Delay-Based Back-Pressure Scheduling in Multi-Hop Wireless Networks

Bo Ji
Department of ECE
The Ohio State University
Email: ji@ece.osu.edu

Changhee Joo
Department of EECE
Korea University of
Technology and Education, Korea
Email: cjoo@kut.ac.kr

Ness B. Shroff
Departments of ECE and CSE
The Ohio State University
Email: shroff@ece.osu.edu

Abstract—Scheduling is a critical and challenging resource allocation mechanism for multi-hop wireless networks. It is well known that scheduling schemes that give a higher priority to the link with larger queue length can achieve high throughput performance. However, this queue-length-based approach could potentially suffer from large (even infinite) packet delays due to the well-known *last packet* problem, whereby packets may get excessively delayed due to lack of subsequent packet arrivals. Delay-based schemes have the potential to resolve this last packet problem by scheduling the link based on the delay for the packet has encountered. However, the throughput performance of delay-based schemes has largely been an open problem except in limited cases of single-hop networks. In this paper, we investigate delay-based scheduling schemes for multi-hop traffic scenarios. We view packet delays from a different perspective, and develop a scheduling scheme based on a new delay metric. Through rigorous analysis, we show that the proposed scheme achieves the optimal throughput performance. Finally, we conduct extensive simulations to support our analytical results, and show that the delay-based scheduler successfully removes excessive packet delays, while it achieves the same throughput region as the queue-length-based scheme.

I. INTRODUCTION

Link scheduling is a critical resource allocation component in multi-hop wireless networks, and also perhaps the most challenging. The celebrated Queue-length-based Back-Pressure (Q-BP) scheduler [1] has been shown to be throughput-optimal and can stabilize the network under any feasible load. Since the development of Q-BP, there have been numerous extensions that have integrated it in an overall optimal cross-layer solution. Further, easier-to-implement queue-length-based scheduling schemes have been developed and shown to be throughput-efficient (see [2] and references therein). Some recent attempts [3], [4] focus on designing real-world wireless protocols using the ideas behind these algorithms.

While these queue-length-based schedulers have been shown to achieve excellent throughput performance, they are usually evaluated under the assumption that flows have an infinite amount of data and keep injecting packets into the network. However, in practice accounting for multiple time scales [5]–[7], there also exist other types of flows that have

a finite number of packets to transmit, which can result in the well-known *last packet* problem: consider a queue that holds the last packet of a flow, then the packet does not see any subsequent packet arrivals, and thus the queue length remains very small and the link may be starved for a long time, since the queue-length-based schemes give a higher priority to links with a larger queue length. In such a scenario, it has also been shown in [5] that the queue-length-based schemes may not even be throughput-optimal.

Recent works in [8]–[13] have studied the performance of delay-based scheduling algorithms that use the Head-of-Line (HOL) delay instead of queue length as link weight. One desirable property of the delay-based approach is that they provide an intuitive way around the last packet problem. The schedulers give a higher priority to the links with a larger weight as before, but the weight (i.e., the HOL delay) of a link increases with time until the link gets scheduled. Hence, if the link with the last packet is not scheduled at this moment, it is more likely to be scheduled in the next time. However, the throughput performance of the delay-based scheduling schemes is not fully understood, and has merely been shown for limited cases of single-hop networks.

The delay-based approach was introduced in [8] for scheduling in Input-Queued switches. The results have been extended to wireless networks for single-hop traffic, providing throughput-optimal delay-based MaxWeight scheduling algorithms [10], [11], [14]. It is also shown that delay-based schemes with appropriately chosen weight parameters also provide good Quality of Service (QoS) [9], and can be used as an important component in a cross-layer protocol design [13]. The performance of the delay-based MaxWeight scheduler has been further investigated in a single-hop network with flow dynamics [12]. The results show that, when flows arrive at the base station carrying a finite amount of data, the delay-based MaxWeight scheduler achieves the optimal throughput performance while its queue-length-based counterpart does not.

However, in multi-hop wireless networks, the throughput performance of these delay-based schemes has largely been an open problem. To the best of our knowledge, there are no prior works that employ delay-based algorithms to address the important issue of throughput-optimal scheduling in wireless

networks with multi-hop traffic. The problem turns out to be far more challenging in the multi-hop scenario due to the following reason. In [11], the key idea in showing throughput optimality of the delay-based MaxWeight scheduler is to exploit the following property: after a finite time, there exists a linear relation between queue lengths and HOL delays, where the ratio is the mean arrival rate. Hence, the delay-based MaxWeight scheme is basically equivalent to its queue-length-based counterpart, and thus achieves the optimal throughput. This property holds for the single-hop traffic, since given that the exogenous arrival processes follow the Strong Law of Large Numbers (SLLN) and the fluid limits exist, the arrival processes turn out to be deterministic processes with constant rates in the fluid limits. However, such a linear relation does not necessarily hold for the multi-hop traffic, since the packet arrival rate at a non-source node (or a relay node) is not a constant and depends on the underlying scheduler's dynamics. To this end, we investigate delay-based scheduling schemes that achieve the optimal throughput performance in multi-hop wireless networks.

Unlike previous delay-based schemes, we view packet delay as a sojourn time in the network, and re-design the delay metric of a queue as the delay difference between the queue's HOL packet and the HOL packet of its previous hop (see Eq. (34) for the formal definition). Using this new metric, we can establish a linear relation between queue lengths and delays in the fluid limit model. Then the linear relation plays the key role in showing that the proposed Delay-based Back-Pressure (D-BP) scheduling scheme is throughput-optimal in multi-hop networks.

In summary, the main contributions of our paper are as follows:

- We re-visit throughput optimality of Q-BP using fluid limit techniques. Throughput optimality of Q-BP has been originally shown using the standard Lyapunov technique in a stochastic sense. We re-derive throughput optimality of Q-BP itself using fluid limit techniques so that we can extend the analysis to D-BP using the linear relation between queue lengths and delays in the fluid limit model.
- We devise a new delay metric for D-BP and show that it achieves the optimal throughput performance in multi-hop wireless networks. Calculating a link weight as sojourn time difference of the HOL packet, we establish a linear relation between queue lengths and delays in the fluid limit model, which leads to throughput-optimality of D-BP following the same analytical procedure of Q-BP.
- We conduct extensive simulations to evaluate the performance of delay-based schedulers. Through simulations, we observe that the last packet problem can cause excessive delays for certain flows under Q-BP, while the problem is eliminated under D-BP. Further, in the case of Q-BP, even though the average delays experienced in the network may be similar to D-BP, the tail of the delay distribution could be substantially longer. We also show that, D-BP can not only achieve the same throughput

region as Q-BP, but also guarantee better fairness by scheduling the links based on delays and not starving certain flows that lack subsequent packet arrivals (or have very large inter-arrival times between groups of packet arrivals).

The remainder of the paper is organized as follows. In Section II, we present a detailed description of our system model. In Section III, we show throughput optimality of Q-BP using fluid limit techniques, and extend the analysis to D-BP in Section IV. We evaluate the performance of delay-based schedulers through simulations in Section V, and conclude our paper in Section VI.

