

Joint Asynchronous Congestion Control and Distributed Scheduling for Multi-Hop Wireless Networks

Loc Bui, Atilla Eryilmaz, R. Srikant, and Xinzhou Wu

Abstract—We consider a multi-hop wireless network shared by many users. For an interference model that only constrains a node to either transmit or receive at a time, but not both, we propose an architecture for fair resource allocation that consists of a distributed scheduling algorithm operating in conjunction with an asynchronous congestion control algorithm. We show that the proposed joint congestion control and scheduling algorithm supports at least one-third of the throughput supportable by any other algorithm, including centralized algorithms.

Index Terms—Congestion Control, Fair Resource Allocation, Totally Asynchronous Algorithm, Distributed Scheduling, Wireless Networks.

I. INTRODUCTION

THE operation of a wireless network differs from its wireline counterpart in many aspects. Interference, time-varying channels, and limited resources are a few of the distinguishing characteristics of wireless networks. A number of papers have addressed the problem of resource allocation in wireless networks while taking into account many of the features. It was shown in [1] that scheduling algorithms that appropriately use the queue length information can stabilize the queues in the network, provided that the set of arrival rates from the various users of the network lies within what is referred to as the stability region of the network. Such a scheduling rule is called *throughput optimal* in the sense that, for any set of flow rates for which the queues can be stabilized, the throughput-optimal scheduler will stabilize them. Later, there has been a large body of literature that extended these ideas to more general systems and classes of schedulers [2]–[7]. All of these works assumed that the incoming flow rates are *inelastic*, and that the buffer length information is available at a central coordinator instantaneously, which then determines the allocation of the resources and informs all the nodes, again in an instantaneous fashion.

Motivated by the works for wireline networks [8]–[10], recently there has been much interest in incorporating congestion control into the system in addition to the queue-

length based schedulers for the purpose of fair-resource allocation [11]–[17]. These algorithms determine the rate at which each user is allowed to inject data into the network as a function of the current congestion level of the network. The congestion level information is fed back to the controller from the nodes. It has been shown that fair allocation¹ can be achieved through the joint operation of these two mechanisms – scheduling and congestion control. However, each of these papers has one or more of the following assumptions: (i) scheduling mechanism is ignored [16], [17], or (ii) a centralized scheduling algorithm is assumed [11]–[13], [15], or (iii) it is assumed that the congestion price information can be instantaneously exchanged between all of the nodes (an assumption made by all the previous papers). The requirement of decentralized scheduling is obvious in a multi-hop wireless network. However, the importance of the ability of the network to exchange information between the nodes does not seem to have been addressed previously. In particular, two nodes can exchange congestion price information only when they can successfully transfer information between them. Unlike a wireless network, information transfer between nodes is subject to interference constraints and hence neither data nor congestion price information can be transferred between two nodes without taking into account the activity of other nodes in their vicinity. This is similar to the asynchronous model considered in [9], except that in [9] it is assumed that nodes exchange information with a bounded delay which may not be the case in wireless networks, as we will demonstrate later.

In this work, one of our goals is to relax some of these assumptions and study the performance of a distributed resource allocation architecture for a specific, but widely used interference model. Under this interference model, called the node-exclusive model, each node can either transmit or receive at a given time, but not both. No other constraints are imposed on the transmission. Thus, each feasible schedule is a matching² of the underlying graph of the network. Such an interference model is appropriate for Bluetooth networks [18] or in FH-CDMA networks [19], [20].

The main contributions of this paper are as follows:

- We propose an architecture for joint congestion control and scheduling under the node-exclusive model. Congestion control is performed by exchanging congestion

This research was supported by AFOSR URI F49620-01-1-0365, NSF Grant ECS-0401125, and a VEF Fellowship.

An earlier version of this paper appeared in the Proceedings of IEEE INFOCOM, Spain, April 2006.

L. Bui and R. Srikant are with the Department of Electrical and Computer Engineering and Coordinated Science Lab, University of Illinois at Urbana-Champaign, USA (emails: {locbui, rsrikant}@uiuc.edu).

A. Eryilmaz is with Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, USA (email: eryilmaz@mit.edu).

X. Wu is with Flarion Technologies, NJ, USA (email: x.wu@flarion.com).

¹An allocation is said to be *fair* if the sum of the utilities of the users is maximized over all possible allocations.

²A matching is a set of edges that no pair is incident to the same node.

price and users' packet arrival rate information in an asynchronous fashion. A distributed maximal matching algorithm proposed in [21] is used to perform the scheduling.

- Unlike [21], we do not assume that the user arrival rates are known at each node on its route instantaneously. Instead, we use a modification of an algorithm in [22], [23] to stabilize the network.
- We consider a deterministic fluid model and show the stability of the joint congestion control and scheduling architecture.
- As mentioned earlier, congestion information can be exchanged between two nodes only when they are scheduled to transmit to each other. To ensure that every pair of neighbors communicate with each other infinitely often (almost surely), we introduce a slight modification to the maximal matching algorithm whereby every node attempts to make a connection with each one of its neighbors with a small probability.
- With the above modification to the maximal matching scheduling algorithm, the resulting congestion control becomes an asynchronous algorithm with possibly unbounded delays in exchanging congestion information. To the best of our knowledge, the approach in [24] does not seem to apply to our asynchronous computation model, and therefore, we provide a new proof of the convergence of the asynchronous congestion control algorithm.

It should be noted that we consider a time-slotted model, which assumes that all nodes have a common notion of the beginning and the end of a time slot. We refer the reader to [25] for a discussion about a time-slotted model, which uses guard intervals in the context of FH-CDMA networks to enforce this requirement at the cost of slightly reduced throughput.

The paper is organized as follows. Section II introduces the system model we consider in this paper, including the interference model, the description of the regulated maximal matching scheduler, and the problem statement. We describe and analyze a heuristic fluid model analysis of the synchronous version of the system in Section III. In Section IV, we provide the description of the asynchronous congestion controller and prove its convergence properties. We finish with conclusions in Section VI.

II. SYSTEM MODEL

Consider a graph, $\mathcal{G} = (\mathcal{N}, \mathcal{L})$, representing a wireless network where \mathcal{N} is the set of nodes and \mathcal{L} is the set of directed links. We will use both notations (n, m) and l to refer to a link in \mathcal{L} . If a link (n, m) is in \mathcal{L} , then it is possible to send packets from node n to node m subject to the interference constraints to be described shortly. We assume that time is slotted, and let $c_l \geq 1$ be the fixed capacity (the number of packets per time-slot that can be transferred over the link) of each link l . Let \mathcal{F} denote the set of flows that share the network resources. The main goal of this paper is to derive a fully distributed asynchronous algorithm that achieves fair allocation of system resources among the competing flows.

We assume that each flow, f , has a unique, loop-free route and a utility function associated with it. We use H_l^f to denote

the indicator function that is equal to one when link l is in the route of flow f , and zero otherwise. The utility function, denoted by $U_f(\cdot)$, is assumed to satisfy the following set of conditions:

- $U_f(\cdot)$ is a strictly concave, nondecreasing, twice differentiable function.
- The second derivative of the utility function is bounded, i.e., for every $\bar{M} \in (0, \infty)$, there exists a constant $m < \infty$ such that

$$0 \leq -\frac{1}{U_f''(x)} \leq m \quad \forall x \in [0, \bar{M}]. \quad (1)$$

- For every $\bar{M} \in (0, \infty)$, there exists a constant a such that

$$\left| U_f'' \left(U_f'^{-1}(y) \right) \right| \geq ay \quad \forall y \geq \bar{M}. \quad (2)$$

- $U_f'^{-1}(\cdot)$ is a convex function, and satisfies³

$$1 - \frac{U_f'^{-1}(\kappa + \frac{\beta}{L^{1-\sigma}})}{U_f'^{-1}(\kappa)} = O(L^{-\gamma}) \quad (3)$$

for any fixed $\kappa, \beta > 0$ and for some $\gamma \in (0, 1)$ that is determined as a function of $\sigma \in (0, 1)$.

We note that these conditions are not restrictive and hold for the following class of utility functions.

$$U_f(x) = \frac{x^{1-\alpha_f}}{1-\alpha_f} \quad \forall \alpha_f > 0 \quad (4)$$

if x is upper bounded. This class of utility functions are known to characterize a large class of fairness concepts [26].

A. Interference Model

In this subsection, we describe the interference model, referred to as the *node-exclusive interference model*, assumed in this paper and its implications. According to this model, each node can either transmit or receive, but not both, at a given time. There are no other constraints on the transmission.

The arrival processes are assumed to be discrete as follows: at each time slot t , the number of arrivals for flow f is a Poisson random variables with mean x_f . At this point, let us define the *capacity region*, Λ , of the network simply as the set of flow rates $\mathbf{x} = \{x_f\}_f$ that are supportable by the network. We will provide a more precise description of Λ at the end of this section, after we describe the interference model. Let $\mathcal{E}(n)$ be the set of links that are incident on node n , i.e. $\mathcal{E}(n) = \{(h, k) \in \mathcal{L} : h = n \text{ or } k = n\}$. Then, the following facts can be asserted.

Fact 1:

$$\Lambda \subset \left\{ \mathbf{x} : \sum_{l \in \mathcal{E}(n)} \frac{\sum_{f \in \mathcal{F}} H_l^f x_f}{c_l} \leq 1, \forall n \in \mathcal{N} \right\}.$$

This fact simply states that no node can be active more than 100% of the time and is proved in [27].

³ $f(x) = O(g(x))$ implies that $\limsup_{x \rightarrow \infty} \frac{f(x)}{g(x)} < \infty$.

Fact 2: Any set of flow rates \mathbf{x} that satisfies

$$\sum_{l \in \mathcal{E}(n)} \frac{\sum_{f \in \mathcal{F}} H_l^f x_f}{c_l} \leq \frac{2}{3} \quad \forall n \in \mathcal{N} \quad (5)$$

lies within the capacity region of the network.

This fact is discussed in [20], [27], and is based on a work by Shannon [28]. Notice that the condition in (5) is equivalent to the requirement that any node in the network should not be scheduled more than two-thirds of the time. Also note that under the node-exclusive interference model, any feasible schedule corresponds to a *matching* of the graph \mathcal{G} . Let $\mathcal{M} = \{M_1, M_2, \dots, M_K\}$ be the set of all possible matchings, where M_i is a 0-1 vector of $|\mathcal{L}|$ dimensions that denotes the set of links that are active for the i^{th} matching. Here, \mathcal{M} is a finite set since we have a graph with finite number of links. Let $\text{co}(\mathcal{M})$ denote the convex-hull⁴ of the set of matchings. Then, the capacity region can be expressed as

$$\Lambda = \left\{ \mathbf{x} : \left[\frac{\sum_{f \in \mathcal{F}} H_l^f x_f}{c_l} \right]_{l \in \mathcal{L}} \in \text{co}(\mathcal{M}) \right\}.$$

Also, for each link $(n, m) \in \mathcal{L}$, let Υ_{nm} denote the set of all links interfering with (n, m) plus link (n, m) itself. For the node-exclusive interference model, Υ_{nm} is the set of (n, m) and all one-hop neighbor links of (n, m) :

$$\Upsilon_{nm} = \{(h, k) \in \mathcal{L} : h \neq k, h \in \{m, n\} \text{ or } k \in \{m, n\}\}.$$

B. Distributed Scheduler

In this subsection, we introduce the fully distributed scheduler that is implemented at the MAC layer. The scheduler determines which links to activate and which packets to serve at a given slot. It is assumed that the amount of time required to perform the scheduling task compared to the actual packet transmission is small.

We introduce the following notation to describe a scheduling rule $\pi = \{\pi_{nm}, (n, m) \in \mathcal{L}\}$: we let $\pi_{nm}(t)$ be equal to one if link $(n, m) \in \mathcal{L}$ is scheduled at slot t and zero otherwise. Also, we define $P_{nm}(t)$ to be the number of packets served over link (n, m) at slot t , and $P_{nm}^f(t)$ to be the number of flow f 's packets served over link (n, m) at slot t . Clearly,

$$P_{nm}(t) = \sum_{f \in \mathcal{F}} H_{(n,m)}^f P_{nm}^f(t).$$

It is assumed that for each node n , a queue is maintained for each of the outgoing links $(n, m) \in \mathcal{L}$. This queue stores and serves all the packets that are forwarded to node m in a first-come, first-served (FCFS) fashion at those slots when the scheduler sets $\pi_{nm}(t) = 1$. We let $Q_{nm}(t)$ denote the total number of packets at the beginning of slot t that are waiting to be transmitted over link (n, m) .

