The Impact of Imperfect Scheduling on Cross-Layer Rate Control in Multihop Wireless Networks

Xiaojun Lin and Ness B. Shroff

Center for Wireless Systems and Applications (CWSA)

School of Electrical and Computer Engineering, Purdue University

West Lafayette, IN 47907, U.S.A.

{linx, shroff}@ecn.purdue.edu *

Abstract

In this paper, we study cross-layer design for rate control in multihop wireless networks. In our previous work, we have developed an optimal cross-layered rate control scheme that jointly computes both the rate allocation and the stabilizing schedule that controls the resources at the underlying layers. However, the scheduling component in this optimal cross-layered rate control scheme has to solve a complex global optimization problem at each time, and hence is too computationally expensive for online implementation. In this paper, we study how the performance of cross-layer rate control will be impacted if the network can only use an imperfect (and potentially distributed) scheduling component that is easier to implement. We study both the case when the number of users in the system is fixed and the case with dynamic arrivals and departures of the users, and we establish desirable results on the performance bounds of cross-layered rate control with imperfect

*This work has been partially supported by the NSF grant ANI-0099137 and the Indiana 21st Century Center for Wireless Communications and Networking.

scheduling. Compared with a layered approach that does not design rate control and scheduling together, our cross-layered approach has provably better performance bounds, and usually substantially outperforms the layered approach. The insights drawn from our analyses also enable us to design a *fully distributed* cross-layered rate control and scheduling algorithm for a restrictive interference model.

Keyword: Cross-layer design, rate control, multihop wireless networks, stability, imperfect scheduling, mathematical programming/optimization, stochastic processes/queueing theory.

1 Introduction

Cross-layer design is becoming increasingly important for improving the performance of multihop wireless networks (see, e.g., [1, 2, 3, 4, 5, 6, 7, 8, 9, 10] and the reference therein). By simultaneously optimizing the control across multiple layers of the network, cross-layer design can substantially increase the network capacity, reduce interference and power consumption.

In this paper, we study the issues involved in the cross-layer design of multihop wireless networks *that employ rate control* [8, 9, 10]. Rate control (or congestion control) is a key functionality in modern communication networks to avoid congestion and to ensure fairness among the users. Although rate control has been studied extensively for *wireline* networks (see [11] for a good survey), these results cannot be applied directly to multihop wireless networks. In wireline networks, the *capacity region* (i.e., the set of feasible data rates) is of a simple form, i.e., the sum of the data rates at each link should be less than the link capacity, which is known and fixed. In multihop wireless networks, the capacity of each radio link depends on the signal and interference levels, and thus depends on the power and transmission schedule at other links. Hence, the capacity region is usually of a complex form that critically depends on the way in which resources at the underlying physical and MAC layers are scheduled. One possible way to address this difficulty is to choose a *rate region* within the capacity region, which has a simpler set of constraints similar to that of wireline networks, and compute the rate allocation within this simpler rate region [12, 13, 14]. This approach essentially attempts to make rate control oblivious of the dynamics of the underlying layers. Hence, we will refer to this approach as the *layered approach* to rate control. However, it requires prior knowledge of the capacity region in order to choose such a rate region. For many network settings, even such a rate region is difficult to find. Further, because the rate region reduces the set of feasible rates that rate control can utilize, the layered approach results in a *conservative* rate allocation.

On the other hand, the *cross-layered approach* to rate control can allocate data rates without requiring precise prior knowledge of the capacity region. Here, by the "cross-layered" approach to rate control, we mean that the network jointly optimizes both the data rates of the users and the resource allocation at the underlying layers, which include modulation, coding, power assignment and link schedules, etc. (For the rest of the paper, we will use the term *scheduling* to refer to the joint allocation of these resources at layers under rate control.) In our previous work [8], we have presented an optimal cross-layered rate control scheme and we have shown that our scheme can fully utilize the capacity of the network, maintain fairness, and improve the quality of service to the users.

However, the *scheduling* component in the optimal cross-layered rate control scheme of [8] requires solving at each iteration a global optimization problem that is usually quite difficult. In some cases, the optimization problem does not even have a polynomial-time solution. In this work, our objective is to develop a framework for cross-layered rate control that is suitable for on-line (and potentially distributed) implementation. The complexity of the scheduling component has become the main *obstacle* to developing such a solution.

To overcome this difficulty, in this paper we take a different approach. We accept the possibility that only suboptimal solutions to the scheduling problem may be computable, which we will refer to as *imperfect schedules*. Instead, we will study the impact of imperfect scheduling on the optimality of cross-layered rate control. In this paper, we have studied this impact for a large class of imperfect scheduling policies, both for the case when the number of users in the system is fixed, and for the case when users dynamically arrive and leave the network. When the number of users in the system is fixed, we are able to obtain some desirable, but weak, results on the fairness and convergence properties of cross-layered rate control with imperfect scheduling. Surprisingly, we are able to obtain far stronger results on the performance of the system when we consider dynamic arrivals and departures of the users. Our numerical results suggest that, in many network configurations, cross-layered rate control with *imperfect* scheduling can perform comparably to that with *perfect* scheduling, while significantly reducing the computation overhead of the scheduling component. Further, we find that our cross-layered approach can substantially outperform the layered approach. Finally, the insights drawn from our analysis allow us to develop a *fully distributed* rate control and scheduling scheme in a more restrictive network setting.

The rest of the paper is structured as follows. The system model is presented in Section 2. We review results with perfect scheduling in Section 3, and study the impact of imperfect scheduling in Section 4 and 5. In Section 6, we present a fully distributed cross-layered rate control algorithm. Simulation results are presented in Section 7, and the conclusion is given in Section 8.

