] that the service rate of link k is given by

$$s_k(\mathbf{r}) = \frac{\exp(r_k)}{1+\exp(r_H)+\exp(r_L)},\ k = H, L. \tag{2}$$

Let us refer to the interval between the beginnings of time slots jT and $(j+1)T$ as "frame j". It is not difficult to see that $\{\mathbf{Q}[j], \sigma[j]\}_{j=0,1,\dots}$ forms a Markov chain [10], where $\sigma[j]$ is the transmission state (indicating whether link k ($k = H, L$) is transmitting) just before slot jT. It has been shown in [10] that the "empirical service rate" $\tilde{s}_k[j]$ of link k (i.e., the total number of packets transmitted on link k during frame j, divided by T) satisfies

$$|\mathbb{E}_j\{\tilde{s}_k[j]\} - s_k(\mathbf{r}[j])| \leq b/T \tag{3}$$

for some constant $b > 0$, where $\mathbb{E}_j\{\cdot\}$ is a shorthand for the conditional expectation $\mathbb{E}\{\cdot|\mathbf{Q}[j], \sigma[j]\}$. Let us also define the quantity

$$\lambda^* = \frac{\exp(r_{\max})}{1+2\exp(r_{\max})}. \tag{4}$$

In light of (2), λ^* can be understood as the service rate afforded to each of the queues, when both links attempt transmission with the maximum permissible TA of $r_{\max}$. The quantity λ^* will play an important role in our analysis.

The queue dynamics is given by

$$Q_k[j+1] = \{Q_k[j] - T\cdot\tilde{s}_k[j]\}_+ + a_k[j],\ k = H, L,$$

where $a_k[j]$ is the total number of packet arrivals to link k during frame j. Recall that $a_L[j]$ is assumed to be a light-tailed random variable, and $a_H[j]$ is assumed to be heavy-tailed.

B. Stability

Our first result shows that the adaptive CSMA algorithm can stabilize arrival rates that lie arbitrarily close to the largest possible stability region boundary, as long as the parameters $r_{\max}, T$ and α are chosen appropriately.

Theorem 1: Suppose that $r_{\max} > 1$, and that

$$\lambda_L + \lambda_H \leq \frac{\exp(r_{\max})}{1+\exp(r_{\max})+\exp(1)} - 2\epsilon$$

where $\epsilon > 0$. Choose $T \geq 2b/\epsilon$. Then both queues are stable.

Remark 1: : Clearly, the stability region approaches the largest possible region ($\lambda_L + \lambda_H < 1$) as $r_{\max} \to \infty$ and $\epsilon \to 0$. This stability result is shown to hold in very general single-hop networks in [5]. However, for the simple system under consideration, a proof involving a linear Lyapunov function suffices, as shown in Appendix A.

C. Queue length behavior at steady-state

We are now ready to state the main results regarding the steady-state queue lengths under adaptive CSMA. First, we show that when λ_L is less than λ^*, the steady-state queue occupancy at the light queue is light-tailed.

Theorem 2: Suppose that $\lambda_L \leq \lambda^* - \epsilon$ where $\epsilon > 0$. Choose $T \geq 2b/\epsilon$. Then, the steady-state queue occupancy Q_L of the light queue is light-tailed.

Proof: We first show that whenever $Q_L[j] > r_{\max}\frac{T}{\alpha}$, the light queue has negative drift. Suppose that $Q_L[j] > r_{\max}\frac{T}{\alpha}$. Then, according to (1), $r_L[j] = r_{\max}$. Since $r_H[j] \leq r_{\max}$, by (2), we have

$$s_L(\mathbf{r}[j]) \geq \frac{\exp(r_{\max})}{1 + 2\exp(r_{\max})} = \lambda^*.$$

Therefore, by (3),

$$\mathbb{E}_j\{\tilde{s}_L[j]\} \geq s_L(\mathbf{r}[j]) - b/T \geq \frac{\exp(r_{\max})}{1 + 2\exp(r_{\max})} - \frac{\epsilon}{2}.$$

So $\lambda_L - \mathbb{E}_j\{\tilde{s}_L[j]\} \leq -\epsilon/2 < 0$ (i.e., there is a negative drift). Next, since the arrival process is light-tailed, we can invoke Theorem 2.3 (Eq. (2.8)) in [11], to conclude that the light queue distribution is light-tailed in steady-state. ■

Our next result is a converse to Theorem 2. In other words, we show that when the arrival rate of the light-tailed traffic is greater than the threshold value λ^*, the steady-state queue occupancy at the light queue is heavy-tailed.