The scheduler we consider in this paper is a slightly modified version of the *Regulated Maximal Matching Scheduler* in [23]. The key idea is to introduce regulators associated with each link, one for each flow passing through the link.

⁴A convex hull of a set S is the intersection of all convex sets that cover each element of S .

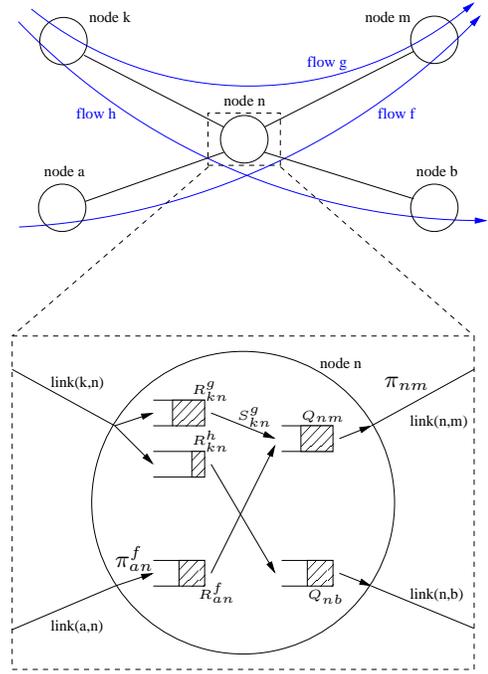


Fig. 1. An example network

Flow f packets that are served over link (n, m) are buffered at the corresponding regulator of link (n, m) before they are transferred to the queue that will relay them to the next link. We let $R_{nm}^f(t)$ denote the length of link (n, m) 's regulator buffer corresponding to flow f at the beginning of time slot t . A $\lambda(t)$ -regulator is a logical device which allows packets to pass through it at a maximum rate of $\lambda(t)$ at slot t . Specifically, at slot t , a $\lambda(t)$ -regulator associated with link l checks its buffer size, and if its buffer size exceeds link capacity c_l , it transfers c_l packets from its buffer probability $\frac{\lambda(t)}{c_l}$; otherwise, no transfer occurs (other implementations of the regulator are also possible). Finally, we let $S_{nm}^f(t)$ denote the number of packets that leave regulator R_{nm}^f at slot t . We refer the reader to Figure 1 for an example network.

With these definitions, the evolution of the number of packets in the regulators and the queues can be described with the following difference equations.

$$\begin{aligned} R_{nm}^f(t+1) &= R_{nm}^f(t) - S_{nm}^f(t) + P_{nm}^f(t) \\ Q_{nm}(t+1) &= Q_{nm}(t) - P_{nm}(t) + \sum_{f \in \mathcal{F}} S_n^f(t) H_{nm}^f, \end{aligned}$$

where $S_n^f(t)$ is the output of the regulator for flow f that is maintained at node n .

In our model, the value of $\lambda_f(t)$ used at a regulator is based on the rate at which flow f is generating data. In general, this information is not available instantaneously to all nodes in f 's path, and hence it is both inaccurate (old) and fluctuating (due to the fact that flow f 's rate is determined by a dynamic congestion control algorithm to be described later). In comparison, the scheduler considered in [23] assumes the knowledge of the mean flow rates at all the regulators and uses this fixed value in its implementation. In our model, however, the current flow rate information is passed from the sources

to each of the nodes with a random propagation delay. To account for the delay, we let $\tau_f^n(t) \in [0, t]$ be the time slot at which the rate information of source f (i.e. $x_f(\cdot)$) was sent, given that it is received by node n at slot t . Further, suppose that node n is the k^{th} node on the route of flow f . Then, the regulator for flow f at node n is a $(x_f(\tau_f^n(t)) + (k-1)\epsilon_r)$ -regulator, where $\epsilon_r > 0$ can take arbitrarily small values. In other words, the regulator uses the most recent update of the rate (plus a very small amount) as $\lambda_f(t)$.

The scheduler we consider uses the values of $Q_{nm}(t)$ to find either a *Maximal Matching*⁵(MM) with a high probability or some matching with a small probability. In particular, a maximal matching is selected in a distributed fashion among those links that have at least one backlogged packet in their buffers. The formal description is provided next.

DISTRIBUTED SCHEDULER - DISCRETE-TIME

At the beginning of each time slot, each node, say n , determines the *eligible* set of links according to:

- With probability $\epsilon_s > 0$ allow all links $(n, m) \in \mathcal{L}$ to be eligible.
- Otherwise, allow only those $(n, m) \in \mathcal{L}$, with $Q_{nm}(t) \geq c_{(n,m)}$ to be eligible.

Then, which set of links to activate at node n is determined by the following distributed strategy:

- If n has at least one eligible neighbor (a neighbor that is at the other end of an eligible link) that is not *matched* with another node, then choose any one randomly, say m , and match them with each other. After this operation n and m are said to be matched.
- Otherwise stop.

At the end of this algorithm, those links that have matched end nodes, say (n, m) , will be scheduled to transmit $c_{(n,m)}$ packets within slot t , i.e. $\frac{P_{nm}(t)}{c_{(n,m)}} = \pi_{nm}(t) = 1$. It is not difficult to see that the algorithm will result in a set of scheduled links that is a maximal matching of \mathcal{G} , because any node can be matched with at most one other node, and no pair of unmatched nodes with a packet to transmit is left at the end. The ϵ_s parameter is included in this algorithm to assure a positive probability of activating a link even if it has no backlogged packet. This is necessary to ensure that every link is active occasionally, to allow its nodes to exchange congestion information.

It can be seen that if a scheduling rule $\pi(t)$ is a maximal matching then it satisfies

$$\sum_{\{(h,k) \in \mathcal{T}_{nm}\}} \pi_{hk}(t) \geq 1 \quad (6)$$

for every $(n, m) \in \mathcal{L}$ with $Q_{nm}(t) \geq c_{(n,m)}$.

Also, notice that the MM algorithm is performed using the lengths of the actual queues, not the lengths of the regulator buffers. The following fact is due to [23].

Fact 3: For the multi-hop wireless network model introduced above, the Regulated MM Scheduling Algorithm can

achieve stability if the set of mean arrival rates of the flows, \mathbf{x} , satisfies

$$\sum_{l \in \mathcal{E}(n)} \frac{\sum_{f \in \mathcal{F}} H_l^f x_f}{c_l} \leq \frac{1}{3} \quad \forall n \in \mathcal{N}$$

C. Problem Statement

Given the above model, our goal is to have the mean flow rate vector $\mathbf{x}^* \geq 0$ satisfy:

$$\begin{aligned} \mathbf{x}^* &\in \arg \max_{\mathbf{x}} \sum_{f \in \mathcal{F}} U_f(x_f) \\ \text{s.t.} \quad &\sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f}{c_l} H_l^f \leq \frac{1}{3} \quad \forall n \in \mathcal{N}, \end{aligned} \quad (7)$$

provided that the regulators and the queues are kept stable. The strict concavity assumption of the utility functions implies that \mathbf{x}^* is unique. We note that the constraint set of this optimization problem contains $\Lambda/3$ due to Fact 1. In (7), we have formulated the wireless network rate allocation problem as an optimization problem by taking the interference and distributed scheduling constraint into account. Such an approach has been taken for our interference model in [15], [16] before. However, it is assumed in these works that the current rate of each source and the current price of each node are immediately available at all the nodes. Obviously, in an actual operation, such information can only be conveyed along with the data transmissions, and hence is delayed by random amounts for each source-node pair. Updates with delays have been taken into account in [9] in the wireline and [29] in the wireless setting, but in our work we link the delay to the scheduler and also allow for unbounded delays as long as the updates occur infinitely often.

Let us describe a new optimization problem whose optimum point converges to \mathbf{x}^* . For any $\epsilon > 0$, let

$$\begin{aligned} \mathbf{x}^*(\epsilon) &\in \arg \max_{\mathbf{x}} \sum_{f \in \mathcal{F}} U_f(x_f) \\ \text{s.t.} \quad &\sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f}{c_l} H_l^f \leq \frac{1}{3} - \epsilon \quad \forall n \in \mathcal{N} \end{aligned} \quad (8)$$

It is not difficult to see that $\mathbf{x}^*(\epsilon) \rightarrow \mathbf{x}^*$ as $\epsilon \rightarrow 0$. We introduced this new problem to make sure that the optimum point lies strictly inside the feasible region of (7). This is necessary to ensure that the queues and the regulators are stable as will be discussed in Section III.

In the remainder of this paper, we will propose a congestion controller mechanism that operates on top of the Regulated MM Scheduler that will provide mean rates that are arbitrarily close to \mathbf{x}^* . Furthermore, we incorporate all the asynchronous components that exist in the operation of the system. In particular, we model the random nature of the scheduling operation, which results in potentially unbounded delays in the information communication between different components of the network. Noting that information feedback is critical in the operation of our congestion controller, it is crucial to answer the question as to whether the asynchronism inherent in the network will significantly affect the performance.

⁵A matching is a collection of links such that no two links are incident on the same node. A matching is said to be maximal if it is a matching and no new link can be added to the set without losing the matching property.

III. CONTINUOUS-TIME FLUID MODEL ANALYSIS

In this section, we consider a continuous-time, deterministic fluid model of the system, in which all stochastic processes are approximated by their mean values. In Section III-A, we introduce a congestion control algorithm for this fluid model, and analyze its convergence properties. Then, in Section III-B, we describe a heuristic continuous-time version of the scheduler described in Section II-B and prove its stabilizing characteristics. We will build upon these analyses later to investigate the asynchronous, discrete-time algorithm.

A. Continuous-time Congestion Controller

We start by noting that the optimization problem in (8) can be solved by using Lagrange multipliers. The Lagrangian and the Dual functions of the problem (8) are:

$$\begin{aligned} \mathbb{L}_\epsilon(\mathbf{x}, \mu) &= \sum_{f \in \mathcal{F}} U_f(x_f) - \sum_{n \in \mathcal{N}} \left\{ \mu_n \right. \\ &\quad \times \left. \left(\sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f}{c_l} H_l^f - \left(\frac{1}{3} - \epsilon \right) \right) \right\} \\ \mathbb{D}_\epsilon(\mu) &= \max_{\mathbf{x} \geq 0} \mathbb{L}_\epsilon(\mathbf{x}, \mu) \\ &= \sum_{f \in \mathcal{F}} \max_{x_f \geq 0} \left(U_f(x_f) - x_f \sum_{(n,m) \in R(f)} \frac{\mu_n + \mu_m}{c_{(n,m)}} \right) \\ &\quad + \left(\frac{1}{3} - \epsilon \right) \sum_n \mu_n \end{aligned}$$

where μ_n is the Lagrange multiplier associated with the n^{th} constraint of (8), and $R(f)$ is the set of links which are along the route of flow f .