2 The System Model

We consider a multihop wireless network with N nodes. Let $\mathcal{L}$ denote the set of node pairs (i, j) (i.e., links) such that direct transmission from node i to node j is allowed. The links are assumed to be directional. Due to the shared nature of the wireless media, the data rate r_{ij} of a link (i, j) depends not only on its own modulation/coding scheme and power assignment P_{ij}, but also on the interference due to the power assignments on other links. Let $\vec{P} = [P_{ij}, (i, j) \in \mathcal{L}]$ denote the vector of global power assignments and let $\vec{r} = [r_{ij}, (i, j) \in \mathcal{L}]$ denote the vector of data rates. We assume that $\vec{r} = u(\vec{P})$, i.e., the data rates are completely determined by the global power assignment*. The function $u(\cdot)$ is called the *rate-power function* of the system. Note that the global power assignment $\vec{P}$ and the rate-power function $u(\cdot)$ summarize the *cross-layer* control capability of the network at both the physical layer and the MAC layer. Precisely, the global power assignment determines the Signal-to-Interference-Ratio (SIR) at each link. Given the SIR, each link can choose appropriate modulation and coding schemes to achieve the data rate specified by $u(\vec{P})$. Finally, the network can schedule different sets of links to be active (and to use different power assignments) at different time to achieve maximum capacity [3]. There may be constraints on the feasible power assignment. For example, if each node has a total power constraint $P_{i,\max}$, then $\sum_{j:(i,j)\in\mathcal{L}} P_{ij} \leq P_{i,\max}$. Let Π denote the set of feasible power assignments, and let $\mathcal{R} = \{u(\vec{P}), \vec{P} \in \Pi\}$. We assume that $\text{Co}(\mathcal{R})$, the convex hull of $\mathcal{R}$, is closed and bounded. We assume that time is divided into slots and the power assignment vector $\vec{P}(t)$ is fixed during each time slot t. We will refer to $\vec{r}(t) = u(\vec{P}(t))$ as the *schedule* at time slot t.

In the rest of the paper, it is usually more convenient to index the links numerically (e.g., links $1, 2, ..., L$) rather than as node-pairs (e.g., link (i, j)). The power assignment vector and the rate vector should then be written as $\vec{P} = [P_1, ..., P_L]$ and $\vec{r} = [r_1, ..., r_L]$, respectively.

There are S users and each user $s = 1, ..., S$ has one path through the network†. Let $H = [H_s^l]$ denote the routing matrix, i.e., $H_s^l = 1$, if the path of user s uses link l, and $H_s^l = 0$, otherwise. Let x_s be the rate with which user s injects data into the network. Each user is associated with a utility function $U_s(x_s)$, which reflects the level of "satisfaction" of user s when its data rate is x_s. As is typically assumed in the rate control literature, we assume that each user s has a maximum data rate M_s and the utility function $U_s(\cdot)$ is strictly concave, non-decreasing and twice continuously differentiable on $(0, M_s]$.

*Although we have not considered channel variation, e.g., due to fading, our main results may be generalized to those cases.

†Extensions to the case with multipath routing are also possible (see [8]).

3 Cross-Layer Rate Control with Perfect Scheduling

In this section, we review the optimal cross-layered rate control scheme that we presented in [8]. We first define the *capacity region* of the system. We say that a system is *stable* if the queue lengths at all links remain finite. We say that a user rate vector $\vec{x} = [x_1, ..., x_s]$ is *feasible* if there exists a scheduling policy that can *stabilize* the system under user rates $\vec{x}$. We define the *capacity region* to be the set of *feasible* rates $\vec{x}$. It has been shown in [3, 4, 6] that the *optimal capacity region* Λ is a convex set and is given by

$$\Lambda = \left\{ \vec{x} \middle| [\sum_{s=1}^{S} H_s^l x_s] \in \mathrm{Co}(\mathcal{R}) \right\}. \tag{1}$$

where $\sum_{s=1}^{S} H_s^l x_s$ can be interpreted as the total data rate on link l. The convex hull operator $\mathrm{Co}(\cdot)$ is due to a standard time-averaging argument [3, 4, 6]. Λ is optimal in the sense that no vector $\vec{x}$ outside Λ is feasible for any scheduling policy.

In [8], we have formulated and solved the following optimal cross-layered rate control problem.

The Cross-Layered Rate Control Problem:

- Find the user rate vector $\vec{x}$ in Λ that maximizes the total system utility, i.e.,

$$\max_{0 \le x_s \le M_s} \quad \sum_{s=1}^{S} U_s(x_s) \tag{2}$$

$$\text{subject to} \quad \sum_{s=1}^{S} H_s^l x_s \le r_l \text{ for all } l \in \mathcal{L} \tag{3}$$

$$\text{and} \quad [r_l] \in \mathrm{Co}(\mathcal{R}).$$

- Find the associated scheduling policy that *stabilizes* the system.

There are two elements in this *cross-layer* control problem. One is to determine the rates with which users inject data into the network. The other is to determine when and at what rate each link in the network should transmit. Maximizing the total system utility as in (2) has been shown to be equivalent to some *fairness* objectives when the utility functions are appropriately chosen [15]. For example, utility functions of the form

$$U_s(x_s) = w_s \log x_s \tag{4}$$

correspond to *weighted proportional fairness*, where $w_s, s = 1, ..., S$ are the weights. A more general form of utility function is

$$U_s(x_s) = w_s \frac{x_s^{1-\beta}}{1-\beta}, \beta > 0. \tag{5}$$

Maximizing the total utility will corresponds to *maximizing weighted throughput* as $\beta \to 0$, *weighted proportional fairness* as $\beta \to 1$, *minimizing weighted potential delay* as $\beta \to 2$, and *max-min faginess* as $\beta \to \infty$.

We now take a duality approach to solve problem (2). We associate a Lagrange multiplier q^l for each constraint in (3). The Lagrangian is then:

$$\begin{aligned} & L(\vec{x}, \vec{r}, \vec{q}) \\ = & \sum_{s=1}^{S} U_s(x_s) - \sum_{l=1}^{L} q^l \left[\sum_{s=1}^{S} H_s^l x_s - r_l \right] \\ = & \sum_{s=1}^{S} \left[U_s(x_s) - \sum_{l=1}^{L} H_s^l q^l x_s \right] + \sum_{l=1}^{L} q^l r_l. \end{aligned}$$

The objective function of the dual of problem (2) is then:

$$\begin{aligned} D(\vec{q}) &= \max_{0 \le x_s \le M_s, s=1,...,S, \vec{r} \in \mathrm{Co}(\mathcal{R})} L(\vec{x}, \vec{r}, \vec{q}) \\ &= \sum_{s=1}^{S} B_s(\vec{q}) + V(\vec{q}), \end{aligned}$$

where

$$B_s(\vec{q}) = \max_{0 \le x_s M_s} \left[U_s(x_s) - \sum_{l=1}^{L} H_s^l q^l x_s \right], \tag{6}$$

and

$$V(\vec{q}) = \max_{\vec{r} \in \mathrm{Co}(\mathcal{R})} \sum_{l=1}^{L} q^l r_l. \tag{7}$$