II. SYSTEM MODEL

We consider a multi-hop wireless network described by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ denotes the set of nodes and $\mathcal{E}$ denotes the set of links. Nodes are wireless transmitters/receivers and links are wireless channels between two nodes if they can directly communicate with each other. We assume a time-slotted system with a single frequency channel. During a single time slot, multiple links that do not interfere can be active at the same time, and each active link transmits one packet during the time slot if its queue is not empty. Let $\mathcal{S}$ denote the set of flows in the network. We assume that each flow has a single, fixed, and loop-free route. The route of flow s has an $H(s)$-hop length from the source to the destination, where each k-th hop link is denoted by (s, k). Note that the assumption of single route and unit capacity is for ease of exposition, and one can easily extend the results to more general networks with multiple fixed routes and heterogeneous capacities. To specify wireless interference, we consider the k-th hop of each flow s or link-flow-pair (s, k). Let $\mathcal{P}$ denote the set of all link-flow-pairs, i.e., $\mathcal{P} \triangleq \{(s, k) \mid s \in \mathcal{S}, \ 1 \leq k \leq H(s)\}$. The set of link-flow-pairs that interfere with (s, k) can be described as

$$I(s,k) \triangleq \{(r,j) \in \mathcal{P} \mid (s,k) \text{ interferes with } (r,j), \\ \text{or } (r,j) = (s,k)\}. \quad (1)$$

Note that the interference model we adopt is general. A schedule is a set of (active or inactive) link-flow-pairs, and can be represented by a vector $\vec{M} \in \{0,1\}^{|\mathcal{P}|}$, where each link-flow-pair is set to 1 if it is active, and 0 if it is inactive, and $|\cdot|$ denotes the cardinality of a set. A schedule $\vec{M}$ is said to be *feasible* if no two link-flow-pairs of $\vec{M}$ interfere with each other, i.e., $(r,j) \notin I(s,k)$ for all (r,j), (s,k) with $M_{r,j} = 1$ and $M_{s,k} = 1$. Let $\mathcal{M}_{\mathcal{P}}$ denote the set of all feasible schedules in $\mathcal{P}$, and let $Co(\mathcal{M}_{\mathcal{P}})$ denote its convex hull.

Let $A_s(t)$ denote the number of packet arrivals at the source node of flow s at time slot t. We assume that the packet arrival processes satisfy the Strong Law of Large Numbers (SLLN): with probability 1,

$$\lim_{t \to \infty} \frac{\sum_{\tau=0}^{t-1} A_s(\tau)}{t} = \lambda_s, \quad (2)$$

for all flow $s \in \mathcal{S}$, and their fluid limits exist [15]. We call λ_s the arrival rate of flow s, and let $\vec{\lambda} \triangleq [\lambda_1, \lambda_2, \cdots, \lambda_{|\mathcal{S}|}]$ denote its vector.

Let $Q_{s,k}(t)$ denote the number of packets at the queue of (s,k) at the beginning of time slot t. Slightly abusing the notation, we also use $Q_{s,k}$ to denote the queue. We denote the queue length vector at time slot t by $\vec{Q}(t) \triangleq [Q_{s,k}(t), (s,k) \in \mathcal{P}]$, and use $\|\cdot\|$ to denote the L_1-norm of a vector, e.g., $\|\vec{Q}(t)\| = \sum_{(s,k) \in \mathcal{P}} Q_{s,k}(t)$. Let $\Pi_{s,k}(t)$ denote the service of $Q_{s,k}$ at time slot t, which takes either 1 if link-flow-pair (s,k) is active, or 0, otherwise, in our settings. We denote the actual number of packets transmitted from $Q_{s,k}$ at time slot t by $\Psi_{s,k}(t)$. Clearly, we have $\Psi_{s,k}(t) \leq \Pi_{s,k}(t)$ for all time slots $t \geq 0$. Let $P_{s,k}(t) \triangleq \sum_{i=1}^{k} Q_{s,i}(t)$ denote the summed queue length of queues up to the k-th hop for flow s. By setting $Q_{s,H(s)+1} = 0$, we have $P_{s,H(s)+1} = P_{s,H(s)}$. The queue length evolves according to the following equations:

$$Q_{s,k}(t+1) = Q_{s,k}(t) + \Psi_{s,k-1}(t) - \Psi_{s,k}(t), \qquad (3)$$

where we set $\Psi_{s,0}(t) = A_s(t)$.

Let $F_s(t)$ be the total number of packets that arrive at the source node of flow s until time slot $t \geq 0$, including those present at time slot 0, and let $\hat{F}_{s,k}(t)$ be the total number of packets that are served at $Q_{s,k}$ until time slot $t \geq 0$. We by convention set $\hat{F}_{s,k}(0) = 0$ for all link-flow-pairs $(s,k) \in \mathcal{P}$. We let $Z_{s,k,i}(t)$ denote the sojourn time of the i-th packet of $Q_{s,k}$ in the network at time slot t, where the time is measured from when the packet arrives in the network (i.e., when the packet arrives at the source node), and let $W_{s,k}(t)$ denote the sojourn time of the Head-of-Line (HOL) packet of $Q_{s,k}$ in the network at time slot t, i.e., $W_{s,k}(t) = Z_{s,k,1}(t)$. We set $W_{s,0}(t) = 0$ and $W_{s,H(s)+1}(t) = W_{s,H(s)}(t)$, for all $s \in \mathcal{S}$. Further, if $Q_{s,k}(t) = 0$, we set $W_{s,k}(t) = W_{s,k-1}(t)$. Letting $U_{s,k}(t) \triangleq t - W_{s,k}(t)$ denote the time when the HOL packet of $Q_{s,k}$ arrives in the network, we have that

$$U_{s,k}(t) = \inf\{\tau \leq t \mid F_s(\tau) > \hat{F}_{s,k}(t)\}, \text{ for all } t \geq 0. \quad (4)$$

We next define the stability of a network as follows.

Definition 1: A network of queues is said to be *stable* if,

$$\limsup_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \mathbf{E}[\|\vec{Q}(\tau)\|] < \infty. \qquad (5)$$

We define the *throughput region* of a scheduling policy as the set of rates, for which the network remains stable under this policy. Further, we define the *optimal throughput region* (or *stability region*) as the union of the throughput regions of all possible scheduling policies. The optimal throughput region Λ^* can be presented as

$$\Lambda^* \triangleq \{\vec{\lambda} \mid \exists \vec{\phi} \in Co(\mathcal{M}_{\mathcal{P}}) \text{ s.t. } \lambda_s \leq \phi_{s,k}, \text{ for all } (s,k) \in \mathcal{P}\}. \qquad (6)$$

An arrival rate vector is strictly inside Λ^*, if the inequalities above are all strict.

III. Throughput Optimality of Q-BP Using Fluid Limits

It has been shown in [1] that Q-BP stabilizes the network for any feasible arrival rate vector using stochastic Lyapunov techniques. Specifically, we can use a quadratic-form Lyapunov function to show that the function has a negative drift under Q-BP when queue lengths are large enough. In this section, we re-visit throughput optimality of Q-BP using fluid limit techniques. The analysis will be extended later to prove throughput optimality of the delay-based back-pressure algorithm.

To begin with, we define the *queue differential* $\Delta Q_{s,k}(t)$ as

$$\Delta Q_{s,k}(t) \triangleq Q_{s,k}(t) - Q_{s,k+1}(t), \qquad (7)$$

and specify the back-pressure algorithm based on queue lengths as follows.