Then the dual optimization problem to (8) is given by $\min_{\mu \geq 0} \mathbb{D}_\epsilon(\mu)$. It can be shown that for the problem we consider, there is no duality gap [30]. Thus, there exists a nonempty set of Lagrange multipliers, Ξ_ϵ , any element (say $\mu^*(\epsilon)$) of which satisfies $\mathbb{D}_\epsilon(\mu^*(\epsilon)) = \sum_f U_f(x_f^*(\epsilon))$. But for any feasible \mathbf{x} of the primal problem (8), we must have:

$$\sum_{n \in \mathcal{N}} \mu_n^*(\epsilon) \left(\sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f}{c_l} H_l^f - \left(\frac{1}{3} - \epsilon \right) \right) \leq 0$$

because the expression in the parenthesis can never be positive for a feasible rate vector, and $\mu^*(\epsilon)$ is a nonnegative vector. Thus, we must have $\mathbf{x}^*(\epsilon)$ as the optimizer of the Lagrangian $\mathbb{L}_\epsilon(\mathbf{x}, \mu^*(\epsilon))$, and also have the pair $(\mathbf{x}^*(\epsilon), \mu^*(\epsilon))$ satisfy:

$$\sum_{n \in \mathcal{N}} \mu_n^*(\epsilon) \left(\sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f^*(\epsilon)}{c_l} H_l^f - \left(\frac{1}{3} - \epsilon \right) \right) \leq 0, \quad (9)$$

which is also called the complementary slackness condition in the optimization literature. From the Lagrangian, it is easy to see that $(\mathbf{x}^*(\epsilon), \mu^*(\epsilon))$ should also satisfy:

$$x_f^*(\epsilon) = U_f'^{-1} \left(\sum_{(n,m) \in R(f)} \frac{\mu_n^*(\epsilon) + \mu_m^*(\epsilon)}{c_{(n,m)}} \right) \quad (10)$$

Let us define $\mathbf{p}^*(\epsilon)$ and $\mathbf{q}^*(\epsilon)$ as follows: $p_n^*(\epsilon) \triangleq L\mu_n^*(\epsilon)$,

$$q_f^*(\epsilon) \triangleq \sum_{(n,m) \in R(f)} \frac{p_n^*(\epsilon) + p_m^*(\epsilon)}{c_{(n,m)}} \quad (11)$$

where L is some multiplicative factor. We define $\Psi_\epsilon = L\Xi_\epsilon$, i.e., if $\mu^*(\epsilon) \in \Xi_\epsilon$ then the corresponding $\mathbf{p}^*(\epsilon) \in \Psi_\epsilon$. Notice that we can rewrite (10) and the complementary slackness condition (9) in terms of $\mathbf{p}^*(\epsilon)$ and $\mathbf{q}^*(\epsilon)$:

$$\begin{aligned} x_f^*(\epsilon) &= U_f'^{-1} \left(\frac{q_f^*(\epsilon)}{L} \right) \\ 0 &= \left(\sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f^*(\epsilon)}{c_l} H_l^f - \left(\frac{1}{3} - \epsilon \right) \right)_{p_n^*(\epsilon)}^+ \end{aligned} \quad (12)$$

where we define $(y)_z^+$ to be equal to y if $z \geq 0$ and zero if $y \leq 0$ and $z = 0$.

The congestion control algorithm for the fluid model is described as follows.

CONGESTION CONTROLLER - CONTINUOUS-TIME

At time t ,

Source f computes: $x_f(t) = U_f'^{-1} \left(\frac{q_f(t)}{L} \right)$.

Node n computes:

$$\dot{p}_n(t) = \left(\sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f(t)}{c_l} H_l^f - \left(\frac{1}{3} - \epsilon \right) \right)_{p_n(t)}^+$$

where

$$q_f(t) \triangleq \sum_{(n,m) \in R(f)} \frac{p_n(t) + p_m(t)}{c_{(n,m)}} \quad (13)$$

Notice that here we have assumed synchronous computation: information updates at the sources and the nodes occur instantaneously and simultaneously at each time instant. In later sections, when considering the discrete-time model, we will remove this key assumption and develop a fully asynchronous algorithm for congestion control. Nevertheless, the analysis of the continuous-time system will be useful in understanding the more realistic model. Next, we state the theorem that proves the convergence properties of the congestion controller.

Theorem 1: Starting from any initial $\mathbf{p}(0)$, $\mathbf{x}(t)$ eventually reaches $\mathbf{x}^*(\epsilon)$ as $t \rightarrow \infty$.

Proof: Consider the Lyapunov function:

$$V(\mathbf{p}; \mathbf{p}^*(\epsilon)) = \frac{1}{2} \sum_n (p_n - p_n^*(\epsilon))^2,$$

which is defined for some $\mathbf{p}^*(\epsilon) \in \Psi_\epsilon$. For notational convenience, we will occasionally use \mathbf{p}^* and \mathbf{x}^* instead of $\mathbf{p}^*(\epsilon)$ and $\mathbf{x}^*(\epsilon)$. Then, the time derivative of this function at t satisfies

$$\begin{aligned} \dot{V}(\mathbf{p}(t), \mathbf{p}^*) &= \sum_n \{ (p_n(t) - p_n^*) \times \\ &\quad \left(\sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f(t)}{c_l} H_l^f - \left(\frac{1}{3} - \epsilon \right) \right)_{p_n(t)}^+ \} \end{aligned}$$

We first consider the case when $\mathbf{p}(t) \in \Psi_\epsilon$: note that the rate vector associated with $\mathbf{p}(t)$ has to be the unique optimizer of (8), i.e. $\mathbf{x}^*(\epsilon)$. But, by utilizing the complementary slackness condition provided in (12) we can easily conclude that

$$\dot{V}(\mathbf{p}(t), \mathbf{p}^*) = 0 \quad \text{for all } \mathbf{p}(t) \in \Psi_\epsilon. \quad (14)$$

Next, we consider $\dot{V}(\mathbf{p}(t), \mathbf{p}^*)$ for any $\mathbf{p}(t) \geq 0$: define

$$y_n(t) \triangleq \sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f(t)}{c_l} H_l^f \quad (15)$$

Then we have:

$$\begin{aligned} \dot{V}(\mathbf{p}(t), \mathbf{p}^*) &= \sum_n (p_n(t) - p_n^*) \left(y_n - \left(\frac{1}{3} - \epsilon \right) \right)_{p_n(t)}^+ \\ &\leq \sum_n (p_n(t) - p_n^*) \left(y_n(t) - \left(\frac{1}{3} - \epsilon \right) \right) \\ &= \sum_n (p_n(t) - p_n^*) (y_n(t) - y_n^*) \\ &\quad + \sum_n (p_n(t) - p_n^*) \left(y_n^* - \left(\frac{1}{3} - \epsilon \right) \right) \end{aligned}$$

where the inequality follows from the fact that if $p_n(t) = 0$ and $y_n(t) - (\frac{1}{3} - \epsilon) \leq 0$, then $(y_n(t) - (\frac{1}{3} - \epsilon))_{p_n(t)}^+ = 0$.

Otherwise, $(y_n(t) - (\frac{1}{3} - \epsilon))_{p_n(t)}^+ = (y_n(t) - (\frac{1}{3} - \epsilon))$.

Also, from the complementary slackness condition, we have that if $p_n^* > 0$, then $y_n^* = (\frac{1}{3} - \epsilon)$. Otherwise, if $p_n^* = 0$, then $y_n^* \leq (\frac{1}{3} - \epsilon)$. This fact implies

$$\sum_n (p_n(t) - p_n^*(\epsilon)) \left(y_n^*(\epsilon) - \left(\frac{1}{3} - \epsilon \right) \right) \leq 0 \quad (16)$$

Therefore,

$$\begin{aligned} \dot{V}(\mathbf{p}(t), \mathbf{p}^*) &\leq \sum_n (p_n(t) - p_n^*) (y_n(t) - y_n^*) \\ &= \sum_n (p_n(t) - p_n^*) \left(\sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f(t) - x_f^*}{c_l} H_l^f \right) \\ &= \sum_f (x_f(t) - x_f^*) \left(\sum_{(n,m) \in R(f)} \frac{p_n(t) + p_m(t)}{c_{(n,m)}} - \sum_{(n,m) \in R(f)} \frac{p_n^* + p_m^*}{c_{(n,m)}} \right) \\ &= \sum_f (x_f(t) - x_f^*) (q_f(t) - q_f^*) \\ &= L \sum_f (x_f(t) - x_f^*) (U_f'(x_f(t)) - U_f'(x_f^*)) \\ &\leq 0, \end{aligned}$$

with strict inequality when $\mathbf{p}(t) \notin \Psi_\epsilon$, due to the strict concavity assumption of $U_f(\cdot)$. Then, by combining this result with (14) and by invoking LaSalle's theorem [31, Theorem 4.4], we conclude that $\mathbf{p}(t) \xrightarrow{t \rightarrow \infty} \Psi_\epsilon$ and hence $\mathbf{x}(t) \xrightarrow{t \rightarrow \infty} \mathbf{x}^*(\epsilon)$. ■

B. Continuous-time Scheduler

In this subsection we show that the Regulated MM Scheduling Algorithm, along with the Congestion Control Algorithm described in previous subsection, can achieve stability.

Recall that Λ denotes the capacity region of the network. Define

$$\Omega = \left\{ \mathbf{x} : \sum_{l \in \mathcal{E}(n)} \frac{\sum_{f \in \mathcal{F}} H_l^f x_f}{c_l} \leq \frac{1}{3}, \forall n \in \mathcal{N} \right\}.$$

By Fact 2, we have $\Omega \subset \frac{\Lambda}{2}$. Moreover, we know that the optimum point $\mathbf{x}^*(\epsilon)$ is strictly inside Ω . Thus, $\mathbf{x}^*(\epsilon)$ is strictly inside the half of the capacity region $\frac{\Lambda}{2}$.

The evolution of the scheduler's queues in continuous-time can be described as follows.

SCHEDULER - CONTINUOUS-TIME

$$\dot{R}_{nm}^f(t) = (P_{nm}^f(t) - (x_f(t) + (K_f^n + 1)\epsilon_r))_{R_{nm}^f(t)}^+ \quad (17)$$

$$\dot{Q}_{nm}(t) = \left(\sum_f (x_f(t) + K_f^n \epsilon_r) \mathcal{I}_{R_{nm}^f(t) > 0} H_{nm}^f - P_{nm}(t) \right)_{Q_{nm}(t)}^+ \quad (18)$$

where K_f^n is the hop number of node n along the path of flow f . Note that this number is upper-bounded by N_{max} where N_{max} is the maximum number of nodes along any flow's path. Notice that in the fluid limit, the scheduling rule $\pi(t)$ satisfies the condition (6) for every $(n, m) \in \mathcal{L}$ with $Q_{nm}(t) > 0$.

Theorem 2: Starting from any initial $\mathbf{R}(0)$ and $\mathbf{Q}(0)$, we have $(\mathbf{R}(t) + \mathbf{Q}(t)) \rightarrow 0$ as $t \rightarrow \infty$.

Proof: The proof uses the fact that $\mathbf{x}^*(\epsilon)$ lies strictly inside half of the capacity region $\frac{\Lambda}{2}$, and a Lyapunov argument that is based on the discrete-time analysis of [23]. The details are moved to the Appendix. ■

IV. DISCRETE-TIME, ASYNCHRONOUS MODEL ANALYSIS

In this section, we return to the discrete-time system model. Here we do not assume synchronous computation which is unrealistic in an actual network. Next, we describe the discrete-time congestion controller mechanism. Then, we will analyze this system.

A. Description of the Congestion Controller

We assume that when two nodes are matched by the MAC layer scheduler, the relevant price and flow rate information is shared between them. Therefore, the information updates at the sources and the nodes are asynchronous. Then, let us consider the following asynchronous congestion controller.

CONGESTION CONTROLLER - DISCRETE-TIME

At the beginning of each time slot, t ,
Source f computes: $x_f(t) = \min \left\{ M, U_f'^{-1} \left(\frac{q_f(\tau^f(t))}{L} \right) \right\}$

Node n computes:

$$p_n(t+1) = \left(p_n(t) + \sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f(\tau_f^{(n)}(t))}{c_l} H_l^f - \left(\frac{1}{3} - \epsilon \right) \right)^+,$$

where

$$q_f(\tau_f^f(t)) \triangleq \sum_{(n,m) \in R(f)} \frac{p_n(\tau_n^f(t)) + p_m(\tau_m^f(t))}{c_{(n,m)}} \quad (19)$$

and $M > 2 \max_{l \in \mathcal{L}} \{c_l\}$. Here we define $\tau_n^f(t) \in [0, t]$ to be the time slot at which the price information of node n ($p_n(\cdot)$) was sent, given that it is received at source f at slot t . Thus, $t - \tau_n^f(t)$ is the random variable representing the number of slots it took for $p_n(\cdot)$ to be transferred to source f by the underlying scheduling mechanism, given that it is received at time slot t . And similarly, $\tau_f^{(n)}(t) \in [0, t]$ is the time slot at which the rate information of source f ($x_f(\cdot)$) was sent, given that it is received by node n at slot t . Further, let $\tau^f(t)$ be the vector of $\tau_n^f(t)$ for source f at time t , and also, $\tau^{(n)}(t)$ be the vector of $\tau_f^{(n)}(t)$ for node n at time t . Finally, L is a positive constant. We are interested in the behavior of the system when L is large.