Further, because the objective function in (7) is a linear function of $\vec{r}$, the optimal point must lie in the set $\mathcal{R}$, i.e.,

$$V(\vec{q}) = \max_{\vec{r}\in\mathcal{R}} \sum_{l=1}^{L} q^l r_l = \max_{\vec{r}=u(\vec{P}),\vec{P}\in\Pi} \sum_{l=1}^{L} q^l r_l. \tag{8}$$

The dual approach thus results in an elegant decomposition of the original problem. Given the Lagrange multipliers q^l, the rate control problem $B_s(\vec{q})$ and the scheduling problem $V(\vec{q})$ are decomposed. The Lagrange multiplier q^l can be interpreted as the *implicit cost* at link l. Each user s solves its own utility maximization problem $B_s(\vec{q})$ independently as if the "price" for user s is $\sum_{l=1}^{L} H_s^l q^l$. The scheduling problem $V(\vec{q})$ also computes the power assignment $\vec{P}$ and the schedule $\vec{r} = u(\vec{P})$ based on the implicit costs. Note that $V(\vec{q})$ also appears as the optimal scheduling policy in [3, 6].

The dual problem of (2) is then

$$\min_{\vec{q}\geq 0} D(\vec{q}). \tag{9}$$

The dual objective function $D(\vec{q})$ is convex. We can show that its subgradient is given by,

$$\frac{\partial D}{\partial q^l} = -\left(\sum_{s=1}^{S} H_s^l x_s - r_l\right).$$

where $\vec{x} = [x_s]$ and $\vec{r} = [r_l]$ solve (6) and (8), respectively. We can then use the subgradient method to solve the dual problem [16]. The solution to the optimal cross-layered rate control problem can be summarized as follows:

The Optimal Cross-Layered Rate Control Algorithm:

At each iteration t:

- The data rates of the users are determined by

$$x_s(t) = \operatorname*{argmax}_{0\leq x_s\leq M_s} \left[U_s(x_s) - \sum_{l=1}^{L} H_s^l q^l(t) x_s\right]. \tag{10}$$

- The schedule is determined by

$$\vec{r}(t) = \operatorname*{argmax}_{\vec{r}\in\mathcal{R}} \sum_{l=1}^{L} q^l(t) r_l = \operatorname*{argmax}_{\vec{r}=u(\vec{P}),\vec{P}\in\Pi} \sum_{l=1}^{L} q^l(t) r_l. \tag{11}$$

- The implicit costs (i.e., Lagrange multipliers) are updated by

$$q^l(t+1) = \left[q^l(t) + \alpha_l \left(\sum_{s=1}^{S} H_s^l x_s(t) - r_l(t) \right) \right]^+ . \qquad (12)$$

The following proposition is given in [8].

Proposition 1 *a) There is no duality gap, i.e., the minimal value of (9) coincides with the optimal value of (2).*

b) Let Φ be the set of $\vec{q}$ that minimizes $D(\vec{q})$. For any $\vec{q} \in \Phi$, let $\vec{x}$ solve (10), then $\vec{x}$ is the unique optimal solution $\vec{x}^*$ of (2).

c) Assume that $\alpha_l = h\alpha_l^0$. Let $||\vec{q}||_A = \sum_{l=1}^{L} \frac{(q^l)^2}{\alpha_l^0}$ and $d(\vec{q}, \Phi) = \min_{\vec{p} \in \Phi} \sqrt{||\vec{q} - \vec{p}||_A}$. For any $\epsilon > 0$, there exists some $h_0 > 0$ such that, for any $h \leq h_0$ and any initial implicit costs $\vec{q}(0)$, there exists a time T_0 such that for all $t \geq T_0$,

$$d(\vec{q}(t), \Phi) < \epsilon \text{ and } ||\vec{x}(t) - \vec{x}^*|| < \epsilon.$$

Proposition 1 is a consequence of Theorem 2.3 in [16, p26]. The details of the proof is in Appendix A. It shows that, when the stepsizes α_l are small, the user rates $\vec{x}(t)$ will converge within a small neighborhood‡ of the optimal rate allocation $\vec{x}^*$.

The optimal cross-layered rate control algorithm (10)-(12) not only computes the optimal rate allocation, but also generates the stabilizing scheduling policy by solving (11) at each time slot t. In fact, let Q^l denote the queue size at link l. Then Q^l evolves approximately as§:

$$Q^l(t+1) \approx \left[Q^l(t) + \left(\sum_{s=1}^{S} H_s^l x_s(t) - r_l(t) \right) \right]^+ . \qquad (13)$$

Comparing (13) with (12), we can see that $Q^l(t) \approx q^l(t)/\alpha_l$. From here we can infer that $Q^l(t)$ is bounded.

‡An alternate formulation of Proposition 1 is as follows: if the stepsizes are time varying and they are chosen such that $\alpha_l(t) = h_t\alpha_l^0$, $h_t \to 0$ as $t \to \infty$ and $\sum_{t=1}^{+\infty} h_t = +\infty$, then $d(\vec{q}(t), \Phi) \to 0$ and $\vec{x}(t) \to \vec{x}^*$ as $t \to \infty$.

§Note, (13) is an approximation because not all links are active at the same time. Hence, data injected to the network by each user at time t may take several time slots to reach downstream links.

Proposition 2 *If the stepsizes α_l are sufficiently small, then using the schedules determined by solving (11) at each time slot, we have,*

$$\sup_t Q^l(t) < +\infty \text{ for all } l \in \mathcal{L}.$$

We give the proof in Appendix B. Combining Propositions 1 and 2, we conclude that, by choosing the stepsizes α_l sufficiently small, we can obtain user rate allocation $\vec{x}$ as close to $\vec{x}^*$ as we want, and we can obtain the joint stabilizing scheduling policy at the same time.