Queue-length-based Back-Pressure (Q-BP) algorithm:

$$\vec{M}^* \in \operatorname{argmax}_{\vec{M} \in \mathcal{M}_{\mathcal{P}}} \sum_{(s,k) \in \mathcal{P}} \Delta Q_{s,k}(t) \cdot M_{s,k}. \quad (8)$$

The algorithm needs to solve a MaxWeight problem with weights as queue differentials, and ties can be broken arbitrarily if there are more than one schedules that have the largest weight sum.

We establish the fluid limits of the system and prove throughput optimality of Q-BP using fluid limit techniques.

A. Fluid limits

We define the process describing the behavior of the underlying system as $\mathcal{X} = (\mathcal{X}(t), t = 0, 1, 2, \cdots)$, where

$$\mathcal{X}(t) \triangleq \left((Z_{s,k,1}(t), \cdots, Z_{s,k,Q_{s,k}(t)}(t)), (s,k) \in \mathcal{P} \right). \quad (9)$$

The process $\mathcal{X}$ forms a discrete time Markov chain, if a scheduling decision is based on the information of the current time slot only. Clearly, $\mathcal{X}$ forms a Markov chain under Q-BP. Motivated by Definition 1, we define the norm of $\mathcal{X}(t)$ as

$$\|\mathcal{X}(t)\| \triangleq \|\vec{Q}(t)\|. \qquad (10)$$

Let $\mathcal{X}^{(x_n)}$ denote a process $\mathcal{X}$ with an initial configuration such that

$$\|\mathcal{X}^{(x_n)}(0)\| = x_n. \qquad (11)$$

All the processes of $\mathcal{X}^{(x_n)}$ satisfy the properties in the original system $\mathcal{X}$.

The following Lemma was derived in [16] for continuous time countable Markov chains, and it follows from more general results in [17] for discrete time countable Markov chains.

Lemma 1: Suppose there exists an integer $T > 0$ such that, for any sequence of processes $\{\mathcal{X}^{(x_n)}\}$, we have that,

$$\lim_{x_n \to \infty} \mathbf{E}\left[\frac{1}{x_n} \|\mathcal{X}^{(x_n)}(x_n T)\|\right] = 0, \qquad (12)$$

then the system is stable.

A stability criteria of (12) leads to a fluid limit approach [15] to the stability problem of queueing systems. Hence, we start our analysis by establishing the *fluid limit model* as in [11], [15]. We define the process $\mathcal{Y} \triangleq (A, F, \hat{F}, Q, P, \Pi, \Psi, W, U)$, and it is clear that, a sample path of $\mathcal{Y}$ uniquely defines the sample path of $\mathcal{X}$. We then extend the definition of $Y = A$, F, $\hat{F}$, Q, P, Π, Ψ, W and U to continuous time domain as $Y(t) \triangleq Y(\lfloor t \rfloor)$ for time $t \geq 0$. Note that, $Y(t)$ is right continuous having left limits.

As in [11], we extend the definition of $F_s^{(x_n)}(t)$ to the negative interval $t \in [-x_n, 0)$ by assuming that the packets present in the initial state $\mathcal{X}^{(x_n)}(0)$ arrived in the past at some of the time instants $-(x_n - 1), -(x_n - 2), \cdots, 0$, according to their delays in the state $\mathcal{X}^{(x_n)}(0)$. By this convention, $F_s^{(x_n)}(-x_n) = 0$ for all $s \in \mathcal{S}$ and x_n, and $\sum_{s \in S} F_s^{(x_n)}(0) = x_n$, for all x_n.

Then, using the techniques of Theorem 4.1 of [15], we can show that, for almost all sample paths and for all positive sequence $x_n \to \infty$, there exists a subsequence x_{n_j} with $x_{n_j} \to \infty$ such that, for all $s \in \mathcal{S}$ and all $(s,k) \in \mathcal{P}$, the following convergences hold *uniformly over compact (u.o.c)* interval:

$$\frac{1}{x_{n_j}} \int_0^{x_{n_j}t} A_s^{(x_{n_j})}(\tau) d\tau \to \lambda_s t, \tag{13}$$

$$\frac{1}{x_{n_j}} F_s^{(x_{n_j})}(x_{n_j} t) \to f_s(t), \tag{14}$$

$$\frac{1}{x_{n_j}} \hat{F}_{s,k}^{(x_{n_j})}(x_{n_j} t) \to \hat{f}_{s,k}(t), \tag{15}$$

$$\frac{1}{x_{n_j}} Q_{s,k}^{(x_{n_j})}(x_{n_j} t) \to q_{s,k}(t), \tag{16}$$

$$\frac{1}{x_{n_j}} P_{s,k}^{(x_{n_j})}(x_{n_j} t) \to p_{s,k}(t), \tag{17}$$

$$\frac{1}{x_{n_j}} \int_0^{x_{n_j}t} \Pi_{s,k}^{(x_{n_j})}(\tau) d\tau \to \int_0^t \pi_{s,k}(\tau) d\tau, \tag{18}$$

$$\frac{1}{x_{n_j}} \int_0^{x_{n_j}t} \Psi_{s,k}^{(x_{n_j})}(\tau) d\tau \to \int_0^t \psi_{s,k}(\tau) d\tau, \tag{19}$$

and similarly, the following convergences (which are denoted by "$\Rightarrow$") hold at every continuous point of the limit function:

$$\frac{1}{x_{n_j}} W_{s,k}^{(x_{n_j})}(x_{n_j} t) \Rightarrow w_{s,k}(t), \tag{20}$$

$$\frac{1}{x_{n_j}} U_{s,k}^{(x_{n_j})}(x_{n_j} t) \Rightarrow u_{s,k}(t). \tag{21}$$

At almost all points $t \in [0, \infty)$, the derivatives of these limit functions exist. We call such points *regular* time. Moreover, the limits satisfy that

$$\sum_{s \in \mathcal{S}} f_s(0) = 1, \tag{22}$$

$$p_{s,k}(t) = \sum_{i=1}^k q_{s,i}(t), \tag{23}$$

$$p_{s,k}(t) = f_s(t) - \hat{f}_{s,k}(t), \tag{24}$$

$$f_s(t) = f_s(0) + \lambda_s t, \tag{25}$$

$$u_{s,k}(t) = t - w_{s,k}(t), \tag{26}$$

$$\psi_{s,k}(t) \leq \pi_{s,k}(t), \tag{27}$$

$$\Delta q_{s,k}(t) = q_{s,k}(t) - q_{s,k+1}(t), \tag{28}$$

$$\frac{d}{dt} q_{s,k}(t) = \begin{cases} \psi_{s,k-1}(t) - \pi_{s,k}(t), & q_{s,k}(t) > 0, \\ [\psi_{s,k-1}(t) - \pi_{s,k}(t)]^+, & q_{s,k}(t) = 0, \end{cases} \tag{29}$$

where $[z]^+ \triangleq \max(z, 0)$, and we set $\psi_{s,0} = \pi_{s,0} = \lambda_s$.