Also, let us introduce the notation:

$$y_n(\tau^{(n)}(t)) \triangleq \sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f(\tau_f^{(n)}(t))}{c_l} H_l^f \quad (20)$$

Notice that $q_f(\tau^f(t))$ (defined in (19)) is the estimated price of flow f 's path at time t which is computed using delayed versions of the actual prices. On the other hand, $q_f(t)$ (defined in (13)) assumes the instantaneous knowledge of all the prices on flow f 's path. Similar interpretation holds for $y_n(\tau^{(n)}(t))$ and $y_n(t)$ (defined in (20) and (15)).

This model contains the essential components of the asynchronous operation of the network. It is referred to as ‘‘Totally Asynchronous’’ in [24]. Specifically, the amount of time it takes for the flow rate and node price information to reach any node or a source is captured by an unbounded random variable. In the next section, we will prove that it asymptotically solves the resource allocation problem described in (8).

B. Analysis of the Congestion Controller

Let $\Delta(t)$ be the vector $\{|t - \tau_n^f(t)|, |t - \tau_f^{(n)}(t)|\}_{(n,f)}$, and define $\tau_m = \min_{n,f} \{\tau_n^f(t), \tau_f^{(n)}(t)\}$. Also, let $\mathcal{P}(t)$ be the sequence of vectors $(\mathbf{p}(t), \mathbf{p}(t-1), \dots, \mathbf{p}(\tau_m))$. Then $(\mathcal{P}(t), \Delta(t))$ forms the state of a Markov chain.

Consider the Lyapunov function, $V(\cdot)$, used in the continuous-time analysis (Theorem 1). For notational convenience, we will generally omit the ϵ term in our analysis. However, we will make the main statements with ϵ . The following theorem characterizes the drift of this Lyapunov function.

Theorem 3: The mean drift satisfies:

$$\begin{aligned} \mathbf{E}[\Delta V_t] &\triangleq \mathbf{E}[V(\mathbf{p}(t+1)) - V(\mathbf{p}(t)) \mid \mathcal{P}(t), \Delta(t)] \\ &\leq -\frac{\delta}{L^\gamma} \|\mathbf{q}(t) - \mathbf{q}^*(\epsilon)\| + \hat{C} \|\Delta(t)\|^2 + \hat{B} \end{aligned} \quad (21)$$

for some constants $\hat{C}, \hat{B} < \infty$, $\gamma \in (0, 1)$ and $\delta > 0$. Here, $\|\cdot\|$ denotes Euclidean distance.

Furthermore, there exist some $\tilde{C}, \tilde{B} < \infty$, and $c > 0$ such that

$$\mathbf{E}[\Delta V_t] \leq -cL \|\hat{\mathbf{x}}(t) - \mathbf{x}^*(\epsilon)\|^2 + \tilde{C} \|\Delta(t)\|^2 + \tilde{B} \quad (22)$$

where we define $\hat{x}_f(t) = \min \left\{ M, U_f'^{-1} \left(\frac{q_f(t)}{L} \right) \right\}$, i.e., it is the rate of flow f at time t if all the price information were instantaneously available at the sources.

Before we provide the proof of Theorem 3 we give several lemmas that will be used in the proof. First, we observe that $t - \tau_n^f(t)$ is the amount of time it takes for node n 's price information to reach source f . Similarly, $t - \tau_f^{(n)}(t)$ is the amount of time it takes for the source f 's rate information to reach node n . These are random variables that depend on the matchings chosen in the operation of the scheduler. To ensure that each matching is chosen at least once with a positive probability, recall that we slightly modified the MM Scheduler used in [21], [23] in Section II. Specifically, when the maximal matchings are being determined, with a very small probability, we allow each node to pick a link that does not have any packets to transmit in line. We now establish the following result for the delays in exchanging price and rate information.

Lemma 1: There exists a constant D such that

$$\mathbf{P} \left(|t - \tau_n^f(t)| > TD \right) \leq e^{-\nu(D)T} \quad \forall n, f$$

$$\mathbf{P} \left(|t - \tau_f^{(n)}(t)| > TD \right) \leq e^{-\eta(D)T} \quad \forall n, f$$

where ν and η are some positive numbers depending on D .

Proof: Recall that \mathcal{M} is a finite set containing all the possible matchings. Also, let S_n^f denote the sequence of matchings needed to transmit the price information from node n to source f , and similarly, $S_f^{(n)}$ denote the sequence of matchings needed to transmit the rate information from source f to node n .

Since each node tries to connect to all of its neighbors with some positive probability, ϵ_s , we have: $\mathbf{P}(M_i \text{ occurs at time } t) \geq \delta$, for all i , and for some $\delta > 0$. Hence, for each (n, f) pair, we can state that

$$\mathbf{P} \left(S_n^f \text{ occurs in an interval of length } |S_n^f| \geq \delta^{|S_n^f|} > 0 \right)$$

A similar argument holds for any $S_f^{(n)}$ as well. Let us define $D = \sum_{n,f} |S_n^f| + |S_f^{(n)}|$, and let X be a random variable that equals 1 when all the matchings in S_n^f and $S_f^{(n)}$ occur within the first D slots. Otherwise, X is equal to 0. Then, due to the above discussion, we can find some $\hat{\delta} \in (0, 1)$ such that $\mathbf{P}(X = 0) \leq (1 - \hat{\delta})$, which denotes the probability that at least one of the nodes (or sources) has not received any rate (or price) update from one of the sources (or nodes) within the last D slots. Thus, we have

$$\mathbf{P} \left(|t - \tau_n^f(t)| > D \right) \leq (1 - \hat{\delta}) \quad \forall n, f$$

$$\mathbf{P}\left(\left|t - \tau_f^{(n)}(t)\right| > D\right) \leq (1 - \hat{\delta}) \quad \forall n, f$$

Next, suppose we look over TD slots. Let X_i , $i = 1, \dots, T$ be the associated Bernoulli random variable for the i^{th} block of duration D . Then

$$\begin{aligned} \mathbf{P}\left(\left|t - \tau_n^f(t)\right| > TD\right) &= \prod_{i=1}^T \mathbf{P}(X_i = 0) \\ &\leq (1 - \hat{\delta})^T = e^{-\nu(D)T} \quad \forall n, f. \end{aligned}$$

A similar argument applies to $\left|t - \tau_f^{(n)}(t)\right|$, which completes the proof with appropriately defined $\nu(D)$ and $\eta(D)$ parameters. ■

The following two lemmas yield two different upper bounds on similar expressions, and will be useful for the proof of Theorem 3.

Lemma 2: Given any $B < \infty$, we can find some $\gamma \in (0, 1)$ such that for L large enough, we have

$$\begin{aligned} B + \sum_f \{(q_f(t) - q_f^*) (\hat{x}_f(t) - x_f^*)\} \\ \leq -\frac{\delta}{L^\gamma} \|\mathbf{q}(t) - \mathbf{q}^*\| \mathcal{I}_{\|\mathbf{q}(t) - \mathbf{q}^*\| \geq \omega L^\gamma} + \zeta \mathcal{I}_{\|\mathbf{q}(t) - \mathbf{q}^*\| \leq \omega L^\gamma} \end{aligned}$$

where δ , ζ and ω are positive constants which are independent of L .

Proof: This statement is proved in [12] for a large class of utility functions and for the case of a single transmitter transmitting to many receivers. Here, we consider the multi-hop scenario and further generalize the utility functions. Nevertheless, the arguments are very similar to those in [12] and are moved to the Appendix. ■

Lemma 3: We have

$$\sum_f \{(q_f(t) - q_f^*) (\hat{x}_f(t) - x_f^*)\} \leq -cL \|\hat{\mathbf{x}}(t) - \mathbf{x}^*\|^2,$$

where c is a positive constant which is independent of L .

Proof: We start by adding and subtracting $LU_f'(\hat{x}_f(t))$ into the first factor within the summation, which yields

$$\begin{aligned} &\sum_f \{(q_f(t) - q_f^*) (\hat{x}_f(t) - x_f^*)\} \\ &= \sum_f (q_f(t) - LU_f'(\hat{x}_f(t))) (\hat{x}_f(t) - x_f^*) \quad (23) \\ &\quad + \sum_f (LU_f'(\hat{x}_f(t)) - LU_f'(x_f^*)) (\hat{x}_f(t) - x_f^*) \quad (24) \end{aligned}$$

We will analyze the terms (23) and (24) separately. We claim that (23) ≤ 0 . To see this, we first note that, if $\hat{x}_f(t) < M$, then $q_f(t) = LU_f'(\hat{x}_f(t))$ and hence we have

$$(q_f(t) - LU_f'(\hat{x}_f(t))) (\hat{x}_f(t) - x_f^*) = 0.$$

If, on the other hand, we have $\hat{x}_f(t) = M > x_f^*$, then $q_f(t) < LU_f'(\hat{x}_f(t))$ which implies that

$$(q_f(t) - LU_f'(\hat{x}_f(t))) (\hat{x}_f(t) - x_f^*) \leq 0.$$

Combining these two observations proves our claim.

Next, we turn our attention to (24). We start by noting that

$$\begin{aligned} &(LU_f'(\hat{x}_f(t)) - LU_f'(x_f^*)) (\hat{x}_f(t) - x_f^*) \\ &= -L |U_f'(\hat{x}_f(t)) - U_f'(x_f^*)| |\hat{x}_f(t) - x_f^*|, \quad (25) \end{aligned}$$

which follows from the strict concavity assumption on $U_f(\cdot)$. Also, due to Taylor expansion, we can find some $y_f(t)$ between $\hat{x}_f(t)$ and x_f^* for which,

$$U_f'(\hat{x}_f(t)) - U_f'(x_f^*) = (\hat{x}_f(t) - x_f^*) U_f''(y_f(t)).$$

Using the assumption in (1), we can thus claim that there exists some $c > 0$ which yields

$$|U_f'(\hat{x}_f(t)) - U_f'(x_f^*)| \geq c |\hat{x}_f(t) - x_f^*|.$$

Substituting this into (25) and then (25) into (24) yields the result. ■

Proof (Theorem 3): Notice that we can write

$$p_n(t+1) = p_n(t) + y_n(\tau^n(t)) - \left(\frac{1}{3} - \epsilon\right) + u_n(t),$$

where $u_n(t)$ is a nonnegative parameter that assures the non-negativity of $p_n(t+1)$. We first start by showing that we can ignore the $u_n(t)$ term in the iteration. Towards this end, we can write

$$(p_n(t+1) - p_n^*)^2 =$$

$$\left(p_n(t) + y_n(\tau^n(t)) - \left(\frac{1}{3} - \epsilon\right) - p_n^*\right)^2 \quad (26)$$

$$+ 2 \left(p_n(t) + y_n(\tau^n(t)) - \left(\frac{1}{3} - \epsilon\right)\right) u_n(t) \quad (27)$$

$$+ u_n^2(t) \quad (28)$$

$$- 2u_n(t)p_n^*, \quad (29)$$

for any n . Since $p_n^*, u_n(t) \geq 0$, we have (29) ≤ 0 . We also claim that (27)+(28) ≤ 0 . To see this, we observe that: $u_n(t) = 0$ if $p_n(t) + y_n(\tau^n(t)) - \left(\frac{1}{3} - \epsilon\right) > 0$, and that $u_n(t) = -(p_n(t) + y_n(\tau^n(t)) - \left(\frac{1}{3} - \epsilon\right))$ if $u_n(t) > 0$. These two observations imply that (27)+(28) $= -u_n^2(t) \leq 0$. This proves that $(p_n(t+1) - p_n^*)^2 \leq (26)$.