Remark: The duality approach that we used here (and in [8]) shares some similarities to the approach in [1, 9, 10]. However, there are also some major differences. The network models in [1] and [10] assume a restrictive set of rate-power functions. They either assume that the data rate at each link is a concave function of its *own* power assignment, or assume a special form of rate-power functions that are concave after a change of variables. In this paper, we impose no such restrictions. Further, a consequence of the assumption in [10] is that, at their optimal solution, all links will be transmitting at the same time. In the more general network model of this paper, it usually requires different sets of links to transmit at different time in order to achieve optimality. In [9], the authors propose a column generation approach for solving (2). This approach appears to be more suitable for *offline* computation as it requires solving a sequence of approximate problems to (2), each of which requires an iterative solution by itself. In contrast, in this paper we are more interested in solutions suitable for *on-line* implementation. Finally, these previous works have not addressed the joint stabilizing scheduling policy as we did in Proposition 2.

4 The Impact of Imperfect Scheduling on Cross-Layered Rate Control: The Static Case

In this paper, we are interested in developing cross-layered rate control solutions that are suitable for online implementation. The main difficulty in implementing the optimal solution of Section 3 is the complexity of the scheduling component. Depending on the rate-power function $u(\cdot)$, the scheduling problem (11) is usually a difficult global optimization problem. In some cases, this optimization problem does not even have a polynomial-time solution. Hence, solving (11) exactly at every time slot is too time-consuming.

As discussed in the Introduction, in this paper, we take a different approach from that of finding *optimal* rate allocations. We will only compute suboptimal solutions to the scheduling problem (11), which we will refer to as *imperfect schedules*. We will instead study how imperfect scheduling impacts the optimality of cross-layered rate control. Our objective is to find some imperfect scheduling policies that are easy to implement and that, when properly designed with rate control, result in good overall performance.

We will particularly be interested in the following class of imperfect scheduling policies:

Imperfect Scheduling Policy S_γ:

Fix $\gamma \in (0, 1]$. At each time slot t, compute a schedule $\vec{r}(t) \in \mathcal{R}$ that satisfies:

$$\sum_{l=1}^{L} r_l(t) q^l(t) \geq \gamma \max_{\vec{r} \in \mathcal{R}} \sum_{l=1}^{L} r_l q^l(t). \tag{14}$$

With an imperfect scheduling policy S_γ, the dynamics of cross-layered rate control are summarized by the following set of equations:

$$x_s(t) = \operatorname*{argmax}_{0 \leq x_s \leq M_s} \left[U_s(x_s) - \sum_{l=1}^{L} H_s^l q^l(t) x_s \right], \tag{15}$$

$$\vec{r}(t)^T \vec{q}(t) \geq \gamma \max_{\vec{r} \in \mathcal{R}} \vec{r}^{\,T} \vec{q}(t), \quad \vec{r}(t) \in \mathrm{Co}(\mathcal{R}), \tag{16}$$

$$q^l(t+1) = \left[q^l(t) + \alpha_l \left(\sum_{s=1}^{S} H_s^l x_s(t) - r_l(t) \right) \right]^+. \tag{17}$$

The parameter γ in (14) can be viewed as a tuning parameter indicating the degree of precision of the imperfect schedule. The complexity of finding a schedule $\vec{r}(t)$ satisfying (14) usually decreases as γ is reduced. When $\gamma = 1$, the dynamics (15)-(17) reduce to the case with perfect scheduling (as in Section 3). Let $\vec{x}^{\,*,0}$ denote the solution to the original optimal cross-layered rate control problem (2). The solution to the following problem turns out to be a good reference point for studying the dynamics (15)-(17) when $\gamma < 1$:

The γ-Reduced Problem:

$$\begin{aligned} \max_{0 \le x_s \le M_s} \quad & \sum_{s=1}^{S} U_s(x_s) \\ \text{subject to} \quad & \vec{x} \in \gamma\Lambda. \end{aligned} \tag{18}$$

Let $\vec{x}^{\,*,\gamma}$ denote the solution to the γ-reduced problem. The choice of $\gamma\Lambda$ in the constraint of the γ-reduced problem is motivated by the following proposition, which shows that an imperfect scheduling policy S_γ at most reduces the capacity region by a factor of γ. The proof is given in Appendix C.

Proposition 3 *If the user rates $\vec{x}$ lie strictly inside $\gamma\Lambda$, then any imperfect scheduling policy S_γ can stabilize the system.*

Motivated by Proposition 3, we would expect that the rate allocation computed by the dynamics (15)-(17) will be "no worse than" $\vec{x}^{\,*,\gamma}$. However, this assertion is not quite true. As we will see soon, the interaction between cross-layered rate control and imperfect scheduling is much more complicated. As the data rates of the users are reacting to the same implicit costs as the scheduling component is, there is a possibility that the system gets stuck into local sub-optimal areas. We will construct examples where, for a subset of the users, their data rates determined by the dynamics (15)-(17) can be much smaller than the corresponding rate allocation computed by the γ-reduced problem. Nonetheless, we will be able to show certain weak but desirable results on the fairness and convergence properties of cross-layer rate control with imperfect scheduling.

4.1 Dominance

We begin our analysis by studying whether the rate allocation computed by the dynamics (15)-(17) will dominate $\vec{x}^{\,*,\gamma}$. (Note: a vector $[x_1, ..., x_S]$ dominates another vector $[y_1, ..., y_S]$ if $x_i \ge y_i$ for all $i = 1, ..., S$.) It is easy to check that, if we let

$$\vec{r}(t) = \gamma \vec{r}_0(t), \tag{19}$$

the dynamics (15)-(17) will solve the γ-reduced problem. Hence, we can use (19) as a special case of the imperfect scheduling policy S_γ, and study first whether the rate allocation $\vec{x}^{\,*,0}$ of the original problem (2) dominates that of the γ-reduced problem (18). The following proposition shows that such dominance holds if the utility function is logarithmic. (Recall that logarithmic utility functions are of the form

$$U_s(x_s) = w_s \log x_s \text{ for all user } s,$$

where w_s is the weight for user s. In this case, the rate allocation computed by the original problem (2) is *weighted proportionally fair* [15].)

Proposition 4 *Assume that the utility function is logarithmic. Let $\vec{x}^{\,*,0}$ be the solution to the original problem (2). Then the solution to the γ-reduced problem is*

$$\vec{x}^{\,*,\gamma} = \gamma \vec{x}^{\,*,0}.$$

Proof: In the γ-reduced problem (18), do a change of variables $\vec{x}' = \vec{x}/\gamma$. Using the fact that

$$U_s(x_s) = w_s \log x'_s + w_s \log \gamma,$$

one can show that the γ-reduced problem becomes equivalent to the original problem (2). Hence, $\vec{x}^{\,*,\gamma} = \gamma \vec{x}^{\,*,0}$. *Q.E.D.*

However, as shown in the following example, if the utility function is not logarithmic, dominance will not hold in general.