Theorem 3: Suppose $\lambda_L > \lambda^*$, and that the CSMA parameters are chosen such that the system is stable. Then, the steady-state occupancy of the light queue is heavy-tailed. with tail coefficient at most $C_H - 1$.

Proof: We need to show that for any $\delta > 0$,

$$\mathbb{E}\left[Q_L^{C_H-1+\delta}\right] = \infty. \tag{5}$$

Instead, we will first show that

$$\lim_{j\to\infty} \mathbb{E}\left[Q_L[j]^{C_H-1+\delta}\right] = \infty. \tag{6}$$

The above (Eq. 6) concerns the limit of the expectation of a sequence of random variables, whereas what we really want to show in (5) is regarding the expectation of the limiting random variable Q_L. Although it is by no means obvious that the limit and the expectation can be interchanged here, we will ignore this technical detail for now.

To prove (6), we first note that the time intervals between two successive frame boundaries at which the system empties constitute renewal intervals. Let us denote by T_R the random variable representing the number of frames in a renewal interval. Since the system is stable, we have $\mathbb{E}[T_R] < \infty$.

Let us now define the renewal reward function

$$R[j] = Q_L[j]^{C_H-1+\delta}.$$

By the key renewal theorem [12],

$$\lim_{j\to\infty} \mathbb{E}[R[j]] = \frac{\mathbb{E}[R]}{\mathbb{E}[T_R]},$$

where $\mathbb{E}[R]$ denotes the expected reward accumulated over a renewal interval, and $\mathbb{E}[T_R] < \infty$. It is therefore enough to show that[2]

$$\mathbb{E}\left[\sum_{i=0}^{T_R} Q_L[i]^{C_H-1+\delta}\right] = \infty. \tag{7}$$

[2]Without loss of generality, we have considered a busy period that commences at time 0.

The busy period that commences at time 0 can be of three different types. It can commence with (i) a burst arriving to the light-tailed link alone, or (ii) a burst arriving to the heavy-tailed link alone, or (iii) bursts arriving to both links simultaneously. It can be shown that all the three events have positive probabilities. The event that is of interest to us is (ii), i.e., the busy period commencing with a burst at the heavy queue only, so that $Q_H[0] > 0$ and $Q_L[0] = 0$. Let us denote this event by $\mathcal{E}_H = \{Q_H[0]) > 0, Q_L[0] = 0\}$.

We now have the following lower bound on (7)

$$\mathbb{E}\left[\sum_{i=0}^{T_R} Q_L[i]^{C_H-1+\delta}\right] \geq \mathbb{P}\{\mathcal{E}_H\}\,\mathbb{E}\left[\sum_{i=0}^{T_R} Q_L[i]^{C_H-1+\delta}\Big|\mathcal{E}_H\right]$$

$$= \mathbb{P}\{\mathcal{E}_H\}\,\mathbb{E}_b\left[\mathbb{E}\left[\sum_{i=0}^{T_R} Q_L[i]^{C_H-1+\delta}\Big|\mathcal{E}_H,\ Q_H[0] = b\right]\right].$$

In the last step above, we have iterated the expectation over the initial burst size b. The inner expectation above is a function of b; let us denote it by

$$g_\delta(b) := \mathbb{E}\left[\sum_{i=0}^{T_R} Q_L[i]^{C_H-1+\delta}\Big|\mathcal{E}_H,\ Q_H[0] = b\right].$$

Thus,

$$\mathbb{E}\left[\sum_{i=0}^{T_R} Q_L[i]^{C_H-1+\delta}\right] \geq \mathbb{P}\{\mathcal{E}_H\}\,\mathbb{E}_b[g_\delta(b)]$$
$$\geq \mathbb{P}\{\mathcal{E}_H\}\,\mathbb{E}_b\left[g_\delta(b)1_{b>b_0}\right], \tag{8}$$

$\forall b_0 \geq 1$. Since the above bound is true for any b_0, we can make b_0 as large as we want.

For large enough b, the TA of the heavy queue will saturate at $r_{\max}$ starting from frame 1, and remain at $r_{\max}$ until the occupancy of the heavy queue falls to $\frac{Tr_{\max}}{\alpha}$. In other words, the TA of the heavy queue will remain at $r_{\max}$ for at least τ_b frames, where

$$\tau_b = \frac{b - Tr_{\max}/\alpha}{T} = \frac{b}{T} - \frac{r_{\max}}{\alpha}.$$

Note that τ_b is $O(b)$. During these τ_b time frames, the light queue receives service at a rate of λ^* at best, according to (2), (3) and (4). At the same time, the light queue receives arrivals at the rate λ_L. Since $\lambda_L > \lambda^*$, the light queue will build up to $O(b)$ during these τ_b frames with high probability.