By using this result in the definition of ΔV_t we get $\mathbf{E}[\Delta V_t]$

$$\begin{aligned} &\leq \frac{1}{2} \sum_n \left[y_n(\tau^n(t)) - \left(\frac{1}{3} - \epsilon\right) \right]^2 \\ &\quad + \sum_n (p_n(t) - p_n^*) \left[y_n(\tau^n(t)) - \left(\frac{1}{3} - \epsilon\right) \right] \\ &\leq B + \sum_n (p_n(t) - p_n^*) \left[y_n(\tau^n(t)) - \left(\frac{1}{3} - \epsilon\right) \right] \\ &= B + \sum_n (p_n(t) - p_n^*) \left[y_n(\tau^n(t)) - y_n^* \right] \\ &\quad + \sum_n (p_n(t) - p_n^*) \left(y_n^* - \left(\frac{1}{3} - \epsilon\right) \right) \\ &\stackrel{(a)}{\leq} B + \sum_n (p_n(t) - p_n^*) \left[y_n(\tau^n(t)) - y_n^* \right] \\ &= B + \sum_n (p_n(t) - p_n^*) [y_n(t) - y_n^*] \\ &\quad + \sum_n (p_n(t) - p_n^*) \left[y_n(\tau^n(t)) - y_n(t) \right], \end{aligned}$$

for some constant B , where inequality (a) follows from (16). Now, looking at the second term:

$$\begin{aligned}
& \sum_n (p_n(t) - p_n^*) (y_n(t) - y_n^*) \\
&= \sum_n (p_n(t) - p_n^*) \left(\sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f(t) - x_f^*}{c_l} H_l^f \right) \\
&= \sum_f (x_f(t) - x_f^*) \left(\sum_{(n,m) \in R(f)} \frac{p_n(t) + p_m(t)}{c_{(n,m)}} - \sum_{l \in R(f)} \frac{p_n^* + p_m^*(t)}{c_{(n,m)}} \right) \\
&= \sum_f (x_f(t) - x_f^*) (q_f(t) - q_f^*) \\
&= \sum_f (q_f(t) - q_f^*) (x_f(t) - \hat{x}_f(t)) \\
&\quad + \sum_f (q_f(t) - q_f^*) (\hat{x}_f(t) - x_f^*)
\end{aligned}$$

Therefore, we can rewrite the upper bound of $\mathbf{E}[\Delta V_t]$ as:

$$\mathbf{E}[\Delta V_t] \leq B + \sum_f (q_f(t) - q_f^*) (\hat{x}_f(t) - x_f^*) \quad (30)$$

$$+ \sum_f (q_f(t) - q_f^*) (x_f(t) - \hat{x}_f(t)) \quad (31)$$

$$+ \sum_n (p_n(t) - p_n^*) [y_n(\tau^{(n)}(t)) - y_n(t)], \quad (32)$$

where we recall that $x_f(t) = \min \left\{ M, U_f^{-1} \left(\frac{q_f(\tau^f(t))}{L} \right) \right\}$ and $\hat{x}_f(t) = \min \left\{ M, U_f^{-1} \left(\frac{q_f(t)}{L} \right) \right\}$.

By the Lemma 2, we know that for some $\gamma \in (0, 1)$,

$$\begin{aligned}
(30) &\leq -\frac{\delta}{L^\gamma} \|\mathbf{q}(t) - \mathbf{q}^*\| \mathcal{I}_{\|\mathbf{q}(t) - \mathbf{q}^*\| \geq \omega L^\gamma} \\
&\quad + \zeta \mathcal{I}_{\|\mathbf{q}(t) - \mathbf{q}^*\| \leq \omega L^\gamma}
\end{aligned} \quad (33)$$

Alternatively, by Lemma 3, we can write

$$(30) \leq -cL \|\hat{\mathbf{x}}(t) - \mathbf{x}^*\|^2 + B. \quad (34)$$

Next, let us consider (31). From the Taylor's expansion,

$$|x_f(t) - \hat{x}_f(t)| \leq \left| \frac{q_f(t) - q_f(\tau^f(t))}{L U_f''(\tilde{x})} \right|$$

for some $\tilde{x} \in [0, M]$.

It is not difficult to see that we can find some $\bar{B} < \infty$ which satisfies $|p_n(t) - p_n(t-1)| \leq \bar{B}, \forall n$. Then we have:

$$\begin{aligned}
& (q_f(t) - q_f^*) (x_f(t) - \hat{x}_f(t)) \\
&\leq |q_f(t) - q_f^*| \frac{|q_f(t) - q_f(\tau^f(t))|}{L |U_f''(\tilde{x})|} \\
&\leq |q_f(t) - q_f^*| \frac{2N_{max} |p_{\tilde{n}(f)}(t) - p_{\tilde{n}(f)}(\tau_{\tilde{n}(f)}^f(t))|}{L |U_f''(\tilde{x})| c_{\tilde{l}(f)}}
\end{aligned}$$

$$\begin{aligned}
&\leq |q_f(t) - q_f^*| \frac{2N_{max} \bar{B} |t - \tau_{\tilde{n}(f)}^f(t)|}{L |U_f''(\tilde{x})| c_{\tilde{l}(f)}} \\
&\leq |q_f(t) - q_f^*| \frac{2N_{max} \bar{B} |t - \tau_{\tilde{n}(f)}^f(t)|}{L |U_f''(\tilde{x})|}
\end{aligned} \quad (35)$$

where N_{max} is the maximum number of nodes along any flow's path, and

$$(\tilde{l}(f), \tilde{n}(f)) = \arg \max_{l \in R(f), n \in l} \left| \frac{p_n(t) - p_n(\tau_n^f(t))}{c_l} \right|.$$

Here the notation $n \in l$ means that link l is incident on node n . To simplify the notation, let us define $\Delta_f^t \triangleq |t - \tau_{\tilde{n}(f)}^f(t)|$. Noting that because $p_n^* = L\mu_n^*$ and q_f^* satisfies (11), we have $q_f^* = \rho L$ for some constant ρ . Then, for each flow f , we get the following cases:

Case 1: $q_f(t) \leq LU_f'(M)$. Using the condition (1) on utility functions, we have:

$$\begin{aligned}
(35) &= |q_f(t) - q_f^*| \frac{2N_{max} \bar{B} \Delta_f^t}{L |U_f''(\tilde{x})|} \\
&\leq L |U_f'(M) + \rho| \frac{2mN_{max} \bar{B} \Delta_f^t}{L} = C_1 \Delta_f^t
\end{aligned}$$

where the constant $C_1 = 2 |U_f'(M) + \rho| mN_{max} \bar{B} < \infty$.

Case 2: $LU_f'(M) < q_f(t) \leq q_f(\tau^f(t))$. Then, $\tilde{x} = U_f'^{-1}(\frac{\tilde{q}}{L})$ for some $q_f(t) \leq \tilde{q} \leq q_f(\tau^f(t))$. Also, from the condition (2) on utility functions, we have: $|U_f''(U_f'^{-1}(\frac{\tilde{q}}{L}))| \geq c_1 \frac{\tilde{q}}{L}$ for some constant $c_1 > 0$.

Therefore,

$$\begin{aligned}
(35) &= |q_f(t) - q_f^*| \frac{2N_{max} \bar{B} \Delta_f^t}{L |U_f''(U_f'^{-1}(\frac{\tilde{q}}{L}))|} \\
&\leq \left| 1 - \frac{q_f^*}{q_f(t)} \right| \frac{2q_f(t) N_{max} \bar{B} \Delta_f^t}{L c_1 \frac{\tilde{q}}{L}} \\
&\leq \left| 1 + \frac{\rho}{U_f'(M)} \right| \frac{2N_{max} \bar{B} \Delta_f^t}{c_1} = C_2 \Delta_f^t
\end{aligned}$$

where the constant $C_2 = \left| 1 + \frac{\rho}{U_f'(M)} \right| \frac{2N_{max} \bar{B}}{c_1} < \infty$.

Case 3: $LU_f'(M) < q_f(\tau^f(t)) < q_f(t)$. Then, $\tilde{x} = U_f'^{-1}(\frac{\tilde{q}}{L})$ for some $q_f(\tau^f(t)) \leq \tilde{q} \leq q_f(t)$. And we have:

$$\begin{aligned}
(35) &= |q_f(t) - q_f^*| \frac{2N_{max} \bar{B} \Delta_f^t}{L |U_f''(U_f'^{-1}(\frac{\tilde{q}}{L}))|} \\
&\leq |q_f(t) - q_f^*| \frac{2N_{max} \bar{B} \Delta_f^t}{L c_1 \frac{q_f(\tau^f(t))}{L}} \\
&\leq \left| 1 - \frac{q_f^*}{q_f(t)} \right| \frac{2N_{max} \bar{B}}{c_1} \frac{q_f(t)}{q_f(\tau^f(t))} \Delta_f^t \\
&\leq \left| 1 + \frac{\rho}{U_f'(M)} \right| \frac{2N_{max} \bar{B}}{c_1} \left[\frac{q_f(t) - q_f(\tau^f(t))}{U_f'(M)L} + 1 \right] \Delta_f^t \\
&= \frac{C_3}{L} (\Delta_f^t)^2 + C_2 \Delta_f^t
\end{aligned}$$

where the constant $C_3 = \left|1 + \frac{\rho}{U'_f(M)}\right| \frac{(2N_{max}\bar{B})^2}{c_1 U'_f(M)} < \infty$.

Case 4: $q_f(\tau^f(t)) \leq LU'_f(M) < q_f(t)$. Also using the condition (1) on the utility functions,

$$\begin{aligned}
(35) &\leq |q_f(t) - q_f^*| \frac{2mN_{max}\bar{B}\Delta_f^t}{L} \\
&\leq (|q_f(t) - q_f(\tau^f(t))| + |q_f(\tau^f(t)) - q_f^*|) \times \\
&\quad \frac{2mN_{max}\bar{B}\Delta_f^t}{L} \\
&\leq (N_{max}\bar{B}\Delta_f^t + L|U'_f(M) + \rho|) \frac{2mN_{max}\bar{B}\Delta_f^t}{L} \\
&= \frac{C_4}{L} (\Delta_f^t)^2 + C_5 \Delta_f^t
\end{aligned}$$

where the constant $C_4 = 2m(N_{max}\bar{B})^2 < \infty$, and $C_5 = 2mN_{max}\bar{B}|U'_f(M) + \rho| < \infty$.

Thus, combining the four cases by defining $C_6 = \max\{C_3, C_4\}$ and $C_7 = \max\{C_1, C_2, C_5\}$, we have

$$(31) \leq \frac{C_6}{L} \sum_f (\Delta_f^t)^2 + C_7 \sum_f \Delta_f^t.$$

Finally, we consider (32). Recall that

$$y_n(\tau^{(n)}(t)) - y_n(t) = \sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f(\tau_f^{(n)}(t)) - x_f(t)}{c_l} H_l^f$$

Then we can write (32) as

$$\begin{aligned}
&\sum_n (p_n(t) - p_n^*) \left[y_n(\tau^{(n)}(t)) - y_n(t) \right] \\
&\leq \sum_n |p_n(t) - p_n^*| \left| \sum_{l \in \mathcal{E}(n)} \sum_{f \in \mathcal{F}} \frac{x_f(\tau_f^{(n)}(t)) - x_f(t)}{c_l} H_l^f \right| \\
&\leq \sum_n |p_n(t) - p_n^*| 2F_{max} \left| \frac{x_{\hat{f}(n)}(\tau_{\hat{f}(n)}^{(n)}(t)) - x_{\hat{f}(n)}(t)}{c_{\hat{l}(n)}} \right|
\end{aligned}$$

where F_{max} is the maximum number of flows which go through any node, and

$$(\hat{l}(n), \hat{f}(n)) = \arg \max_{l \in \mathcal{E}(n), f \in \mathcal{F}} \left| \frac{x_f(\tau_f^{(n)}(t)) - x_f(t)}{c_l} H_l^f \right|.$$

Also, for every flow f that goes through node n , we always have $p_n \leq q_f$. Therefore,

$$\begin{aligned}
(32) &\leq \sum_n \frac{2F_{max}}{c_{\hat{l}(n)}} \left| q_{\hat{f}(n)}(t) + p_n^* \right| \\
&\quad \times \left| x_{\hat{f}(n)}(\tau_{\hat{f}(n)}^{(n)}(t)) - x_{\hat{f}(n)}(t) \right| \\
&\leq \sum_n 2F_{max} \left| q_{\hat{f}(n)}(t) + p_n^* \right| \\
&\quad \times \left| x_{\hat{f}(n)}(\tau_{\hat{f}(n)}^{(n)}(t)) - x_{\hat{f}(n)}(t) \right|
\end{aligned}$$

Let

$$(\bar{l}(n), \bar{n}(n)) = \arg \max_{l \in R(\hat{f}(n)), m \in l} \left| \frac{p_m(t) - p_m(\tau_m^{\hat{f}(n)}(t))}{c_l} \right|,$$

and $\Delta_n^t \triangleq \left| t - \tau_{\bar{n}(n)}^{\hat{f}(n)}(t) \right|$. Then, using a similar technique as in the analysis of (31), we can finally argue that:

$$(32) \leq \frac{C_8}{L} \sum_n (\Delta_n^t)^2 + C_9 \sum_n \Delta_n^t$$

for some constants $C_8, C_9 < \infty$.