Example 1: Consider the following wireline network (note that a wireline network can be viewed as a special case of our network model where the capacity of each link is fixed). There are two links, whose capacities are 2 and 7, respectively. There are three users. The first user uses both links, the second user uses only the first link, and the third user uses only the second link. Their utility functions are

$$U_1(x) = \log x + 6x$$

$$U_2(x) = \log x$$

$$U_3(x) = 36 \log x.$$

The γ-reduced problem is then

$$\begin{aligned} \max_{x_1, x_2, x_3 \geq 0} \quad & (\log x_1 + 6x_1) + \log x_2 + 36 \log x_3 \\ \text{subject to} \quad & x_1 + x_2 \leq 2\gamma \\ & x_1 + x_3 \leq 7\gamma. \end{aligned}$$

When $\gamma = 1$, the solution is $\vec{x}^{\,*,0} = [1 \quad 1 \quad 6]^T$. When $\gamma = 0.95$, the solution becomes $\vec{x}^{\,*,\gamma} = [0.8551 \quad 1.0449 \quad 5.7949]^T$. Note that the rate of the second user *increases* as γ is reduced. This example shows that $\vec{x}^{\,*,0}$ does not dominate $\vec{x}^{\,*,\gamma}$ in general.

4.2 A Weak Fairness Property

For the rest of the paper, we will focus on logarithmic utility functions, although most of the results that follow can also be extended to utility functions of other forms (as in (5)). Note that even though $\vec{x}^{\,*,0}$ dominates $\vec{x}^{\,*,\gamma}$ when the utility function is logarithmic (as shown in Proposition 4), it does not imply that the rate allocation computed by the cross-layered rate control algorithm with an *arbitrary* imperfect scheduling policy S_γ will dominate $\vec{x}^{\,*,\gamma}$.

In the following proposition, we characterize the likely rate allocation under imperfect scheduling *provided that the dynamics (15)-(17) converges.* The proof is given in Appendix D.

Proposition 5 *Assume that the utility function is logarithmic (i.e., of the form in (4)). If the dynamics (15)-(17) converges, i.e., $\vec{x}(t) \to \vec{x}^{\,*,I}$ and $\vec{q}(t) \to \vec{q}_I^{\,*}$ as $t \to \infty$, then*

$$\vec{x}^{\,*,I} \in \Lambda \text{ and } \sum_{s=1}^{S} \frac{w_s x_s^{*,\gamma}}{x_s^{*,I}} \leq \sum_{s=1}^{S} w_s. \tag{20}$$

Proposition 5 can be generalized to other forms of utility functions (as in (5)). This result can be viewed as a *weak fairness property*. It shows that, if the dynamics (15)-(17) converge, the

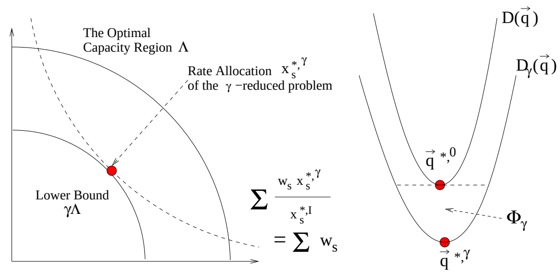

Figure 1: The weak fairness property (left) and the set Φ_γ (right).

rate allocation of the users will lie in a strip defined by (20) (see Fig. 1). Hence, the rate of each user is unlikely to be too unfair compared to $\vec{x}^{*,\gamma}$. In particular, if $w_s = 1$ for all s, then by (20), $x_s^{*,I}$ will be no smaller than $x_s^{*,\gamma}/S$. On the negative side, the rates of some users can still be substantially smaller than their rates computed by the γ-reduced problem, which indicates that cross-layered rate control with imperfect scheduling may indeed get stuck into local sub-optimal regions.

4.3 Convergence

We next study the question whether the dynamics (15)-(17) converge in the first place. Using a duality approach analogous to that in Section 3, we can define the dual of the γ-reduced problem as

$$D_\gamma(\vec{q}) = \sum_{s=1}^{S} B_s(\vec{q}) + \gamma V(\vec{q}),$$

where $B_s(\vec{q})$ and $V(\vec{q})$ are still defined as in (6) and (7), respectively. Note that both $D(\vec{q})$ and $D_\gamma(\vec{q})$ are convex functions and $D(\vec{q}) \geq D_\gamma(\vec{q})$.

Let $\vec{q}^{*,0}$ denote a minimizer of $D(\vec{q})$ and $\vec{q}^{*,\gamma}$ denote a minimizer of $D_\gamma(\vec{q})$. Further, let

$$\Phi_\gamma = \{\vec{q} : D_\gamma(\vec{q}) \leq D(\vec{q}^{*,0})\}.$$

Proposition 6 *Assume that* $\alpha_l = h\alpha_l^0$. *Let* $||\vec{q}||_A = \sum_{l=1}^{L} \frac{(q^l)^2}{\alpha_l^0}$. *For any* $\epsilon > 0$, *there exists some* $h_0 > 0$ *such that, for any* $h \leq h_0$ *and any initial implicit costs* $\vec{q}(0)$, *there exists a time* T_0 *such* *that for all* $t \geq T_0$,

$$\sqrt{||\vec{q}(t) - \vec{q}^{\,*,0}||_A} < \max_{\vec{p} \in \Phi_\gamma} \sqrt{||\vec{p} - \vec{q}^{\,*,0}||_A} + \epsilon.$$

The proof is provided in Appendix E. Proposition 6 shows that, if the stepsizes α_l are sufficiently small, the dynamics (15)-(17) will eventually enter a neighborhood of the set Φ_γ. Note that both $\vec{q}^{\,*,0}$ and $\vec{q}^{\,*,\gamma}$ belong to the set Φ_γ (see Fig. 1). Hence, in a weak sense, the dynamics of the system are moving in the right direction. However, in general the set Φ_γ is quite large and does not provide much further insights on the eventual rate allocation. We next present two examples illustrating the possible behaviors of the dynamics.