To make the above argument precise, choose $\xi > 0$ such that $\lambda_L - \lambda^* - 2\xi = \eta > 0$, and choose an arbitrarily small $\kappa > 0$. For any sample path of arrivals to the light queue, the TA of the light queue is at most $r_{\max}$. Since the heavy queue is also attempting at $r_{\max}$ during the first τ_b frames, it is clear that the long term service rate given to the light queue can be at most λ^*. Indeed, using the machinery in [10], it can be shown that for *every* arrival sample path during the first τ_b time frames, and for suitably large T and b, the empirical service afforded to the light queue satisfies[3]

$$\mathbb{P}\left\{\frac{1}{T\tau_b}\sum_{j=1}^{\tau_b} T\tilde{s}_L[j] \leq \lambda^* + \xi\right\} \geq 1-\kappa. \qquad (9)$$

Next, we invoke the weak law of large numbers for the light-tailed arrival process to assert that the empirical rate is very likely to be around λ_L. That is, for any $\kappa > 0$ and large enough b,

$$\mathbb{P}\left\{\frac{1}{T\tau_b}\sum_{j=1}^{\tau_b} a_L[j] \geq \lambda_L - \xi\right\} \geq 1-\kappa. \qquad (10)$$

Note that (9) holds for all arrival sample paths, and in particular, for the 'typical' sample paths characterized in (10). Therefore, we can assert that the event corresponding to the arrivals being typical (in the sense of (10)) as well as the empirical service being typical has high probability. Thus,

$$\mathbb{P}\left\{\frac{1}{T\tau_b}\sum_{j=1}^{\tau_b}(a_L[j] - T\tilde{s}_L[j]) \geq \lambda_L - \lambda^* - 2\xi\right\} \geq 1-2\kappa, \qquad (11)$$

which implies that

$$\mathbb{P}\left\{Q_L[\tau_b] \geq T\tau_b\eta\right\} \geq 1-2\kappa. \qquad (12)$$

The above lower bound asserts that the light queue at least grows to $O(b)$ with high probability, by the end of the frame τ_b. It is now inevitable that the light queue stays above $Q_L[\tau_b]/2$ for another $O(b)$ time slots, since at most one packet can be served in a slot. In particular,

$$Q_L[j] \geq T\tau_b\eta/2;\ \tau_b \leq j \leq \tau_b + \tau_b\left(\frac{\eta}{2}\right),$$

with probability at least $1-2\kappa$. Thus, the light queue stays at $O(b)$ for $O(b)$ time frames with high probability.

We can thus write the following lower bound for large enough b_0 and $b > b_0$

$$\begin{aligned} g_\delta(b)1_{\{b>b_0\}} &= \mathbb{E}\left[\sum_{i=0}^{T_R} Q_L[i]^{C_H-1+\delta}\Big|\mathcal{E}_H, Q_H[0]=b\right]1_{\{b>b_0\}} \\ &\geq \left[(1-2\kappa)\sum_{i=\tau_b}^{\tau_b+\tau_b\left(\frac{\eta}{2}\right)}(T\tau_b\eta/2)^{C_H-1+\delta}\right]1_{\{b>b_0\}} \\ &= (1-2\kappa)T\left(\tau_b\eta/2\right)^{C_H+\delta}1_{\{b>b_0\}} = K_1 b^{C_H+\delta}1_{\{b>b_0\}}, \end{aligned} \qquad (13)$$

for some positive constant K_1. Now, we use (13) in (8) to get

$$\mathbb{E}\left[\sum_{i=0}^{T_R} Q_L[i]^{C_H-1+\delta}\right] \geq \mathbb{P}\left\{\mathcal{E}_H\right\}\mathbb{E}_b\left[K_1 b^{C_H+\delta}1_{b>b_0}\right] = \infty.$$

The last step is because the initial burst size b has tail coefficient C_H, so that $\mathbb{E}_b\left[b^{C_H+\delta}\right] = \mathbb{E}_b\left[b^{C_H+\delta}1_{b>b_0}\right] = \infty$ for all b_0 and $\delta > 0$. Therefore, we are done proving (7), from which (6) follows.