Thus, if we use the upper bound in (33), we have:

$$\begin{aligned}
\mathbf{E}[\Delta V_t] &\leq -\frac{\delta}{L\gamma} \|\mathbf{q}(t) - \mathbf{q}^*\| \mathcal{I}_{\|\mathbf{q}(t) - \mathbf{q}^*\| \geq \omega L\gamma} + \zeta \mathcal{I}_{\|\mathbf{q}(t) - \mathbf{q}^*\| \leq \omega L\gamma} \\
&\quad + \sum_f \left[\frac{C_6}{L} (\Delta_f^t)^2 + C_7 \Delta_f^t \right] + \sum_n \left[\frac{C_8}{L} (\Delta_n^t)^2 + C_9 \Delta_n^t \right] \\
&\leq -\frac{\delta}{L\gamma} \|\mathbf{q}(t) - \mathbf{q}^*\| + \hat{B} + \hat{C} \|\Delta(t)\|^2,
\end{aligned}$$

for some $\hat{B}, \hat{C} < \infty$.

Instead, if we use the upper bound in (34), we get

$$\mathbf{E}[\Delta V_t] \leq -cL \|\mathbf{x}(t) - \mathbf{x}^*\|^2 + \tilde{B} + \tilde{C} \|\Delta(t)\|^2.$$

This completes the proof of Theorem 3. \square

Corollary 1:

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{E} \left[\frac{\|\mathbf{q}(t) - \mathbf{q}^*(\epsilon)\|}{L} \right] \leq \frac{B}{\delta L^{1-\gamma}}$$

Proof:

We start by taking the expectation of both sides of the expression (21) over $\mathcal{P}(t)$, and then over $\Delta(t)$:

$$\mathbf{E}[V(\mathbf{p}(t+1)) - V(\mathbf{p}(t)) \mid \Delta(t)]$$

$$\leq -\frac{\delta}{L\gamma} \mathbf{E}[\|\mathbf{q}(t) - \mathbf{q}^*(\epsilon)\| \mid \Delta(t)] + \hat{C} \|\Delta(t)\|^2 + \hat{B}$$

$$\mathbf{E}[V(\mathbf{p}(t+1)) - V(\mathbf{p}(t))]$$

$$\leq -\frac{\delta}{L\gamma} \mathbf{E}[\|\mathbf{q}(t) - \mathbf{q}^*(\epsilon)\|] + \hat{C} \mathbf{E}[\|\Delta(t)\|^2] + \hat{B}$$

Then apply the Lemma 1:

$$\begin{aligned}
\mathbf{E} \left[\left| t - \tau_n^f(t) \right|^2 \right] &\leq \sum_T (TD)^2 e^{-\nu(D)T} \\
&= D^2 \sum_T (T)^2 e^{-\nu(D)T} = C_{10}
\end{aligned}$$

for some constant $C_{10} < \infty$. We can also obtain the similar bound:

$$\begin{aligned}
\mathbf{E} \left[\left| t - \tau_f^{(n)}(t) \right|^2 \right] &\leq D^2 \sum_T (T)^2 e^{-\eta(D)T} \\
&= C_{11}
\end{aligned}$$

for some constant $C_{11} < \infty$. Therefore, we will have $\mathbf{E}[\|\Delta(t)\|^2]$ is bounded by some constant $C < \infty$, or

$$\mathbf{E}[V(\mathbf{p}(t+1)) - V(\mathbf{p}(t)) \mid \tau(t)]$$

$$\leq -\frac{\delta}{L\gamma} \mathbf{E}[\|\mathbf{q}(t) - \mathbf{q}^*(\epsilon)\|] + \hat{C}C + \hat{B}$$

Let $B = C\hat{C} + \hat{B}$, and we vary t from 0 up to T :

$$\mathbf{E}[V(1) - V(0)] \leq -\frac{\delta}{L\gamma} \mathbf{E}[\|\mathbf{q}(0) - \mathbf{q}^*(\epsilon)\|] + B$$

\vdots

$$\mathbf{E}[V(T) - V(T-1)] \leq -\frac{\delta}{L\gamma} \mathbf{E}[\|\mathbf{q}(T-1) - \mathbf{q}^*(\epsilon)\|] + B$$

Therefore, adding both sides of the inequalities and rearranging the terms, we get

$$\begin{aligned} \frac{1}{T} \sum_{t=0}^T E[\|\mathbf{q}(t) - \mathbf{q}^*(\epsilon)\|] &\leq \frac{\mathbf{E}[V(0) - V(T)]L^\gamma}{T\delta} + \frac{BL^\gamma}{\delta} \\ &\leq \frac{\mathbf{E}[V(0)]L^\gamma}{T\delta} + \frac{BL^\gamma}{\delta} \end{aligned}$$

Taking the lim sup as T goes to infinity yields the result. \blacksquare

Similarly, we can get an upper bound on the rate vectors.

Corollary 2:

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} E[\|\hat{\mathbf{x}}(t) - \mathbf{x}^*(\epsilon)\|^2] \leq \frac{B}{L}$$

Proof: The proof follows the exact same arguments as in Corollary 1, applied to (22). \blacksquare

Corollaries 1 and 2 respectively argue that as L increases, $\frac{\mathbf{q}(t)}{L}$ and $\mathbf{x}(t)$ tend to Ξ_ϵ and $\mathbf{x}^*(\epsilon)$ in the stated sense. Next, we have the result for the positive recurrence of the Markov chain.

Theorem 4: The Markov chain $(\mathcal{P}(t), \Delta(t))$ is irreducible, aperiodic and positive recurrent.

Proof: It is easy to see that this Markov chain is irreducible and aperiodic. From the result of Corollary 1, we can find some $B_1 < +\infty$ such that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{E}[\|\mathbf{q}(t)\|] \leq B_1 L^\gamma.$$

Thus, there exists some $B_2 < +\infty$ such that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbf{E} \left[\sum_n p_n(t) \right] \leq B_2 L^\gamma. \quad (36)$$

If the Markov chain $(\mathcal{P}(t), \Delta(t))$ is not positive recurrent,

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbf{P} \left(\sum_{i=0}^{t-\tau_m} \sum_n p_n(t-i) + \sum_{n,f} |t - \tau_n^f(t)| \right. \\ \left. + \sum_{n,f} |t - \tau_f^{(n)}(t)| \leq \hat{M} \right) = 0, \quad \forall \hat{M} > 0. \end{aligned}$$

That means, for every $\epsilon \in (0, 1)$, we can find a K such that

$$\begin{aligned} \mathbf{P} \left(\sum_{i=0}^{t-\tau_m} \sum_n p_n(t-i) + \sum_{n,f} |t - \tau_n^f(t)| \right. \\ \left. + \sum_{n,f} |t - \tau_f^{(n)}(t)| > \hat{M} \right) \geq 1 - \epsilon, \quad \forall t \geq K. \end{aligned}$$

Also, $p_n(t) - \frac{1}{3} \leq p_n(t+1) \leq p_n(t) + \hat{M}$. Let us define $\Delta t \triangleq t - \tau_m$, then we have:

$$\begin{aligned} \sum_{i=0}^{t-\tau_m} \sum_n p_n(t-i) + \sum_{n,f} |t - \tau_n^f(t)| + \sum_{n,f} |t - \tau_f^{(n)}(t)| \\ \leq (t - \tau_m + 1) \sum_n p_n(t) + N \sum_{i=0}^{t-\tau_m} \frac{i}{3} + 2N(t - \tau_m) \\ = (\Delta t + 1) \sum_n p_n(t) + N \frac{\Delta t(\Delta t + 1)}{6} + 2N\Delta t \end{aligned}$$

where N is the number of all flows and nodes in the network. Therefore, $\forall t \geq K$,

$$\begin{aligned} \mathbf{P} \left((\Delta t + 1) \sum_n p_n(t) + N \frac{\Delta t(\Delta t + 13)}{6} > \hat{M} \right) \\ \geq \mathbf{P} \left(\sum_{i=0}^{t-\tau_m} \sum_n p_n(t) + \sum_{n,f} |t - \tau_n^f(t)| \right. \\ \left. + \sum_{n,f} |t - \tau_f^{(n)}(t)| > \hat{M} \right) \geq 1 - \epsilon. \end{aligned}$$

Also, $\mathbf{P} \left((\Delta t + 1) \sum_n p_n(t) + N \frac{\Delta t(\Delta t + 13)}{6} > \hat{M} \right)$

$$\begin{aligned} = \sum_{\tau=0}^{\infty} \mathbf{P}(\Delta t = \tau) \mathbf{P} \left((\Delta t + 1) \sum_n p_n(t) \right. \\ \left. + N \frac{\Delta t(\Delta t + 13)}{6} > \hat{M} \mid \Delta t = \tau \right) \\ \leq \sum_{\tau=R}^{\infty} \mathbf{P}(\Delta t = \tau) + \sum_{\tau=0}^R \mathbf{P}(\Delta t = \tau) \\ \mathbf{P} \left((R+1) \sum_n p_n(t) + \frac{NR(R+13)}{6} > \hat{M} \right) \end{aligned}$$

Notice that $\mathbf{P}(\Delta t = \tau)$ is exponentially decayed (see Lemma 1). Hence, for every $\delta \in (0, 1)$, we can find a R such that $\sum_{\tau=R}^{\infty} \mathbf{P}(\Delta t = \tau) < \delta$. Also, let $\tilde{M} = \frac{\hat{M}}{R+1} - \frac{NR(R+13)}{6(R+1)}$. Finally, we get

$$\mathbf{P} \left(\sum_n p_n(t) > \tilde{M} \right) \geq 1 - \epsilon - \delta, \quad \forall t \geq K.$$

Therefore,

$$\begin{aligned} \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=K}^{T-1} \mathbf{E} \left[\sum_n p_n(t) \right] \geq \lim_{T \rightarrow \infty} \frac{(T-K)}{T} \tilde{M}(1 - \epsilon - \delta) \\ = \tilde{M}(1 - \epsilon - \delta) \end{aligned}$$

Choose $\hat{M}, \epsilon, \delta$ such that $\tilde{M}(1 - \epsilon - \delta) > B_2 L^\gamma$, then we get a contradiction with (36). Thus, the Markov chain $(\mathcal{P}(t), \Delta(t))$ is positive recurrent. \blacksquare

V. DISCUSSION ON THE STABILITY OF THE SYSTEM

In this section, based on the analysis in Sections III-B and IV-B, we will provide a discussion why the regulators and the queues are stable for the asynchronous, discrete-time model.