It is clear from (8) that Q-BP will not schedule link-flow-pair (s,k) if $Q_{s,k}(t) - Q_{s,k+1}(t) < 0$. This implies that, if

$$Q_{s,k}(t) \geq Q_{s,k+1}(t) - 2 \tag{30}$$

initially holds for all (s,k) at time slot 0, then the inequality holds for every time slot $t \geq 0$. This further implies that

$$q_{s,k}(t) \geq q_{s,k+1}(t), \text{ i.e., } \Delta q_{s,k}(t) \geq 0, \tag{31}$$

for all (scaled) time $t \geq 0$. Without loss of generality, we assume that, at time slot 0, all queues on each route are empty, except for the first queue, then it follows that (30) holds for all (scaled) time $t \geq 0$, and thus, $\Delta q_{s,k}(t) \geq 0$ holds, for all $t \geq 0$.

B. Throughput optimality of Q-BP

Proposition 2: Q-BP can support any traffic with arrival rate vector that is strictly inside Λ^*.

Proof: We prove the stability using the standard Lyapunov technique. We consider a quadratic-form Lyapunov function in the fluid limit model of the system, and show that it has a negative drift, which implies that the fluid limit model and thus the original system is stable.

Let $V(\vec{q}(t))$ denote the Lyapunov function defined as

$$V(\vec{q}(t)) \triangleq \frac{1}{2} \sum_{(s,k) \in \mathcal{P}} (q_{s,k}(t))^2. \tag{32}$$

Suppose $\vec{\lambda}$ is strictly inside Λ^*, then there exists a vector $\vec{\phi} \in Co(\mathcal{M}_\mathcal{P})$ such that $\vec{\lambda} < \vec{\phi}$, i.e., $\lambda_s < \phi_{s,k}$ for all $(s,k) \in \mathcal{P}$. Since $\vec{q}(t)$ is differentiable, for any regular time $t \geq 0$ such that $V(\vec{q}(t)) > 0$, we can obtain the derivative of $V(\vec{q}(t))$ as

$$
\begin{aligned}
&\frac{D^+}{dt^+} V(\vec{q}(t)) \\
&= \sum_{(s,k) \in \mathcal{P}} q_{s,k}(t) \cdot (\psi_{s,k-1}(t) - \pi_{s,k}(t)) \\
&\leq \sum_{(s,k) \in \mathcal{P}} q_{s,k}(t) \cdot (\pi_{s,k-1}(t) - \pi_{s,k}(t)) \\
&= \sum_{(s,k) \in \mathcal{P}} \Delta q_{s,k}(t) \cdot \lambda_s - \sum_{(s,k) \in \mathcal{P}} \Delta q_{s,k}(t) \cdot \pi_{s,k}(t) \\
&= \sum_{(s,k) \in \mathcal{P}} \Delta q_{s,k}(t) \cdot (\lambda_s - \phi_{s,k}) \\
&\quad + \sum_{(s,k) \in \mathcal{P}} \Delta q_{s,k}(t) \cdot (\phi_{s,k} - \pi_{s,k}(t)),
\end{aligned} \tag{33}
$$

where $\frac{D^+}{dt^+} V(\vec{q}(t)) = \lim_{\delta \downarrow 0} \frac{V(\vec{q}(t+\delta)) - V(\vec{q}(t))}{\delta}$, and the first equality and the inequality are from (29) and (27), respectively.

Note that in the final result of (33), we obtain that i) the first term is negative because i) $\vec{\lambda} < \vec{\phi}$, and $\Delta q_{s,k}(t) \geq 0$ for all $(s,k) \in \mathcal{P}$ and $\Delta q_{s',k'}(t) > 0$ for at least one link-flow-pair (s', k') since $V(\vec{q}(t)) > 0$, and that ii) the second term becomes non-positive since Q-BP chooses schedules that maximize the queue differential weight sum (8), its fluid limit $\vec{\pi}(t)$ satisfies that

$$\vec{\pi}(t) \in \text{argmax}_{\vec{\phi} \in Co(\mathcal{M}_\mathcal{P})} \sum_{(s,k) \in \mathcal{P}} \Delta q_{s,k}(t) \cdot \phi_{s,k},$$

which implies that

$$\sum_{(s,k) \in \mathcal{P}} \Delta q_{s,k}(t) \cdot \phi_{s,k} \leq \sum_{(s,k) \in \mathcal{P}} \Delta q_{s,k}(t) \cdot \pi_{s,k}(t),$$

for all $\vec{\phi} \in Co(\mathcal{M}_\mathcal{P})$. Therefore, we have $\frac{D^+}{dt^+} V(\vec{q}(t)) < 0$ and the fluid limit model of the system is stable, which implies that the original system is also stable by Theorem 4.2 of [15]. ∎

IV. THROUGHPUT OPTIMALITY OF D-BP

A. Algorithm description

Next, we develop Delay-based Back-Pressure (D-BP) policy that can establish a linear relation between queue lengths and delays in the fluid limit model. The idea has appeared first in [11] for single-hop networks. However, when packets travel multiple hops before leaving the system, the analytical approach in [11] (i.e., using HOL delay in the queue as the metric) cannot capture queueing dynamics of multi-hop traffic and the resultant solutions cannot guarantee the linear relation. This is because the arrival rate of a relay node is not a constant and depends on the system dynamics (i.e., depends on the underlying scheduling policies). In this section, we carefully design link weights using a new delay metric, and re-establish the linear relation between queue lengths and delays under multi-hop traffic.

Recall that $W_{s,k}(t)$ denotes the sojourn time of the HOL packet of queue $Q_{s,k}(t)$ in the network, where the time is measured from when the packet arrives in the network. We define the delay metric $\hat{W}_{s,k}(t)$ as

$$\hat{W}_{s,k}(t) \triangleq W_{s,k}(t) - W_{s,k-1}(t), \tag{34}$$

and also define the *delay differential* as

$$\Delta \hat{W}_{s,k}(t) \triangleq \hat{W}_{s,k}(t) - \hat{W}_{s,k+1}(t). \tag{35}$$

The relations between these delay metrics are illustrated in Fig. 1. We specify the back-pressure algorithm with the new delay metric as follows.

Delay-based Back-Pressure (D-BP) algorithm:

$$\vec{M}^* \in \text{argmax}_{\vec{M} \in \mathcal{M}_\mathcal{P}} \sum_{(s,k) \in \mathcal{P}} \Delta \hat{W}_{s,k}(t) \cdot M_{s,k}. \tag{36}$$

D-BP computes the weight of (s,k) as the delay differential $\Delta \hat{W}_{s,k}(t)$ and solves the MaxWeight problem, i.e., finds a set of non-interfering link-flow-pairs that maximizes weight sum. Ties can be broken arbitrarily if there are more than one schedules that have the largest weight sum. An intuitive interpretation of the new delay metric $\hat{W}_{s,k}(t)$ is as follows. Note that the queue length $Q_{s,k}(t)$ is roughly the number of packets arriving at the source of flow s during the time slots between $[U_{s,k}(t), U_{s,k}(t) + \hat{W}_{s,k}(t))$, and $Q_{s,k}(t)$ is in the order of $\lambda_s \hat{W}_{s,k}(t)$ when $\hat{W}_{s,k}(t)$ is large. Hence, a large $\hat{W}_{s,k}(t)$ implies a large queue length $Q_{s,k}(t)$, and similarly, a large delay differential $\Delta \hat{W}_{s,k}(t)$ implies a large queue length differential $\Delta Q_{s,k}(t)$. Therefore, being favorable to the delay weight sum in (36) is in some sense "equivalent" to being favorable to the queue length weight sum in (8) as Q-BP. We later formally establish the linear relation between the fluid limits of queue lengths and delays in Section IV-B.