Finally, to prove that the limit and expectation can legitimately be interchanged in (6), one needs to use a truncation argument, followed by repeated use of the monotone convergence theorem and the dominated convergence theorem. In particular, we can imitate the methodology used in proving [13, Proposition 5.4]. ■

[3]The probabilities in (9) through (12) are conditioned on $\mathcal{E}_H$, and on some $b > b_0$, where b_0 is chosen suitably large. We refrain from explicitly notating the conditioning, to avoid cumbersome expressions.

D. Discussion

In this section, we studied the performance of adaptive CSMA under heavy-tailed traffic. We identified a threshold arrival rate λ^*, such that when the arrival rate of the light-tailed traffic is less than λ^*, the light queue has light-tailed queue backlog in steady-state (Theorem 2). When the arrival rate of the light-tailed traffic exceeds λ^*, the light-tailed traffic suffers a heavy-tailed queue backlog in steady-state (Theorem 3). Since λ^* is close to one half for large $r_{\max}$, our result is tantamount to saying that adaptive CSMA induces heavy-tailed backlog for the light queue only if the light-tailed traffic is responsible for more than half the total supportable traffic rate in the system.

On the other hand, it was shown in [4] that maximum weight scheduling and its generalized version called max-weight-α scheduling induce a heavy-tailed queue backlog at the light queue, for all non-zero arrival rates of the heavy-tailed and light-tailed traffic. Furthermore, it was shown that max-weight scheduling induces the worst possible asymptotic behavior on the light queue among all non-idling policies. In comparison, the adaptive CSMA algorithm performs better in terms of the backlog faced by the light-tailed traffic.

As explained in [4], max-weight scheduling induces very poor queue backlog on the light-tailed link because large burst arrivals to the heavy-tailed link can starve the light queue for extended durations. On the other hand, with adaptive CSMA, all links have bounded TA values. As a result, even when large bursts arrive at a link carrying heavy-tailed traffic, the link cannot take over the server by attempting to transmit with arbitrary aggressiveness. This has the effect of 'shielding' the light-tailed traffic from the large bursts, at least when the arrival rate is smaller than the threshold value. Furthermore, adaptive CSMA does not need any a priori information about traffic statistics. In contrast, the policies proposed in [4] to mitigate the effect of heavy-tailed traffic need to have a priori information about which flow is heavy-tailed.

We wish to point out that the cap on the TA values of each link in adaptive CSMA was originally intended as a mechanism to bound the mixing time of the CSMA Markov chain. In other words, capped TA values imply bounded 'fugacities' in the underlying Glauber dynamics, which leads to bounded mixing time. However in our context, the bounded TA values help in another way as well, by preventing the heavy-tailed link from attempting too aggressively.

IV. Maximum Weight Scheduling with Capped Queue Lengths

As we just discussed, adaptive CSMA is able to mitigate the effects of the heavy-tailed traffic by capping the transmission aggressiveness (TA) parameter, thus preventing the heavy-tailed link from attempting too aggressively. In this section we show that a similar idea can also be used with max-weight scheduling. In particular, capping the queue-length values can similarly be used to prevent the heavy-tailed traffic from saturating the channel. We develop a simple variation of the well known max-weight scheduling policy, where the nodes only report their capped queue lengths to the scheduler. In particular, a queue length cap $Q_{\max}$ is chosen in advance, and the nodes report their queue lengths accurately when it is less than or equal to $Q_{\max}$, otherwise nodes simply report $Q_{\max}$. In the proposed scheme, the scheduler performs max-weight scheduling as usual, except it only has access to the capped queue length values from each node.

Since the scheduler cannot discern large queue build up beyond $Q_{\max}$ in any queue, the max-weight policy with capped queue lengths is not throughput optimal in general. Interestingly however, we show in this section that the stability region of max-weight policy with capped queue lengths approaches the stability region of the system, as the cap value increases. Specifically, in a general single-hop network and for any arrival rate vector stabilizable by the max-weight scheduling policy, there exists a cap value $Q_{\max}$ for which the corresponding max-weight policy with capped queue lengths can stabilize the system under that arrival rate vector.

We now describe a simple single-hop network model and describe the operation of the max-weight policy with capped queue lengths. Consider a network represented by a graph $G = (V, E)$, where E represents the set of links in the network, and V denotes the set of nodes. Let $\mathcal{S}$ denote the set of all feasible link activations in the network. Note that $\mathcal{S}$ is usually a proper subset of the power set of E, since not all links can be simultaneously activated, due to interference constraints. A scheduling policy selects a valid link schedule during each time slot from the feasible set $\mathcal{S}$. The stability region Λ of the system is defined as the set of all arrival rate vectors that can be stably supported by some scheduling policy.