Recall that in Theorem 2, we proved the stability of the continuous-time, fluid limit system by using the fact that $\mathbf{x}(t)$ will stay inside a δ neighborhood of $\mathbf{x}^*(\epsilon)$ for t large enough. Subsequently, in Section IV-B, we proved that $\mathbf{x}(t)$ of the asynchronous, discrete-time model can be made to be arbitrarily close to $\mathbf{x}^*(\epsilon)$ in an asymptotic and expected manner (c.f. Corollary 2). Moreover, in Theorem 4, we proved the positive recurrence of the Markov chain and can consequently assume that the arrival processes are stationary. In fact, by

using Jensen's inequality and Ergodic theorem, we can argue that for any $\tilde{\epsilon} > 0$, and $k \geq 0$,

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=k}^{k+T-1} \sum_f E [|x_f(t) - x_f^*(\epsilon)|] \leq \sqrt{\frac{B}{L}} + \tilde{\epsilon}$$

Hence, given any $\tilde{\delta} > 0$, we can choose ϵ and $\tilde{\epsilon}$ small enough, and T and L large enough, so that for any $k \geq 0$,

$$\frac{1}{T} \sum_{t=k}^{k+T-1} E [|x_f(t) - x_f^*|] < \tilde{\delta}, \quad \forall f \in \mathcal{F}, \text{ and } T \geq T.$$

Then, following the argument in [7] which essentially claims that if the continuous-time, fluid limit system is stable and the arrival processes satisfy those above conditions, then the discrete-time, stochastic system is also stable. However, we do not provide the details here due to space limitation.

VI. CONCLUSIONS

In this paper, we have considered the fair resource allocation problem in multi-hop wireless networks with a specific interference model, and developed a cross-layer algorithm to solve it. More specifically, we proposed a congestion control algorithm for transport layer, and a fully distributed scheduling algorithm for MAC layer. The main contribution of the paper is to allow for unbounded delays in the feedback between the components of the network. This is a crucial step towards being able to actually implement a congestion control mechanism in a real network, for time-varying delay is an inseparable ingredient of a wireless network.

We proved that even when all the sources and nodes operate in a totally asynchronous manner, our algorithms can achieve flow rates that are arbitrarily close to the fair operating point. Extensions to other interference models is a topic for future research.

APPENDIX

Proof: [Theorem 2]

Consider the following Lyapunov function:

$$W(\mathbf{Q}, \mathbf{R}) = W_1(\mathbf{Q}) + \xi W_2(\mathbf{Q}, \mathbf{R}),$$

$$W_1(\mathbf{Q}) = \frac{1}{2} \sum_{(n,m) \in \mathcal{L}} \frac{Q_{nm}}{c(n,m)} \left(\sum_{(h,k) \in \Upsilon_{nm}} \frac{Q_{hk}}{c(h,k)} \right)$$

$$W_2(\mathbf{Q}, \mathbf{R}) = \frac{1}{2} \sum_{(n,m) \in \mathcal{L}} \frac{(\sum_f R_{nm}^f + Q_{nm})^2}{c(n,m)}$$

where ξ is a positive parameter which will be chosen later.

First, consider $W_1(\cdot)$:

$$\begin{aligned} \dot{W}_1(\mathbf{Q}(t)) &= \frac{1}{2} \sum_{(n,m) \in \mathcal{L}} \left[\frac{\dot{Q}_{nm}(t)}{c(n,m)} \left(\sum_{(h,k) \in \Upsilon_{nm}} \frac{Q_{hk}(t)}{c(h,k)} \right) \right. \\ &\quad \left. + \left(\frac{Q_{nm}(t)}{c(n,m)} \right) \left(\sum_{(h,k) \in \Upsilon_{nm}} \frac{\dot{Q}_{hk}(t)}{c(h,k)} \right) \right] \\ &= \sum_{(n,m) \in \mathcal{L}} \left(\frac{Q_{nm}(t)}{c(n,m)} \right) \left(\sum_{(h,k) \in \Upsilon_{nm}} \frac{\dot{Q}_{hk}(t)}{c(h,k)} \right) \end{aligned}$$

We know from our analysis of the congestion controller that $\mathbf{x}(t) \rightarrow \mathbf{x}^*(\epsilon)$ as $t \rightarrow \infty$. Therefore, for every $\rho > 0$, there exists $T < \infty$ such that $|x_f(t) - x_f^*(\epsilon)| < \rho$ for $t \geq T$ and for all f .

Also, noting a standard fact that the projection in (18) can be ignored, we have:

$$\begin{aligned} \dot{W}_1(\mathbf{Q}(t)) &\leq \sum_{\{(n,m) \in \mathcal{L}\}} \left\{ \left(\frac{Q_{nm}(t)}{c(n,m)} \right) \right. \\ &\quad \left. \times \sum_{(h,k) \in \Upsilon_{nm}} \left[\frac{(\sum_f (x_f(t) + N_{max}\epsilon_r) H_{hk}^f)}{c(h,k)} - \pi_{hk} \right] \right\}. \end{aligned}$$

In the fluid limit, the scheduling rule $\pi(t)$ satisfies the condition (6) for every $(n, m) \in \mathcal{L}$ with $Q_{nm}(t) > 0$. Thus,

$$\begin{aligned} \dot{W}_1(\mathbf{Q}(t)) &\leq \sum_{\{(n,m) \in \mathcal{L}\}} \left\{ \left(\frac{Q_{nm}(t)}{c(n,m)} \right) \right. \\ &\quad \left. \times \sum_{(h,k) \in \Upsilon_{nm}} \left[\frac{(\sum_f (x_f(t) + N_{max}\epsilon_r) H_{hk}^f)}{c(h,k)} - 1 \right] \right\}. \end{aligned}$$

Because $\mathbf{x}^*(\epsilon)$ is strictly inside $\frac{\Omega}{2}$, given any $\epsilon > 0$, we can find some $\varphi > 0$ for which $(x_f^*(\epsilon) + \varphi)_f \in \frac{\Omega}{2}$. Then, we can choose $\rho > 0$ and $\epsilon_r > 0$ small such that $\rho + N_{max}\epsilon_r < \varphi$, or $(x_f^*(\epsilon) + \rho + N_{max}\epsilon_r)_f$ is strictly inside $\frac{\Omega}{2}$. Now, using the same argument as in [23], we have:

$$\sum_{(h,k) \in \Upsilon_{nm}} \left[\frac{(\sum_f (x_f(t) + N_{max}\epsilon_r) H_{hk}^f)}{c(h,k)} \right] - 1 \leq -\theta < 0$$

for some $\theta > 0$ if $Q_{nm}(t) > 0$. Thus,

$$\dot{W}_1(\mathbf{Q}(t)) \leq -\theta \sum_{\{(n,m) \in \mathcal{L}\}} \frac{Q_{nm}(t)}{c(n,m)}.$$

Now, let us consider $W_2(\cdot)$. Recall that

$$\begin{aligned} \dot{Q}_{nm}(t) + \sum_f \dot{R}_{nm}^f(t) &= \left(\sum_f (x_f(t) + K_f^n \epsilon_r) \mathcal{I}_{R_{nm}^f(t) > 0} H_{nm}^f - P_{nm}(t) \right)_{Q_{nm}}^+ \\ &\quad + \sum_f (P_{nm}^f(t) - (x_f(t) + (K_f^n + 1) \epsilon_r))_{R_{nm}^f}^+ \end{aligned}$$

Note that $P_{nm}^f(t)$ can only be positive if $Q_{nm} > 0$. Then we have the following cases:

- If $Q_{nm} > 0$ and $R_{nm}^f > 0$ for all f , we remove the projections:

$$\dot{Q}_{nm}(t) + \sum_f \dot{R}_{nm}^f(t) \leq -\sum_f \epsilon_r$$

- If $Q_{nm} = 0$ and $R_{nm}^f > 0$ for all f , then $P_{nm}(t) = 0$, $\dot{Q}_{nm}(t) + \sum_f \dot{R}_{nm}^f(t)$

$$\begin{aligned} &= \left(\sum_f (x_f(t) + K_f^n \epsilon_r) \mathcal{I}_{R_{nm}^f(t) > 0} H_{nm}^f \right) \\ &\quad - \sum_f (x_f(t) + (K_f^n + 1) \epsilon_r) \leq -\sum_f \epsilon_r \end{aligned}$$

- If $Q_{nm} > 0$ and $R_{nm}^g = 0$ for some g , then there exists some constant χ such that

$$\begin{aligned} & \dot{Q}_{nm}(t) + \sum_f \dot{R}_{nm}^f(t) \\ &= \sum_f (x_f(t) + K_f^n \epsilon_r) \mathcal{I}_{R_{nm}^f(t) > 0} H_{nm}^f - P_{nm}(t) \\ &+ \sum_{f \neq g} P_{nm}^f(t) - (x_f(t) + (K_f^n + 1) \epsilon_r) \\ &+ (P_{nm}^g(t) - (x_g(t) + (K_g^n + 1) \epsilon_r))^+ \\ &\leq -\epsilon_r \mathcal{I}_{\{P_{nm}^g(t) \geq x_g(t) + (K_g^n + 1) \epsilon_r\}} \\ &+ \chi \mathcal{I}_{\{P_{nm}^g(t) < x_g(t) + (K_g^n + 1) \epsilon_r\}} \leq \chi \end{aligned}$$

Therefore, we can write

$$\begin{aligned} & \dot{W}_2(\mathbf{Q}(t), \mathbf{R}(t)) \\ &= \sum_{(n,m) \in \mathcal{L}} \frac{(\sum_f R_{nm}^f(t) + Q_{nm}(t))}{c(n,m)} \\ &\times \left(\dot{Q}_{nm}(t) + \sum_f \dot{R}_{nm}^f(t) \right) \\ &\leq -\epsilon_r \sum_{(n,m) \in \mathcal{L}} \frac{(\sum_f R_{nm}^f(t) + Q_{nm}(t))}{c(n,m)} \mathcal{I}_{\{R_{nm}^f > 0\}} \\ &+ \chi \sum_{(n,m) \in \mathcal{L}} \frac{Q_{nm}(t)}{c(n,m)} \end{aligned}$$

Therefore,

$$\begin{aligned} & \dot{W}(\mathbf{Q}(t), \mathbf{R}(t)) = \dot{W}_1(\mathbf{Q}(t), \mathbf{R}(t)) + \xi \dot{W}_2(\mathbf{Q}(t), \mathbf{R}(t)) \\ &\leq -(2\theta - \xi\chi) \sum_{(n,m) \in \mathcal{L}} \frac{Q_{nm}(t)}{c(n,m)} \\ &- \xi \epsilon_r \sum_{(n,m) \in \mathcal{L}} \frac{(\sum_f R_{nm}^f(t) + Q_{nm}(t))}{c(n,m)} \mathcal{I}_{\{R_{nm}^f > 0\}} \end{aligned}$$

We can easily choose ξ such that $2\theta - \xi\chi > 0$. Thus,

$$\begin{aligned} \dot{W}(\mathbf{Q}(t), \mathbf{R}(t)) &< 0 & \text{if } \mathbf{Q} + \mathbf{R} \neq \mathbf{0} \\ &= 0 & \text{if } \mathbf{Q} + \mathbf{R} = \mathbf{0} \end{aligned}$$

Then the result follows from Lyapunov's stability theorem. \blacksquare

Proof: [Lemma 2] We define

$$\Phi(t) = \sum_f (q_f(t) - q_f^*(\epsilon)) (\hat{x}_f(t) - x_f^*(\epsilon)),$$

and

$$\hat{f} = \arg \max_f |q_f(t) - q_f^*|.$$

Again, we will omit (ϵ) for notational convenience. Noting that for all f we have

$$(q_f(t) - q_f^*) (\hat{x}_f(t) - x_f^*) \leq 0$$

due to the fact that $U_f'^{-1}(\cdot)$ is decreasing in its parameter, and that $x_f^*(t) = U_f'^{-1}(\frac{q_f^*(t)}{L})$, we can write

$$\Phi(t) \leq -|q_{\hat{f}}(t) - q_{\hat{f}}^*| \left| \hat{x}_{\hat{f}}(t) - x_{\hat{f}}^* \right|$$

There are two cases to consider: If $\hat{x}_{\hat{f}}(t) = M$, then $\hat{x}_{\hat{f}}(t) - x_{\hat{f}}^* > M - c_{max} > c_{max}$, since M is chosen to be larger than $2 \max_l c_l = 2c_{max}$.