Clearly, D-BP also will not schedule link-flow-pair (s,k) if $\hat{W}_{s,k}(t) - \hat{W}_{s,k+1}(t) < 0$. Let $B_{s,k}(t)$ denote the inter-arrival time between the HOL packet of $Q_{s,k}(t)$ and the packet that arrives immediately after it. The aforementioned operation of D-BP implies that, if inequality

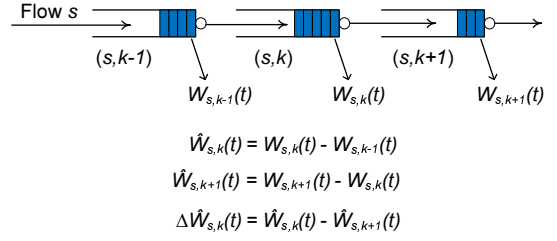$$\hat{W}_{s,k}(t) \geq \hat{W}_{s,k+1}(t) - 2B_{s,k}(t), \tag{37}$$

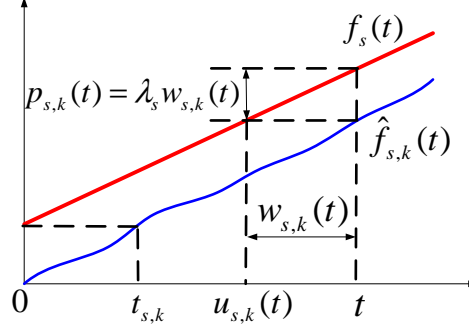Fig. 1. Delay differentials using new delay metric.

Fig. 2. Linear relation between queue lengths and delays in the fluid limits.

initially holds for all (s,k) at time slot 0, then the inequality holds for all time slot $t \geq 0$. This further leads to

$$\hat{w}_{s,k}(t) \geq \hat{w}_{s,k+1}(t), \text{ i.e., } \Delta \hat{w}_{s,k}(t) \geq 0, \tag{38}$$

for all (scaled) time $t \geq 0$, in the fluid limits, since $\frac{1}{x_{n_j}} B_{s,k}^{(x_{n_j})}(x_{n_j} t) \to 0$, as $x_{n_j} \to \infty$, otherwise we will arrive a contradiction to the fact that the arrival process satisfies the Strong Law of Large Numbers. Recall that we assume that all queues on each route are empty, except for the first queue at time slot 0, then (37) and (38) follow.

B. Analysis of throughput performance

We first establish the linear relation between the fluid limits of queue lengths and delays in the following lemma. We will use the lemma later to show that D-BP achieves the optimal throughput.

Lemma 3: For any fixed $t_{s,k} > 0$, for any link-flow-pair $(s,k) \in \mathcal{P}$, the two conditions $u_{s,k}(t_{s,k}) > 0$ and $\hat{f}_{s,k}(t_{s,k}) > f_s(0)$ are equivalent. Further, if these conditions hold, we have

$$p_{s,k}(t) = \lambda_s w_{s,k}(t), \tag{39}$$

$$q_{s,k}(t) = \lambda_s \hat{w}_{s,k}(t), \tag{40}$$

for all $t \geq t_{s,k}$, with probability 1.

Fig. 2 describes the relations between the variables.

Proof: Since the first part, i.e., the two conditions are equivalent, is straightforward from the definition of fluid limits and (4), we focus on the second part, i.e., if $\hat{f}_{s,k}(t_{s,k}) > f_s(0)$, then (39) and (40) follow.

Suppose that

$$\hat{f}_{s,k}(t_{s,k}) > f_s(0). \tag{41}$$

Then, by definition of $u_{s,k}(t)$, we have

$$\hat{f}_{s,k}(t) = f_s(u_{s,k}(t)), \qquad (42)$$

for all $t \geq t_{s,k}$. Hence, we obtain that

$$
\begin{aligned}
p_{s,k}(t) &= f_s(t) - \hat{f}_{s,k}(t) \\
&\overset{(a)}{=} (f_s(0) + \lambda_s t) - (f_s(0) + \lambda_s u_{s,k}(t)) \qquad (43) \\
&\overset{(b)}{=} \lambda_s w_{s,k}(t),
\end{aligned}
$$

where (a) is from (42) and (25), and (b) is from (26). Further, (40) follows from (23) and the fluid limit version of (34). ∎

We emphasize the importance of Lemma 3. Lemma 3 implies that after a finite time (i.e., $\max_{(s,k)\in\mathcal{P}} t_{s,k}$), queue lengths are λ_s times delays in the fluid limit model. Then the schedules of D-BP are very similar to those of Q-BP, which implies that D-BP achieves the optimal throughput region Λ^*. In the following, we show that such a finite time exists.

Lemma 4: Consider a system under the D-BP policy. For $\vec{\lambda}$ strictly inside Λ^*, there exists a time $T > 0$ such that the fluid limits satisfy the following property with probability 1,

$$\hat{f}_{s,k}(T) > f_s(0), \qquad (44)$$

for all link-flow-pairs $(s,k) \in \mathcal{P}$.

We can prove Lemma 4 by induction following the techniques described in Lemma 7 of [11]. We omit the proof and refer readers to our online technical report [18]. We next outline an informal discussion, which highlights the main idea of the proof. First, we consider the base case. D-BP chooses one of the feasible schedules in $\mathcal{M}_{\mathcal{P}}$ (we omit the term "feasible" in the following, whenever there is no confusion) at each time slot. Each schedule receives a fraction of the total time and there must exist a schedule that gets at least $\frac{1}{|\mathcal{M}_{\mathcal{P}}|}$ fraction of the total time. Thus, after a large enough time $T_1 > 0$, there must exist a schedule $\vec{M}^*$ that is chosen for at least $\frac{T_1}{|\mathcal{M}_{\mathcal{P}}|}$ amount of time. The amount of initial packets of $\vec{M}^*$ is bounded from (22), thus, for a large enough T_1, all initial packets of at least one link-flow-pair of $\vec{M}^*$ must be completely served, i.e., $\hat{f}_{s,k}(T_1) > f_s(0)$, for at least one (s,k) with $M^*_{s,k} = 1$.

Next, we consider the inductive step. Suppose there exists a $T_l > 0$, such that for at least one subset $S_l \subset \mathcal{P}$ of cardinality l, we have

$$\hat{f}_{s,k}(T_l) > f_s(0), \qquad (45)$$

for all $(s,k) \in S_l$. Then there exists $T_{l+1} \geq T_l$ such that

$$\hat{f}_{s,k}(T_{l+1}) > f_s(0), \qquad (46)$$

holds for all link-flow-pairs (s,k) within at least one subset $S_{l+1} \subset \mathcal{P}$ of cardinality $l+1$. Note that, if $(s,k) \in S_l$, then $(s,i) \in S_l$ for $1 \leq i \leq k$. Let

$$
\begin{aligned}
S_l^* \triangleq \{(r,j) \mid &(r,j) \notin S_l, (r,j-1) \in S_l, \text{ for } j > 1; \\
&\text{or } (r,j) \notin S_l, \text{ for } j = 1\}
\end{aligned} \qquad (47)
$$

denote the set of link-flow-pairs (r,j) such that $(r,j) \in \mathcal{P} \backslash S_l$ is the closest hop to the source of r. To avoid unnecessary

complications, we discuss the induction step for $l = 1$. The generalization for $l > 1$ is straightforward. We show that for given S_1 and T_1, there exists a finite $T_2 \geq T_1$ such that (46) with T_2 holds for at least two different link-flow-pairs.