Let $\mu_k,\ k \in E$ denote the transmission rate on link k, if the link were to be scheduled. The celebrated max-weight policy [1], [2] picks the following schedule $s^*(t)$ at each time slot, where

$$s^*(t) = \arg\max_{s(t)\in\mathcal{S}} \sum_{k\in s(t)} Q_k(t)\mu_k.$$

The max-weight policy with capped queue lengths operates as follows. Let $Q_{\max}$ be a cap value chosen a priori. During time slot t, each link k reports its capped queue length

$$\tilde{Q}_k(t) = \min(Q_k(t), Q_{\max})$$

to the scheduler. The scheduler picks the schedule $\tilde{s}(t)$ according to

$$\tilde{s}(t) = \arg\max_{s(t)\in\mathcal{S}} \sum_{k\in s(t)} \tilde{Q}_k(t)\mu_k. \tag{14}$$

In case of a tie, the scheduler picks uniformly at random among the tied schedules.

The max-weight policy with capped queue lengths is not throughput optimal in general. However, we show next that the policy can come arbitrarily close to being throughput optimal, for a large enough cap value.

Theorem 4: Let $\underline{\lambda}$ be any arrival rate vector in the interior of the stability region Λ defined above. Then, there exists a value of $Q_{\max}$ such that the max-weight policy with queue lengths capped at $Q_{\max}$ can keep the queues in the single-hop network stable, when the arrival rate is $\underline{\lambda}$.

The proof of Theorem 4 utilizes a piecewise quadratic-linear Lyapunov function, similar to [14, Section 3.11], and is given in Appendix B.

We remark that the value of $Q_{\max}$ needed to stabilize the system depends on the arrival rate vector in general. Further, the policy is amenable to finite bit-rate implementation with quantized queue length information, and is particularly well suited to a wireless uplink scenario, where centralized scheduling is not a problem.

Having shown in Theorem 4 that max-weight scheduling with capped queue lengths is asymptotically throughput optimal, we proceed to investigate its performance in the presence of heavy-tailed traffic. To that end, we return to the simple two link network in Fig. 1, and analyze the queue backlogs under max-weight scheduling with capped queue lengths.

A. Impact of heavy-tailed traffic

Recall the system model in Fig. 1, and the assumptions in Section II-B. Suppose now that max-weight scheduling with queue length cap $Q_{\max}$ is performed in this system. In particular, for this simple system, the scheduler will serve the longer queue in each slot, as long as the shortest queue is smaller than $Q_{\max}$. If both queues are longer than $Q_{\max}$ during a particular slot, the scheduler will serve one of the queues uniformly at random.

We show that for any fixed $Q_{\max}$, the light queue suffers a heavy-tailed queue backlog, if its arrival rate exceeds $1/2$, and a light-tailed backlog if its arrival rate is less than $1/2$.

Theorem 5: Suppose that a max-weight policy with queue lengths capped at $Q_{\max}$ is used to schedule the links in Fig. 1. Then, if $\lambda_L > 1/2$, the steady-state queue backlog Q_L at the light queue is heavy-tailed with tail coefficient $C_H - 1$. If $\lambda_L < 1/2$, Q_L is light-tailed.

Proof: The key observation is that whenever both queues have more than $Q_{\max}$ packets, the max-weight policy with capped queue lengths behaves essentially like a random scheduling policy. Therefore, it should not be surprising that the asymptotic steady-state queue length distributions should be similar to that under a randomized scheduling policy [3]. Indeed, the result can be proved by using techniques similar to Theorem 3 for the heavy-tailed part and Theorem 2 for the light-tailed part. ∎

Thus, we have shown that under max-weight policy with capped queue lengths, the light queue suffers a heavy-tailed queue backlog only when the arrival rate to the queue exceeds half the server capacity. This is in contrast to the original max-weight policy *always* induces heavy-tailed queue backlog at the light queue, for all non-zero arrival rates [3], [4].

Finally, if we know a priori which of the two links carries heavy-tailed traffic, we show that it is possible to ensure light-tailed backlog at the light queue for all arrival rates. A simple way to achieve this is to assign a larger queue length cap to the light-tailed link, so that when Q_L is large, the light-tailed link effectively gets priority. This is sufficient to induce a light-tailed backlog at the light queue.

Proposition 1: Suppose that the following variant of the max-weight policy with capped queue lengths is implemented in the system shown in Fig. 1. The heavy-tailed link reports its queue length capped by some $Q_{\max}$, while the light-tailed link reports its queue length capped by $Q_{\max}+1$. In this case, the steady-state backlog at the light queue is light-tailed, whenever the system is stable.