If, on the other hand, $\hat{x}_{\hat{f}}(t) < M$, then we have

$$\left| \hat{x}_{\hat{f}}(t) - x_{\hat{f}}^* \right| = x_{\hat{f}}^* \left| \frac{U_{\hat{f}}'^{-1}\left(\frac{q_{\hat{f}}(t)}{L}\right)}{U_{\hat{f}}'^{-1}\left(\frac{q_{\hat{f}}^*}{L}\right)} - 1 \right|. \quad (37)$$

Notice that

$$q_f(t) = \begin{cases} q_f^* - |q_f(t) - q_f^*| \geq 0 & \text{if } q_f(t) - q_f^* \leq 0, \\ q_f^* + |q_f(t) - q_f^*| \geq 0 & \text{if } q_f(t) - q_f^* \geq 0. \end{cases}$$

Assuming that $U_f'^{-1}(\cdot)$ is a decreasing, convex function, we can write

$$\left| \frac{U_{\hat{f}}'^{-1}\left(\frac{q_{\hat{f}}^*}{L} - \frac{|q_{\hat{f}}(t) - q_{\hat{f}}^*|}{L}\right)}{U_{\hat{f}}'^{-1}\left(\frac{q_{\hat{f}}^*}{L}\right)} - 1 \right| \geq \left| \frac{U_{\hat{f}}'^{-1}\left(\frac{q_{\hat{f}}^*}{L} + \frac{|q_{\hat{f}}(t) - q_{\hat{f}}^*|}{L}\right)}{U_{\hat{f}}'^{-1}\left(\frac{q_{\hat{f}}^*}{L}\right)} - 1 \right|.$$

Therefore, we have

$$\left| \frac{U_{\hat{f}}'^{-1}\left(\frac{q_{\hat{f}}(t)}{L}\right)}{U_{\hat{f}}'^{-1}\left(\frac{q_{\hat{f}}^*}{L}\right)} - 1 \right| \geq \left| \frac{U_{\hat{f}}'^{-1}\left(\frac{q_{\hat{f}}^*}{L} + \frac{|q_{\hat{f}}(t) - q_{\hat{f}}^*|}{L}\right)}{U_{\hat{f}}'^{-1}\left(\frac{q_{\hat{f}}^*}{L}\right)} - 1 \right|$$

We consider the set of \mathbf{q} which satisfies $\|\mathbf{q} - \mathbf{q}^*\| \geq \omega L^\sigma$, where ω and L are positive constants and $\sigma \in (0, 1)$. We are interested in the behavior of the system as L tends to infinity. The exact values of ω and σ depend on the utility functions and other system parameters, and will be provided later in the proof.

Notice that if $\|\mathbf{q} - \mathbf{q}^*\| \geq \omega L^\sigma$, then $|q_{\hat{f}}(t) - q_{\hat{f}}^*| \geq \frac{\omega}{\sqrt{|\mathcal{N}|}} L^\sigma$. Then, we can write

$$(37) \geq x_{\hat{f}}^* \left| \frac{U_{\hat{f}}'^{-1}\left(\frac{q_{\hat{f}}^*}{L} + \frac{\omega}{\sqrt{|\mathcal{N}|} L^{1-\sigma}}\right)}{U_{\hat{f}}'^{-1}\left(\frac{q_{\hat{f}}^*}{L}\right)} - 1 \right|$$

Noting that $q_{\hat{f}}^* = \rho L$ for some $\rho > 0$, and invoking the condition (3) on the utility functions, we can write: for $\|\mathbf{q} - \mathbf{q}^*\| \geq \omega L^\sigma$,

$$\Phi(t) + B \leq -|q_{\hat{f}}(t) - q_{\hat{f}}^*| \left(x_{\hat{f}}^* c_1 L^{-\gamma} + \frac{B \sqrt{|\mathcal{N}|}}{\omega} L^{-\sigma} \right).$$

If we choose $\sigma \in (0, 1)$ such that $\gamma = \sigma$, then for large enough ω , we get the following expression for some $\delta > 0$ and $\zeta < \infty$.

$$\begin{aligned} \Phi(t) + B &\leq -\frac{\|\mathbf{q}(t) - \mathbf{q}^*\|}{L^\gamma} \delta \mathcal{I}_{\{\|\mathbf{q}(t) - \mathbf{q}^*\| \geq \omega L^\sigma\}} \\ &+ \zeta \mathcal{I}_{\{\|\mathbf{q}(t) - \mathbf{q}^*\| < \omega L^\sigma\}} \end{aligned}$$

\blacksquare

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, pp. 1936–1948, December 1992.
- [2] L. Tassiulas, "Scheduling and performance limits of networks with constantly changing topology," *IEEE Transactions on Information Theory*, vol. 43, pp. 1067–1073, May 1997.
- [3] S. Shakkottai and A. Stolyar, "Scheduling for multiple flows sharing a time-varying channel: The exponential rule," *Translations of the AMS, Series 2*, A volume in memory of F. Karpelevich, vol. 207, 2002.
- [4] S. Shakkottai, R. Srikant, and A. Stolyar, "Pathwise optimality of the exponential scheduling rule for wireless channels," *Advances in Applied Probability*, vol. 36, no. 4, pp. 1021–1045, 2004.
- [5] A. Eryilmaz, R. Srikant, and J. Perkins, "Stable scheduling policies for fading wireless channels," in *Proceedings of IEEE International Symposium on Information Theory*, 2003, p. 454.
- [6] M. Neely, E. Modiano, and C. Rohrs, "Dynamic power allocation and routing for time varying wireless networks," in *Proceedings of IEEE Infocom*, April 2003, pp. 745–755.
- [7] A. Eryilmaz, R. Srikant, and J. R. Perkins, "Stable scheduling policies for fading wireless channels," *IEEE/ACM Transactions on Networking*, vol. 13, pp. 411–424, April 2005.
- [8] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, pp. 237–252, 1998.
- [9] S. H. Low and D. E. Lapsley, "Optimization flow control, I: Basic algorithm and convergence," *IEEE/ACM Transactions on Networking*, vol. 7, no. 6, pp. 861–875, December 1999.
- [10] R. Srikant, *The Mathematics of Internet Congestion Control*. Boston, MA: Birkhauser, 2004.
- [11] A. Stolyar, "Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm," *Queueing Systems*, vol. 50, no. 4, pp. 401–457, August 2005.
- [12] A. Eryilmaz and R. Srikant, "Fair resource allocation in wireless networks using queue-length based scheduling and congestion control," in *Proceedings of IEEE Infocom*, March 2005, pp. 1794–1803.
- [13] M. Neely, E. Modiano, and C. Li, "Fairness and optimal stochastic control for heterogeneous networks," in *Proceedings of IEEE Infocom*, March 2005, pp. 1723–1734.
- [14] X. Lin and N. Shroff, "Joint rate control and scheduling in multihop wireless networks," in *Proceedings of IEEE Conference on Decision and Control*, December 2004, pp. 1484–1489.
- [15] S. Sarkar and L. Tassiulas, "End-to-end bandwidth guarantees through fair local spectrum share in wireless ad-hoc networks," in *Proceedings of IEEE Conference on Decision and Control*, December 2003, pp. 564–569.
- [16] Y. Yi and S. Shakkottai, "Hop-by-hop congestion control over a wireless multi-hop network," in *Proceedings of the IEEE Infocom*, March 2004, pp. 2548–2558.
- [17] L. Chen, S. H. Low, and J. C. Doyle, "Joint congestion control and media access control design for ad hoc wireless networks," in *Proceedings of IEEE Infocom*, March 2005, pp. 2212–2222.
- [18] B. Miller and C. Bisdikian, *Bluetooth Revealed: The Insider's Guide to an Open Specification for Global Wireless Communication*. Upper Saddle River, NJ: Prentice Hall, 2000.
- [19] D. J. Baker, J. Wieselthier, and A. Ephremides, "A distributed algorithm for scheduling the activation of links in a self-organizing mobile radio network," in *Proceedings of IEEE International Conference on Communications*, June 1982, pp. 2F.6.1–2F.6.5.
- [20] B. Hajek and G. Sasaki, "Link scheduling in polynomial time," *IEEE Transactions on Information Theory*, vol. 34, pp. 910–917, 1988.
- [21] X. Lin and N. Shroff, "The impact of imperfect scheduling on cross-layer rate control in multihop wireless networks," in *Proceedings of IEEE Infocom*, March 2005, pp. 1804–1814.
- [22] C. Humes, "A regulator stabilization technique: Kumar-Seidman revisited," *IEEE Transactions on Automatic Control*, vol. 39, no. 1, pp. 191–196, 1994.
- [23] X. Wu and R. Srikant, "Regulated maximal matching: A distributed scheduling algorithm for multi-hop wireless networks with node-exclusive spectrum sharing," in *Proceedings of the IEEE Conference on Decision and Control*, December 2005, pp. 5342–5347.
- [24] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Belmont, MA: Athena Scientific, 1997.
- [25] M. B. Pursley, H. B. Russell, and P. E. Staples, "Routing multimedia packets in a frequency-hop packet radio network," in *MILCOM 96*, 1996, pp. 220–224.
- [26] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Transactions on Networking*, vol. 8, no. 5, pp. 556–567, October 2000.
- [27] M. Kodialam and T. Nandagopal, "Characterizing achievable rates in multi-hop wireless networks: The joint routing and scheduling problem," in *Proceedings of ACM Mobicom*, September 2003, pp. 42–54.
- [28] C. E. Shannon, "A theorem on coloring the lines of a network," *Journal of Mathematical Physics*, vol. 28, pp. 148–151, 1949.
- [29] L. Ying, R. Srikant, A. Eryilmaz, and G. E. Dullerud, "Distributed fair resource allocation in cellular networks in the presence of heterogeneous delays," in *Proceedings of International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT)*, April 2005, pp. 96–105.
- [30] J. M. Borwein and A. S. Lewis, *Convex Analysis and Nonlinear Optimization: Theory and Examples (CMS Books in Mathematics)*. New York, NY: Springer-Verlag, 2000.
- [31] H. Khalil, *Nonlinear Systems, 3rd ed.* Upper Saddle River, NJ: Prentice Hall, 2002.

Loc Xuan Bui Loc Xuan Bui received the B.Tech. degree in Telecommunications and Electronics Engineering from the Posts and Telecommunications Institute of Technology, Ho Chi Minh City, in 2003. He is currently with the Department of Electrical and Computer Engineering and the Coordinated Science Lab, University of Illinois at Urbana-Champaign, where he is a PhD candidate. His research interests include communication networks, wireless communications, stochastic processes, network control and optimization.

Atilla Eryilmaz Atilla Eryilmaz (S'02, M'05) received the B.S. degree in Electrical and Electronics Engineering from Bogazici University, Istanbul, in 1999, and the M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign in 2001 and 2005, respectively. He is currently with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology as a Postdoctoral Associate. His research interests include communication networks, optimal control of multihop wireless networks, optimization, stochastic processes and network coding.

R. Srikant R. Srikant (S'90, M'91, SM'01, F'06) received his B.Tech. from the Indian Institute of Technology, Madras in 1985, his M.S. and Ph.D. from the University of Illinois in 1988 and 1991, respectively, all in Electrical Engineering. He was a Member of Technical Staff at AT&T Bell Laboratories from 1991 to 1995. He is currently with the University of Illinois at Urbana-Champaign, where he is a Professor in the Department of Electrical and Computer Engineering, and a Research Professor in the Coordinated Science Lab. He was an associate editor of *Automatica* and the *IEEE Transactions on Automatic Control*, and is currently an associate editor of the *IEEE/ACM Transactions on Networking*. He has also served on the editorial boards of special issues of the *IEEE Journal on Selected Areas in Communications* and *IEEE Transactions on Information Theory*. He was the chair of the 2002 *IEEE Computer Communications Workshop* in Santa Fe, NM and will be a program co-chair of *IEEE INFOCOM*, 2007. His research interests include communication networks, stochastic processes, queueing theory, information theory, and game theory.

Xinzhou Wu Xinzhou Wu (S'01) received his B.Tech. from Tsinghua University at Beijing in 1998, his M.S. and Ph.D. from the University of Illinois in 2000 and 2004, respectively, all in Electrical Engineering. He was a Member of Technical Staff at Flarion Technologies from 2005 to 2006. He is currently with Qualcomm Flarion Technologies, where he is a senior engineer. His research interests include wireless communication systems and networks, queueing theory and information theory.