Example 2:

We will first show that, for any vectors $\vec{q}_I^{\,*}$ and $\vec{x}^{\,*,I}$ that satisfy

$$x_s^{*,I} = \frac{w_s}{\sum_{l=1}^{L} H_s^l q_I^{l,*}} \text{ for all } s,\ \vec{x}^{\,*,I} \in \Lambda, \text{ and}$$
$$\sum_{l=1}^{L} q_I^{l,*} \sum_{s=1}^{S} H_s^l x_s^{*,I} > \gamma \max_{\vec{r} \in \mathrm{Co}(\mathcal{R})} \sum_{l=1}^{L} q_I^{l,*} r_l, \tag{21}$$

there exists an imperfect scheduling policy S_γ such that the dynamics (15)-(17) converge to $\vec{q}_I^{\,*}$ and $\vec{x}^{\,*,I}$. Note that the above set of conditions implies (20). In fact, since

$$\left[\sum_{s=1}^{S} H_s^l x_s^{*,\gamma}, l \in \mathcal{L}\right] \in \gamma \mathrm{Co}(\mathcal{R}),$$

we have,

$$\begin{aligned}
\sum_{s=1}^{S} w_s &= \sum_{s=1}^{S} x_s^{*,I} \sum_{l=1}^{L} H_s^l q_I^{l,*} = \sum_{l=1}^{L} q_I^{l,*} \sum_{s=1}^{S} H_s^l x_s^{*,I} \\
&\geq \sum_{l=1}^{L} q_I^{l,*} \sum_{s=1}^{S} H_s^l x_s^{*,\gamma} = \sum_{s=1}^{S} x_s^{*,\gamma} \sum_{l=1}^{L} H_s^l q_I^{l,*} \\
&= \sum_{s=1}^{S} \frac{w_s x_s^{*,\gamma}}{x_s^{*,I}}.
\end{aligned}$$

We now show how a suitable imperfect scheduling policy S_γ can be constructed. It is easy to verify that $\vec{x}^{\,*,I}$ is the solution to the following optimization problem and $\vec{q}_I^{\,*}$ is the corresponding

Lagrange multipliers.

$$\max_{\vec{x}\geq 0} \quad \sum_{s=1}^{S} w_s \log x_s$$
$$\text{subject to} \quad \sum_{s=1}^{S} H_s^l x_s \leq \sum_{s=1}^{S} H_s^l x_s^{*,I}. \tag{22}$$

Hence, if we let

$$r_l(t) = \sum_{s=1}^{S} H_s^l x_s^{*,I} \text{ for all } l \text{ and all } t, \tag{23}$$

then, using a standard gradient descent argument for the dual problem of (22), we can show that the dynamics (15)-(17) will converge to $\vec{q}_I^*$ and $\vec{x}^{\,*,I}$ as $t \to \infty$. It remains to be verified whether the schedule in (23) belongs to the class of imperfect scheduling policies S_γ. To see this, note that if we pick the initial implicit cost vector $\vec{q}(0)$ to be sufficiently close to $\vec{q}_I^*$, then $\vec{q}(t) \approx \vec{q}_I^*$ for all t. Hence,

$$\sum_{l=1}^{L} q^l(t) r_l(t) \approx \sum_{l=1}^{L} q_I^{l,*} \sum_{s=1}^{S} H_s^l x_s^{*,I}$$
$$> \gamma \max_{\vec{r}\in \mathrm{Co}(\mathcal{R})} \sum_{l=1}^{L} q_I^{l,*} r_l \approx \gamma \max_{\vec{r}\in \mathrm{Co}(\mathcal{R})} \sum_{l=1}^{L} q^l(t) r_l,$$

i.e., the schedule in (23) indeed belongs to S_γ if the initial implicit cost vector $\vec{q}(0)$ is sufficiently close to $\vec{q}_I^*$.

Example 3:

We next give another example in which the dynamics (15)-(17) *never converge to any point*. Consider the following simple wireline network with two users, each of which uses one different link. The capacity is c for both links. The solution to the γ-reduced problem is simply $x_1^{*,\gamma} = x_2^{*,\gamma} = \gamma c$. Assume that the vectors $\vec{q}_I^*$ and $\vec{x}^{\,*,I}$ satisfy the conditions in (21) of Example 2. At any time t, define

$$\Delta(t) = \vec{q}(t) - \vec{q}_I^*.$$

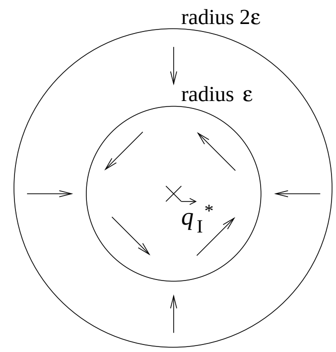

Figure 2: The direction of the update of the implicit costs

Let ϵ be a small positive number. We now use the following scheduling policy:

$$\vec{r}(t) = [r_1 \quad r_2]^T = \begin{cases} [x_1 \quad x_2]^T - \epsilon \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \frac{\Delta(t)}{||\Delta(t)||}, & \text{if } \Delta(t) \leq \epsilon, \\ [x_1 \quad x_2]^T - \epsilon \frac{\Delta(t)}{||\Delta(t)||}, & \text{if } \epsilon \leq \Delta(t) \leq 2\epsilon, \\ c, & \text{otherwise.} \end{cases}$$

With this choice of the schedule $\vec{r}(t)$, the update of the implicit cost $\vec{q}(t)$ will be around a circle when $||\vec{q}(t) - \vec{q}_I^*|| \leq \epsilon$, and it will be towards $\vec{q}_I^*$ when $\epsilon < ||\vec{q}(t) - \vec{q}_I^*|| \leq 2\epsilon$ (see Fig. 2). Provided that the initial $\vec{q}(0)$ satisfies $||\vec{q}(0) - \vec{q}_I^*|| \leq 2\epsilon$ and the stepsizes are sufficiently small, the dynamics (15)-(17) will eventually follow the circle $||\vec{q}(t) - \vec{q}_I^*|| = \epsilon$, and hence will never converge. We can verify as in Example 2 that the schedule $\vec{r}(t)$ does belong to the class S_γ when the stepsizes and ϵ are sufficiently small.