Proof: When Q_L is large enough (in particular, larger than $Q_{\max}+1$), the heavy-tailed link reports $\min\{Q_H, Q_{\max}\} \leq Q_{\max}$ as its queue length, while the light-tailed link reports $Q_{\max}+1$. Thus, the scheduler will give priority to the light-tailed traffic whenever $Q_L \geq Q_{\max}+1$. This implies that the light queue backlog is light-tailed for all arrival rates. ∎

V. Numerical Results

In this section, we use numerical simulations to study the tail behavior of the steady-state backlog at the light-tailed link. In particular, we compare the tail behavior under adaptive CSMA and max-weight scheduling. We simulate a scenario where the heavy queue is fed by a discrete Pareto distributed traffic source, with arrival rate $\lambda_H = 0.3$ packets/slot, and tail coefficient $C_H = 3$. The light queue is fed by a discrete Poisson source, with $\lambda_L = 0.4$ packets/slot. The system was simulated in Matlab over 100 million time slots to obtain the empirical backlog distribution.

Fig. 2 shows a log-log plot of the backlog tail distribution at the light queue (i.e., $\log \mathbb{P}\{Q_L > b\}$ versus $\log b$) under max-weight scheduling, and the discrete time adaptive CSMA algorithm proposed in [6]. As seen from the figure, the plot corresponding to max-weight scheduling is approximately linear with a negative slope, indicating that the backlog distribution is heavy-tailed [3]. In contrast, the plot corresponding to CSMA exhibits a distinctive 'waterfall' shape, which is characteristic of a light-tailed distribution. This verifies our assertion that adaptive CSMA can lead to a light-tailed backlog at the light queue, even if max-weight scheduling leads to a heavy-tailed backlog. Another interesting feature to note in Fig. 2 is that the *average* queue backlog seems to be much larger under CSMA than under max-weight, although the former exhibits superior asymptotics. Indeed, simple CSMA algorithms are known to suffer from poor average delay, and various innovative modifications have been proposed in the literature [6], [10], [15] to counter this problem.

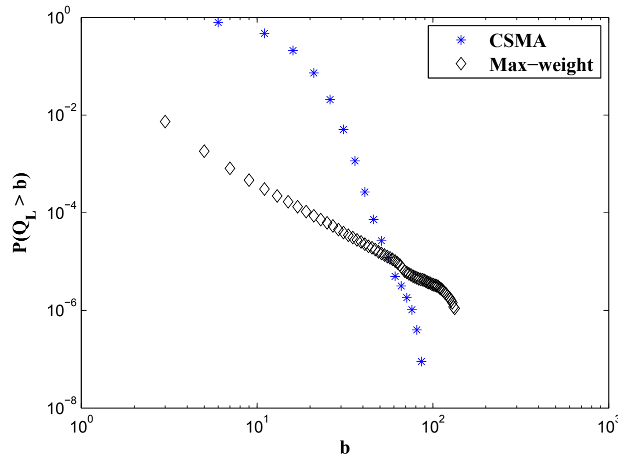

Fig. 2. Log-log plot of the backlog tail distribution $\mathbb{P}\{Q_L > b\}$.

VI. Concluding Remarks

We investigated scheduling policies that are robust to bursty traffic, while also being suitable to practical implementation. In particular, we discussed two scheduling schemes, one based on adaptive CSMA, and another based on max-weight scheduling with capped queue lengths. In a simple queueing network consisting of two links, one of them carrying heavy-tailed traffic, we showed a threshold phenomenon for the asymptotic queue length behavior. Specifically, for both adaptive CSMA and max-weight scheduling with capped queue lengths, we showed that the light queue suffers a heavy-tailed backlog if and only if the arrival rate exceeds about half the server capacity.

Our study suggests that adaptive CSMA has the potential to be more robust than max-weight scheduling, in queueing networks that serve highly heterogeneous traffic. Our results also suggest that max-weight scheduling, when implemented approximately using quantized queue length information, does not perform as poorly as predicted in [4]. Notably, neither the adaptive CSMA algorithm nor the max-weight policy with capped queue lengths needs a priori information regarding which particular flow in the network is heavy-tailed. Finally, if this information is available a priori, then we show that the capped max-weight policy can be suitably modified to ensure light-tailed backlog at the light queue for all arrival rates.