Let $(\hat{s}, \hat{k})$ denote the link-flow-pair that satisfies (45) with T_1. Since $(\hat{s}, \hat{k}) \in S_l$ implies $(\hat{s}, i) \in S_l$ for all $1 \leq i \leq \hat{k}$, we must have $\hat{k} = 1$ and $S_1 = \{(\hat{s}, 1)\}$. From (47), we have that

$$S_1^* = \{(r,1) \mid r \in \mathcal{S} \backslash \{\hat{s}\}\} \cup N_{\hat{s}}, \qquad (48)$$

where $N_{\hat{s}} = \{(\hat{s}, 2)\}$ if $H(\hat{s}) > 1$, and $N_{\hat{s}} = \emptyset$ if $H(\hat{s}) = 1$. We discuss only the case that $H(\hat{s}) > 1$, and the other case can be easily shown following the same line of analysis. Now suppose that

$$\hat{f}_{r,j}(t) \leq f_r(0), \text{ for all } (r,j) \in \mathcal{P} \backslash S_1, \text{ and all } t \geq 0, \quad (49)$$

i.e., for all the link-flow-pairs except those of S_1, the total amount of service up to time t is no greater than the amount of the initial packets for all $t \geq 0$. We show that this assumption leads to a contradiction, which completes the inductive step, and we prove the lemma.

From the base case and Lemma 3, we have $q_{\hat{s},1}(t) = \lambda_{\hat{s}} \hat{w}_{\hat{s},1}(t)$ for all $t \geq T_1$. We view the subset of links S_1 as a generalized system, and consider the time slots when there is at least one packet transmission from the outside of S_1, i.e., $(r,j) \in \mathcal{P} \backslash S_1$. For each of such time slot, we say that the time slot is *unavailable* to S_1.

1) The number of such unavailable time slots is bounded from the above by x_{n_j}, since at every such time slot, at least one initial packet will be transmitted and the total number of initial packets is bounded by $\|\vec{Q}(0)\| = x_{n_j}$ from (11). Hence, the amount of (scaled) time unavailable to S_1 is bounded by $\|\vec{q}(0)\| = 1$.

2) Since the amount of (scaled) time unavailable to S_1 is bounded, there exists a sufficiently large $t \geq T_1$ such that the fraction of time that is given to $(r,j) \in \mathcal{P} \backslash S_1$ is negligible, and we must have $\hat{w}_{\hat{r},\hat{j}}(t) = \Theta(1)$[1] and $\Delta \hat{w}_{\hat{r},\hat{j}}(t) = \Theta(1)$ for $(\hat{r},\hat{j}) \in \mathcal{P} \backslash (S_1 \bigcup S_1^*)$.

3) Then, we can restrict our focus on the generalized system S_1 to time $t \geq T_1$, and ignore the time that is unavailable to S_1. Then Q-BP and D-BP are in some sense "equivalent" in the generalized system S_1 for $t \geq T_1$ with the following properties: First, Q-BP will stabilize the system if the arrival rate vector is strictly inside Λ^*. Second, since the linear relation (40) holds for all link-flow-pairs in S_1 from Lemma 3, D-BP will schedule links similar to Q-BP and also stabilizes the generalized system S_1.

4) Now let us focus on S_1^*. Link-flow-pairs in S_1^* must have some initial packets at $t \geq T_1$ because $S_1 \cap S_1^* = \emptyset$. On the other hand, the generalized network S_1 is stable. This implies that the delay metrics of link-flow-pairs in S_1^* should increase at the same order as we increase t,

[1]We use the standard order notation: $g(n) = o(f(n))$ implies $\lim_{n\to\infty}(g(n)/f(n)) = 0$; and $g(n) = \Theta(f(n))$ implies $c_1 \leq \lim_{n\to\infty}(g(n)/f(n)) \leq c_2$ for some constants c_1 and c_2.

i.e., $\hat{w}_{r^*,j^*}(t) = \Theta(t)$ for $(r^*, j^*) \in S_1^*$. Then we have $\Delta \hat{w}_{r^*,j^*}(t) = \Theta(t)$, since $\hat{w}_{r^*,j^*+1}(t) = \Theta(1)$ from $(r^*, j^* + 1) \in \mathcal{P} \backslash (S_1 \bigcup S_1^*)$ and 2). Since the delay differentials $\Delta \hat{w}_{s,k}(t)$ for all $(s, k) \in S_1$ and $\Delta \hat{w}_{\hat{r},\hat{j}}(t)$ for all $(\hat{r}, \hat{j}) \in \mathcal{P} \backslash (S_1 \cup S_1^*)$ are bounded above from stability of S_1 and 2), respectively, D-BP will choose some of the link-flow-pairs in S_1^* for most of time for a sufficiently large t. This implies that the amount of time unavailable to S_1 is $\Theta(t)$, which conflicts with our previous statement that the fraction of time that is given to $(r, j) \in \mathcal{P} \backslash S_1$ is negligible.

As mentioned earlier, we omit the detailed proof here and refer readers to our online technical report [18].

The following proposition shows throughput optimality of D-BP.

Proposition 5: D-BP can support any traffic with arrival rate vector that is strictly inside Λ^*.

Proof: We show the stability using fluid limits and standard Lyapunov techniques. From Lemmas 3 and 4, we obtain the key property for proving throughput optimality of D-BP in Eq. (40), i.e., after a finite time, there is a linear relation between queue lengths and delays in the fluid limit model. We start with the following quadratic-form Lyapunov function,

$$V(\vec{q}(t)) \triangleq \frac{1}{2} \sum_{(s,k) \in \mathcal{P}} \frac{(q_{s,k}(t))^2}{\lambda_s}. \tag{50}$$

Following the line of analysis in the proof of Proposition 2, we can show that the Lyapunov function has a negative drift if the underlying scheduler maximizes $\sum_{s,k} \frac{\Delta q_{s,k}(t)}{\lambda_s} \cdot \pi_{s,k}(t)$. Now applying the linear relation (40), we can observe that D-BP satisfies such a condition, and obtain the results. We omit the detailed proof. ∎

Although D-BP operates efficiently and achieves the optimal throughput region, it is difficult to implement in practice due to centralized operations and high computational complexity. Therefore, we are interested in simpler approximations to D-BP that can achieve a guaranteed fraction of the optimal performance. The delay-based greedy maximal algorithm[2] is a good candidate algorithm. We can characterize the throughput performance of the delay-based greedy scheme combining our results along with the techniques used in [19], [20], and show that it is as efficient as its queue-length-based counterpart, i.e., the queue-length-based greedy maximal algorithm.