To conclude this section, we have studied the impact of imperfect scheduling on the dynamics of cross-layered rate control *when the number of users in the system is fixed*. We have presented several examples that illustrate the difficulty in characterizing the dynamics precisely. We have shown that the system may not even converge in the first place, or, it may converge to any rate allocation within a fairly large set that does not possess any desirable dominance property. These examples indicate that the interaction between cross-layered rate control and imperfect scheduling are quite complicated, and the system may indeed get stuck into local sub-optimal regions. Nonetheless, we do show two desirable, but weak, results on the fairness and convergence properties of the system. In Proposition 6, we are able to show that the dynamics (15)-(17) appear to move in the right direction globally. In Proposition 5, we show that those local sub-optimal regions are probably "not too bad." In the next section, we will turn to the case when users dynamically arrive and depart the network, and surprisingly, we will be able to show far stronger results on the performance of the system there.

5 Stability Region of Cross-Layered Rate Control

In this section, we turn to the case when the number of users in the system is itself a stochastic process. We will study how imperfect scheduling impacts the *stability region* of the system employing cross-layer rate control. Here, by *stability*, we mean that the number of users in the system and the queue lengths at all links in the network remain finite. The *stability region* of the system is the set of offered loads under which the system is stable. Previous works for wireline networks have shown that, by allocating data rates to the users according to some fairness criteria, the *largest possible* stability region can be achieved [15]. This result is important as it tells us that fairness is not just an *aesthetic* property, but it actually has a strong global *performance* implication, i.e., in achieving the *largest possible* stability region. In this section, we will show that similar but stronger results can be shown for our cross-layered rate control scheme with imperfect scheduling.

To be precise, instead of using the notation s for user s, we now use s to denote a class of users with the same utility function and the same path. We assume that users of class s arrive according to a Poisson process with rate λ_s and that each user brings with it a file for transfer whose size is exponentially distributed with mean $1/\mu_s$. The load brought by users of class s is then $\rho_s = \lambda_s/\mu_s$. Let $\vec{\rho} = [\rho_1, ..., \rho_S]$. Let $n_s(t)$ denote the number of users of class s that are in the system at time t, and let $\vec{n}(t) = [n_1(t), ..., n_S(t)]$. We assume that the rate allocations for users of the same class are identical. Let $x_s(t)$ denote the rate of each user of class s at time t.

In the rate assignment model that follows, the evolution of $\vec{n}(t)$ will be governed by a Markov process. Its transition rates are given by:

$$\begin{aligned} n_s(t) &\to n_s(t)+1, && \text{with rate } \lambda_s, \\ n_s(t) &\to n_s(t)-1, && \text{with rate } \mu_s x_s(t) n_s(t) \\ & && \quad \text{if } n_s(t) > 0. \end{aligned}$$

As in [17], We say that the above system is *stable* if

$$\limsup_{t\to\infty} \frac{1}{t}\int_0^t \mathbf{1}_{\{\sum_{s=1}^{S} n_s(t)+\sum_{l=1}^{L} q^l(t)>M\}} dt \to 0,$$

as $M \to \infty$. This means that the fraction of time that the amount of "unfinished work" in the system exceeds a certain level M can be made arbitrary small as $M \to \infty$. The *stability region* Θ of the system under a given rate control and scheduling policy is the set of offered loads $\vec{\rho}$ such that the system is stable.

We next describe the rate assignment and implicit cost update policy. We assume that time is divided into slots of length T, and the schedules and implicit costs are only updated at the end of each time slot. However, users may arrive and depart in the middle of a time slot. Let $\vec{q}(kT)$ denote the implicit cost at time slot k. The data rates of the users are determined by the current implicit costs as in (10). For simplicity, we assume that the utility function is logarithmic (the result can be readily generalized to utility functions of other forms in (5)). Further, let M_s denote the maximum data rate for users of class s. The rate of each user of class s is then given by

$$x_s(t) = x_s(kT) = \min\left\{ \frac{w_s}{\sum_{l=1}^{L} H_s^l q^l(kT)}, M_s \right\} \tag{24}$$

for $kT \le t < (k+1)T$. The schedule $\vec{r}(kT)$ at time slot k is computed according to an imperfect scheduling policy S_γ based on the current implicit cost $\vec{q}(kT)$. Finally, at the end of each time slot, the implicit costs are updated as

$$q^l((k+1)T) = \left[q^l(kT) + \alpha_l \left(\sum_{s=1}^{S} H_s^l \int_{kT}^{(k+1)T} n_s(t) x_s(kT) dt - r_l(kT)T \right) \right]^+. \tag{25}$$

The following proposition shows that, using the above cross-layered rate control algorithm with imperfect scheduling policy S_γ, the stability region of the system is no smaller than $\gamma\Lambda$.

Proposition 7 *If*

$$\max_{l \in \mathcal{L}} \alpha_l \leq \frac{1}{T\bar{S}\bar{L}} \min_s \frac{w_s}{4\rho_s M_s}, \tag{26}$$

where $\bar{S} = \max_{l \in \mathcal{L}} \sum_{s=1}^{S} H_s^l$ is the maximum number of classes using any link, and $\bar{L} = \max_s \sum_{l=1}^{L} H_s^l$ is the maximum number of links used by any class, then for any offered load $\vec{\rho}$ that resides strictly inside $\gamma\Lambda$, the system described by the Markov process $[\vec{n}(kT), \vec{q}(kT)]$ is stable.

Several remarks are in order: Firstly, Proposition 7 shows that, when imperfect schedules are used, the stability region of the system employing cross-layer rate control is no worse than the capacity region shown in Proposition 3 (and used by the γ-reduced problem). This result is interesting (and somewhat surprising) given the fact that, when the number of users in the system is fixed, the dynamics of cross-layered rate control with imperfect scheduling can form loops or get stuck into local sub-optimal regions. *Nonetheless, Proposition 7 shows that these potential local sub-optimums are inconsequential when the arrivals and departures of the users are taken into account.*

Secondly, we do not need the rates of any users to converge. Previous results on the stability region of rate control typically adopt a *time-scale separation assumption* [15], which assumes that the rate allocation $\vec{x}(t)$ *perfectly solves (2) at each time instant* t. Such an approach is of little value for the model in this paper because the dynamics (15)-(17) with imperfect scheduling do not even converge in the first place! Further, the time-scale separation assumption is rarely realistic in practice: as the number of users in the system is constantly changing, the rate allocation may never have the time to converge. In Proposition 7, we establish the stability region of the system without requiring such a time-scale separation assumption. This result is of independent value. For the special case when $\gamma = 1$, it can be viewed as a stronger version of previous results in the literature (including those for wireline networks, e.g., Theorem 1 in [15]).