Appendix A
Proof of Theorem 1

Consider the Lyapunov function $L[j] := Q_L[j] + Q_H[j] \geq 0$. We will show that if $L[j] \geq 2r_{\max}T/\alpha$, then the following holds:

$$\mathbb{E}_j\{L[j+1] - L[j]\} \leq -T\epsilon. \tag{15}$$

Suppose that $L[j] \geq 2r_{\max}T/\alpha$, then $Q_k[j] \geq r_{\max}T/\alpha$ for some k. Without loss in generality, assume that $Q_L[j] \geq$

$r_{\max}T/\alpha$. So $r_L[j] = r_{\max}$. Since $r_{\max} > 1$ and $\alpha < 1$, one has $Q_L[j] \geq T$. Therefore, $Q_L[j+1] = Q_L[j] - T \cdot \tilde{s}_L[j] + a_L[j]$.

Now consider two cases.

Case 1: If $Q_H[j] \geq T$, then $Q_H[j+1] = Q_H[j] - T \cdot \tilde{s}_H[j] + a_H[j]$. Consequently,

$$\mathbb{E}_j\{L[j+1]\} = L[j] + T\{\lambda_L + \lambda_H - \mathbb{E}_j\{\tilde{s}_L[j] + \tilde{s}_H[j]\}\}.$$

With $T \geq 2b/\epsilon$, the RHS of (3) is less than or equal to $\epsilon/2$. Therefore,

$$\begin{aligned}\mathbb{E}_j\{\tilde{s}_L[j] + \tilde{s}_H[j]\} &\geq \frac{\exp(r_L[j]) + \exp(r_H[j])}{1 + \exp(r_L[j]) + \exp(r_H[j])} - \epsilon \\ &\geq \frac{\exp(r_{\max})}{1 + \exp(r_{\max})} - \epsilon.\end{aligned}$$

So (15) holds.

Case 2: If $Q_H[j] < T$, we have $Q_H[j+1] \leq Q_H[j] + a_H[j]$. So

$$\mathbb{E}_j\{L[j+1]\} \leq L[j] + T\{\lambda_L + \lambda_H - \mathbb{E}_j\{\tilde{s}_L[j]\}\}.$$

Since $Q_H[j] < T$ we also have $r_H[j] < \alpha < 1$. Therefore

$$\begin{aligned}s_L(\mathbf{r}[j]) = &\frac{\exp(r_{\max})}{1 + \exp(r_{\max}) + \exp(r_H[j])} \\ &\geq \frac{\exp(r_{\max})}{1 + \exp(r_{\max}) + \exp(1)}.\end{aligned}$$

So (15) still holds.

Therefore, the Lyapunov function has a negative drift whenever $L[j] \geq 2r_{\max}T/\alpha$. Combined with the fact that $\mathbb{E}_j\{L[j+1]\} - L[j]$ is bounded, by the Foster-Lyapunov Criterion, we conclude that the queues are stable.

APPENDIX B
PROOF OF THEOREM 4

We prove this result for a general single-hop network. Define the Lyapunov function $L(\mathbf{Q}) := \sum_k L_k(Q_k)$ where

$$L_k(Q_k) = (Q_{\max}Q_k)1_{\{Q_k \geq Q_{\max}\}} + \frac{Q_k^2 + Q_{\max}^2}{2} 1_{\{Q_k < Q_{\max}\}}.$$

We have

$$\begin{aligned}\frac{\partial L(\mathbf{Q})}{\partial Q_k} &= Q_{\max}1_{\{Q_k \geq Q_{\max}\}} + Q_k 1_{\{Q_k < Q_{\max}\}} \\ &= Q_k \wedge Q_{\max}. \qquad (16)\end{aligned}$$

Let $x_k(\mathbf{Q}(t)) := \mu_k \cdot 1_{\{k \in \tilde{s}(t)\}}$, where $\tilde{s}(t)$ is defined in (14). Then,

$$\begin{aligned}\Delta(t) &:= E[L(\mathbf{Q}(t+1))|\mathbf{Q}(t))] - L(\mathbf{Q}(t)) \\ &\leq \sum_k \frac{\partial L(\mathbf{Q})}{\partial Q_k}[\underline{\lambda}_k - x_k(\mathbf{Q}(t))] + D/2 \\ &= (\mathbf{Q}(t) \wedge Q_{\max})^T(\underline{\lambda} - \mathbf{x}(\mathbf{Q}(t))) + D/2 \\ &= (\tilde{\mathbf{Q}}(t))^T(\underline{\lambda} - \mathbf{x}(\mathbf{Q}(t))) + D/2,\end{aligned}$$

where $D := \sum_k E[a_k^2(t)] + \sum_k \mu_k^2$, with $a_k(t)$ being the amount of arrivals to queue k in slot t. D is finite since we have assumed that the tail coefficient of $a_k(t)$ is larger than 2.