V. NUMERICAL RESULTS

In this section, we first highlight the *last packet* problem of the queue-length-based back-pressure algorithm. The last packet problem implies that flows that lack packet arrivals at subsequent times may experience excessive delay under Q-BP, which is later confirmed in the simulations. We compare

(a) "H"-type network topology

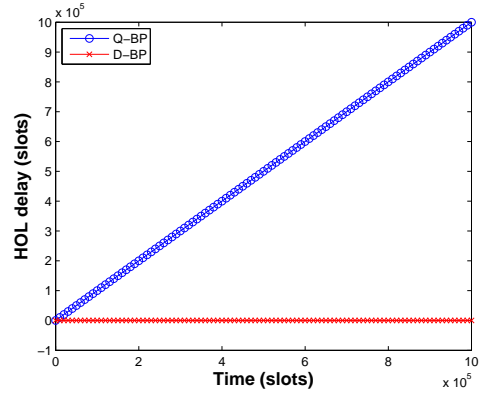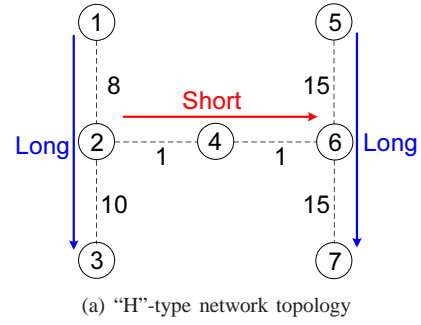(b) HOL delay of short flow $(2 \rightarrow 4 \rightarrow 6)$ when $\lambda = 3$

Fig. 3. Illustration of the last packet problem under Q-BP.

throughput and delay performance of Q-BP and D-BP in a grid network topology under the 2-hop[3] interference model.

We first show the last packet problem of Q-BP through simulations. We observe that several last packets of a short flow that carry a finite amount of data may get stuck, which could cause excessive delay. We consider a scenario consisting of 7 nodes and 6 links as shown in Fig. 3(a), where nodes are represented by circles and links are represented by dashed lines with link capacity[4]. We assume a time-slotted system. We establish three flows: one short flow $(2 \rightarrow 4 \rightarrow 6)$ and two long flows $(1 \rightarrow 2 \rightarrow 3)$ and $(5 \rightarrow 6 \rightarrow 7)$. The short flow arrives at the network with a finite amount of packets at time 0, and the number of packets follows Poisson distribution with mean rate 10. The long flows have an infinite amount of data and keep injecting packets at the source nodes following Poisson distribution with mean rate λ at each time slot. Numerical calculation shows that the feasible rate under the 2-hop interference should satisfy that $\lambda \leq 4.44$. We conduct our simulation for 10^6 time slots, and plot time traces of HOL delay of the short flow when $\lambda = 3$. Fig. 3(b) illustrates the results that the delay linearly increases with time under Q-BP, which implies that several last packets of the

---

[2]A greedy maximal algorithm finds its schedule in a decreasing order of weight (e.g., queue length or delay) conforming to the underlying interference constraints.

[3]In the 2-hop interference model, two links within a 2-hop "distance" interfere with each other. Note that the interference model (Eq. (1)) in the problem setup is general. We consider the 2-hop interference model in the simulations, as it is often used to model the ubiquitous IEEE 802.11 DCF (Distributed Coordination Function) wireless networks [21]–[23].

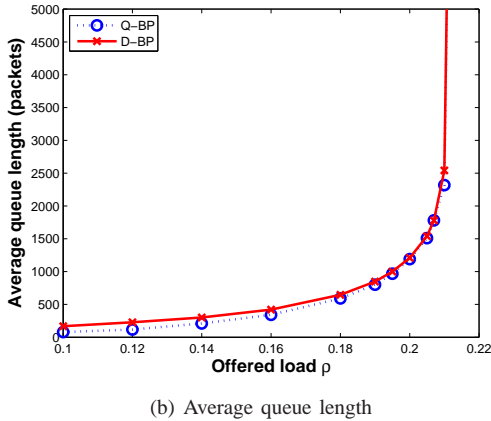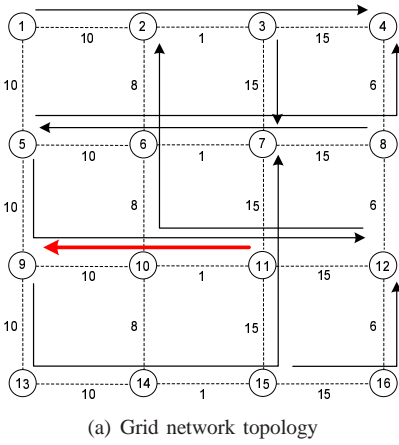[4]Unit of link capacity is packets per time slot.

(a) Grid network topology

Fig. 5. Delay distribution of the bursty flow under $\rho = 0.2$.

(b) Average queue length

Fig. 4. Performance of scheduling algorithms for multi-hop traffic following Poisson distribution.

Fig. 6. Mean delay, top 1% and top 5% largest delay of the bursty flow over offered loads.

short flow get excessively delayed. On the other hand, D-BP succeeds in serving the short flow and keeps the delay close to 0. This also implies that certain flows whose queue lengths do not increase because of lack of future arrivals (or whose inter-arrival times between groups of packets are very large) may experience a large delay under Q-BP, which will be confirmed in the following simulations.

Next, we evaluate the throughput of different schedulers in a grid network that consists of 16 nodes and 24 links as shown in Fig. 4(a), where nodes and links are represented by circles and dashed lines, respectively, with link capacity. We establish 9 multi-hop flows that are represented by arrows. Let $\lambda_1 = 0.1$ and $\lambda_2 = 1$. At each time slot, there is a file arrival with probability $p = 0.01$ for flow ($11 \rightarrow 10 \rightarrow 9$) (represented by the red thick arrow in Fig. 4(a)), and the file size follows Poisson distribution with mean rate[5] $\rho\lambda_1/p$. Note that flow ($11 \rightarrow 10 \rightarrow 9$) has bursty arrivals with a small mean rate (we simply call it the bursty flow in the following part). All the other 8 flows have packet arrivals following Poisson

distribution with mean rate $\rho\lambda_2$ at each time slot. Although these flows share the same stochastic property with an identical mean arrival rate $\rho\lambda_2$, uniform patterns of traffic are avoided by carefully setting the link capacities differently and placing the flows with different number of hops in an asymmetric manner.

We evaluate the scheduling performance by measuring average total queue lengths in the network over time. Fig. 4(b) illustrates average queue lengths under different offered loads to examine the performance limits of scheduling schemes. Each result represents an average of 10 simulation runs with independent stochastic arrivals, where each run lasts for 10^6 time slots. Since the optimal throughput region is defined as the set of arrival rates under which queue lengths remain finite (see Definition 1), we can consider the traffic load, under which the queue length increases rapidly, as the boundary of the optimal throughput region. Fig. 4(b) shows that D-BP achieves the same throughput region as Q-BP, thus supports the theoretical results of throughput optimality.