Since $\underline{\lambda} \in \Lambda^o$, there is $\epsilon > 0$ such that $\underline{\lambda} + \epsilon\mathbf{1} \in \Lambda$. Since $\mathbf{x}(\mathbf{Q}(t)) \in \arg\max_{\mathbf{x}\in\Lambda}(\tilde{\mathbf{Q}}(t))^T\mathbf{x}$, we have

$$(\tilde{\mathbf{Q}}(t))^T\mathbf{x}(\mathbf{Q}(t)) \geq (\tilde{\mathbf{Q}}(t))^T(\underline{\lambda} + \epsilon\mathbf{1}).$$

Therefore,

$$\Delta(t) \leq -\epsilon \cdot (\tilde{\mathbf{Q}}(t))^T\mathbf{1} + D/2.$$

Choose $Q_{\max} = D/\epsilon$. Then, whenever $||\mathbf{Q}(t)||_\infty \geq Q_{\max}$,

$$\Delta(t) \leq -\epsilon Q_{\max} + D/2 = -D/2 < 0. \qquad (17)$$

Also, for any $\mathbf{Q}(t)$, we have $\Delta(t) \leq D/2 < \infty$. Therefore, by the Foster-Lyapunov Criterion, $\mathbf{Q}(t)$ is stable under the policy.

ACKNOWLEDGEMENTS

The first author gratefully acknowledges the IU-ATC project funded by the Department of Science and Technology, Government of India, and the UK EPSRC Digital Economy programme. This work was also supported by NSF Grants CNS-0915988,CNS-1217048 and ARO MURI Grant W911NF-08-1-0238.

REFERENCES

[1] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936–1948, 1992.

[2] ——, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Transactions on Information Theory*, vol. 39, no. 2, pp. 466–478, 1993.

[3] M. Markakis, E. Modiano, and J. Tsitsiklis, "Scheduling policies for single-hop networks with heavy-tailed traffic," in *Allerton Conference, Monticello IL*, 2009.

[4] K. Jagannathan, M. Markakis, E. Modiano, and J. N. Tsitsiklis, "Queue-length asymptotics for generalized max-weight scheduling in the presence of heavy-tailed traffic," *To appear, IEEE/ACM Transactions on Networking*, 2012.

[5] L. Jiang and J. Walrand, "A distributed csma algorithm for throughput and utility maximization in wireless networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 18, no. 3, pp. 960–972, 2010.

[6] J. Ni, B. Tan, and R. Srikant, "Q-csma: Queue-length based csma/ca algorithms for achieving maximum throughput and low delay in wireless networks," in *INFOCOM, 2010 Proceedings IEEE*. IEEE, 2010.

[7] S. Rajagopalan and D. Shah, "Distributed algorithm and reversible network," in *Information Sciences and Systems, 2008. 42nd Annual Conference on*. IEEE, 2008, pp. 498–502.

[8] S. Rajagopalan, D. Shah, and J. Shin, "Network adiabatic theorem: an efficient randomized protocol for contention resolution," in *Proceedings of the eleventh international joint conference on Measurement and modeling of computer systems*. ACM, 2009, pp. 133–144.

[9] K. Jagannathan, L. Jiang, and E. Modiano, "On scheduling algorithms robust to heavy-tailed traffic (invited)," in *ITA Workshop, San Diego*, 2012.

[10] L. Jiang, M. Leconte, J. Ni, R. Srikant, and J. Walrand, "Fast mixing of parallel glauber dynamics and low-delay csma scheduling," in *INFOCOM*. IEEE, 2011, pp. 371–375.

[11] B. Hajek, "Hitting-time and occupation-time bounds implied by drift analysis with applications," *Advances in Applied probability*, pp. 502–525, 1982.

[12] R. Gallager, *Discrete stochastic processes*. Kluwer, 1996.

[13] K. Jagannathan, "Asymptotic performance of queue-length-based network control policies," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, 2010.

[14] L. Jiang and J. Walrand, "Scheduling and congestion control for wireless and processing networks," *Synthesis Lectures on Communication Networks*, vol. 3, no. 1, pp. 1–156, 2010.

[15] M. Lotfinezhad and P. Marbach, "Throughput-optimal random access with order-optimal delay," *Arxiv preprint arXiv:1009.5944*, 2010.