Although Q-BP and D-BP perform similarly in terms of average queue length (or average delay) over the network, the tail of the delay distribution of Q-BP could be substantially longer because certain flows are starved. This could cause enormous unfairness between flows, resulting in very poor

[5]Note that given the network topology, it is hard to find the exact boundary of the optimal throughput region of scheduling policies in a closed form. Hence, we probe the boundary by scaling the amount of traffic. After we choose $\vec{\lambda}$, which determines the direction of traffic load vector, we run our simulations with traffic load $\rho\vec{\lambda}$ changing ρ, which scales the traffic loads.
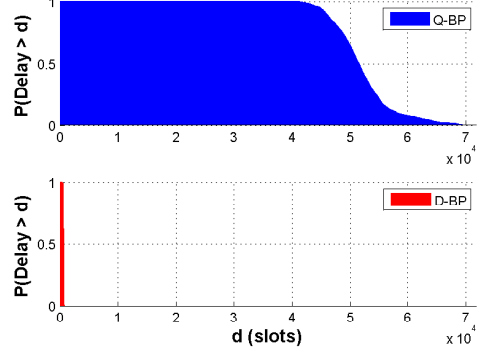
QoS for certain flows. Note that although a bursty flow is a long flow that has an infinite amount of data, the arrivals occur in a dispersed manner (i.e., the inter-arrival times between groups of packets are very large) and we can view this bursty flow as consisting of many short flows. Thus, we expect that the bursty flow may experience a very large delay under Q-BP due to lack of subsequent packet arrivals over long periods of time that does not allow the queue-lengths to grow and thus contributes to the long tail of the delay distribution. However, this phenomenon may not manifest itself in terms of a higher average delay for Q-BP, as can be observed in Fig. 4(b), because the amount of data corresponding to the bursty flow in the simulation is small compared to the other flows. On the other hand, D-BP can achieve better fairness by scheduling the links based on delays and not starving bursty or variable flows. We confirm this in the following observations.

We now illustrate the effectiveness of using D-BP over Q-BP in terms of how each scheme affects the delay distribution of bursty flows. We plot the delay distribution of the bursty flow in Fig. 5 under $\rho = 0.2$. It reveals that the tail of the delay distribution under D-BP vanishes much faster than Q-BP. Further, we plot the mean delay, top[6] 1% and top 5% largest delays of the bursty flow over offered loads in Fig. 6. All these delays under D-BP are substantially less than under Q-BP, which implies that D-BP successfully eliminates the excessive packet delays. The top 0.1% largest delays of the whole network demonstrate similar behaviors in Fig. 6 and the results are omitted. This confirms that, Q-BP causes a substantially long tail for the delay distribution of the network due to the starvation of the bursty flow, while D-BP overcomes this and achieves better fairness among the flows by scheduling the links based on delays.

VI. CONCLUSION

In this paper, we develop a throughput-optimal delay-based back-pressure scheme for multi-hop wireless networks. We introduce a new delay metric suitable for multi-hop traffic and establish a linear relation between queue lengths and delays in the fluid limit model, which plays a key role in the performance analysis and proof of throughput-optimality. Delay-based schemes provide a simple way around the well-known last packet problem that plagues the queue-length-based schedulers, and avoid flow starvation. As a result, the excessively long delays that could be experienced by certain flows under the queue-length-based scheduling schemes are eliminated without any loss of throughput.

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] X. Lin, N. Shroff, and R. Srikant, "A tutorial on cross-layer optimization in wireless networks," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 8, pp. 1452–1463, Aug. 2006.

[3] A. Warrier, S. Janakiraman, S. Ha, and I. Rhee, "DiffQ: Practical differential backlog congestion control for wireless networks," in *Proc. of INFOCOM*, 2009.

[4] A. Sridharan, S. Moeller, and B. Krishnamachari, "Implementing Backpressure-based Rate Control in Wireless Networks," in *Information Theory and Applications Workshop*, 2009.

[5] P. van de Ven, S. Borst, and S. Shneer, "Instability of MaxWeight Scheduling Algorithms," in *Proc. IEEE INFOCOM*, 2009, pp. 1701–1709.

[6] S. Liu, L. Ying, and R. Srikant, "Throughput-Optimal Opportunistic Scheduling in the Presence of Flow-Level Dynamics," in *Proc. IEEE INFOCOM*, 2010.

[7] ——, "Scheduling in multichannel wireless networks with flow-level dynamics," *ACM SIGMETRICS Performance Evaluation Review*, vol. 38, no. 1, pp. 191–202, 2010.

[8] A. Mekkittikul and N. McKeown, "A starvation-free algorithm for achieving 100% throughput in an input-queued switch," in *Proc. of the IEEE International Conference on Communication Networks*, 1996.

[9] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, "Providing quality of service over a shared wireless link," *IEEE Communications magazine*, vol. 39, no. 2, pp. 150–154, 2001.

[10] S. Shakkottai and A. Stolyar, "Scheduling for multiple flows sharing a time-varying channel: The exponential rule," *Translations of the American Mathematical Society-Series 2*, vol. 207, pp. 185–202, 2002.

[11] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, R. Vijayakumar, and P. Whiting, "Scheduling in a queuing system with asynchronously varying service rates," *Probability in the Engineering and Informational Sciences*, vol. 18, no. 02, pp. 191–217, 2004.

[12] B. Sadiq and G. de Veciana, "Throughput optimality of delay-driven MaxWeight scheduler for a wireless system with flow dynamics," in *Forty-seventh Annual Allerton Conference on Communication, Control, and Computing, Allerton House, Monticello, IL*, 2009.

[13] M. Neely, "Delay-based network utility maximization," in *Proc. IEEE INFOCOM*, 2010.

[14] A. Eryilmaz, R. Srikant, and J. Perkins, "Stable scheduling policies for fading wireless channels," *IEEE/ACM Transactions on Networking*, vol. 13, no. 2, p. 424, 2005.

[15] J. Dai, "On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models," *The Annals of Applied Probability*, pp. 49–77, 1995.

[16] A. Rybko and A. Stolyar, "Ergodicity of stochastic processes describing the operation of open queueing networks," *Problems of Information Transmission*, vol. 28, no. 3, pp. 3–26, 1992.

[17] V. Malyshev and M. Menshikov, "Ergodicity, continuity and analyticity of countable Markov chains," *Transactions of the Moscow Mathematical Society*, vol. 39, pp. 3–48, 1979.

[18] B. Ji, C. Joo, and N. Shroff, "Delay-Based Back-Pressure Scheduling in Multi-Hop Wireless Networks," November 2010. [Online]. Available: http://arxiv.org/abs/1011.5674

[19] C. Joo, X. Lin, and N. Shroff, "Understanding the capacity region of the greedy maximal scheduling algorithm in multi-hop wireless networks," in *IEEE INFOCOM*, 2008, pp. 1103–1111.

[20] ——, "Greedy Maximal Matching: Performance Limits for Arbitrary Network Graphs Under the Node-exclusive Interference Model," *IEEE Transactions on Automatic Control*, 2009.

[21] X. Wu, R. Srikant, and J. Perkins, "Scheduling Efficiency of Distributed Greedy Scheduling Algorithms in Wireless Networks," *IEEE Transactions on Mobile Computing*, pp. 595–605, 2007.

[22] M. Leconte, J. Ni, and R. Srikant, "Improved bounds on the throughput efficiency of greedy maximal scheduling in wireless networks," in *MobiHoc '09*. New York, NY, USA: ACM, 2009, pp. 165–174.

[23] C. Joo and N. Shroff, "Performance of random access scheduling schemes in multi-hop wireless networks," in *IEEE INFOCOM*, 2007, pp. 19–27.

---

[6] Suppose there are N packets sorted by their delays from the largest to the smallest, the top $X\%$ largest delay is defined as the delay of the $\lfloor \frac{NX}{100} \rfloor$-th packet. If $\frac{NX}{100} \le 1$, it means the maximum delay. For example, if the delays are $[3, 2, 1, 1, 1]$, the top 20% largest delay is 2.