Finally, a simple stepsize rule is provided in (26). Note that when the number of users in the system is fixed, we typically require the stepsizes to be driven to zero for convergence to occur (see Proposition 1). However, in (26) the stepsizes can be chosen bounded away from zero. In fact, as the set $\gamma\Lambda$ is bounded, the stepsizes can be chosen independently of the offered load. The simplicity in the stepsize rule is another benefit we obtain by studying the dynamic arrivals and departures of the users.

5.1 The Main Idea of the Proof of Proposition 7

We now sketch the main idea of the proof for Proposition 7 so that the reader can gain some insight on the dynamics of the system. Define the following Lyapunov function,

$$\mathcal{V}(\vec{n},\vec{q}) = V_n(\vec{n}) + V_q(\vec{q}),$$

where $V_n(\vec{n}) = \sum_{s=1}^{S} \frac{w_s n_s^2}{2\lambda_s}$, and $V_q(\vec{q}) = \sum_{l=1}^{L} \frac{(q^l)^2}{2\alpha_l}$. We shall show below that $\mathcal{V}(\vec{n},\vec{q})$ has a negative drift. As a crude first-order approximation, assume that users arrive and depart only at the end of each time slot. Thus, $n_s(t) = n_s(kT)$ during the k-th time slot. We can show that (see Appendix F for the details),

$$\begin{aligned} & \mathbf{E}[V_n(\vec{n}((k+1)T)) - V_n(\vec{n}(kT))|\vec{n}(kT),\vec{q}(kT)] \\ \leq \;& T\sum_{s=1}^{S}\left[\frac{w_s}{x_s(kT)}\right][\rho_s - n_s(kT)x_s(kT)] + E_1(k), \end{aligned}$$

where $E_1(k)$ is an error term that is roughly on the order of $|\rho_s - n_s(kT)x_s(kT)|$. Since the rate allocation is determined by (24), we have (ignoring the maximum data rate M_s),

$$\begin{aligned} & \mathbf{E}[V_n(\vec{n}((k+1)T)) - V_n(\vec{n}(kT))|\vec{n}(kT),\vec{q}(kT)] \\ \leq \;& T\sum_{s=1}^{S}\left[\sum_{l=1}^{L} H_s^l q^l(kT)\right][\rho_s - n_s(kT)x_s(kT)] \\ & +E_1(k). \end{aligned} \tag{27}$$

We can also show that

$$\mathbf{E}[V_q(\vec{q}((k+1)T) - V_q(\vec{q}(kT))|\vec{n}(kT),\vec{q}(kT)]$$

$$\leq \; T\sum_{l=1}^{L} q^l(kT)\left[\sum_{s=1}^{S} H_s^l n_s(kT)x_s(kT) - r_l(kT)\right] + E_2(k), \tag{28}$$

where $E_2(k)$ is an error term that is roughly on the order of $\left[\sum_{s=1}^{S} H_s^l n_s(kT)x_s(kT) - r_l(kT)\right]^2$. Hence, by adding (27) and (28), and by changing the order of the summation, we have

$$\begin{aligned}
&\mathbf{E}[\mathcal{V}(\vec{n}((k+1)T), \vec{q}((k+1)T)) \\
&\quad -\mathcal{V}(\vec{n}(kT), \vec{q}(kT))|\vec{n}(kT), \vec{q}(kT)] \\
\leq \; & T\sum_{l=1}^{L} q^l(kT)\left[\sum_{s=1}^{S} H_s^l \rho_s - r_l(kT)\right] \\
&\quad +E_1(k) + E_2(k).
\end{aligned} \tag{29}$$

By assumption, $\vec{\rho}$ lies strictly inside $\gamma\Lambda$. Hence, there exists some $\epsilon > 0$ such that

$$[(1+\epsilon)\sum_{s=1}^{S} H_s^l \rho_s] \in \gamma \mathrm{Co}(\mathcal{R}).$$

By the definition of the imperfect scheduling policy S_γ,

$$\sum_{l=1}^{L} q^l(kT) r_l(kT) \geq (1+\epsilon)\sum_{l=1}^{L} q^l(kT)\sum_{s=1}^{S} H_s^l \rho_s.$$

Substituting into (29), we have,

$$\begin{aligned}
&\mathbf{E}[\mathcal{V}(\vec{n}((k+1)T), \vec{q}((k+1)T)) \\
&\quad -\mathcal{V}(\vec{n}(kT), \vec{q}(kT))|\vec{n}(kT), \vec{q}(kT)] \\
\leq \; & -T\epsilon \sum_{l=1}^{L} q^l(kT)\sum_{s=1}^{S} H_s^l \rho_s + E_1(k) + E_2(k).
\end{aligned} \tag{30}$$

This shows that $\mathcal{V}(\cdot,\cdot)$ would drift towards zero when $||\vec{q}(kT)||$ is large and when the error terms $E_1(k)$ and $E_2(k)$ are bounded. We would then apply Theorem 2 of [17] to establish the stability of the system.

To complete the proof, however, we have to address several difficulties:

- In order to apply Theorem 2 of [17], a stronger negative drift is required. Instead of (30), we need,

$$\begin{aligned} &\mathbf{E}[\mathcal{V}(\vec{n}((k+1)T), \vec{q}((k+1)T)) \\ &\qquad -\mathcal{V}(\vec{n}(kT), \vec{q}(kT))|\vec{n}(kT), \vec{q}(kT)] \\ \leq\ & -\epsilon'(||\vec{n}(kT)|| + ||\vec{q}(kT)||) + E_0 \end{aligned}$$

  for some positive constants ϵ' and E_0.

- Further, in order to apply Theorem 2 of [17], the error terms $E_1(k)$ and $E_2(k)$ have to be bounded, which is not true in (30) since they both can become large as $n_s(kT)$ increases.

- Finally, users could arrive and depart at any time (not only at the end of a time slot).

The complete proof that addresses these difficulties is given in Appendix F.

We now give two examples showing how efficient cross-layer rate control schemes can be constructed by applying Proposition 7 to different network settings.

5.2 The Node Exclusive Interference Model

Proposition 7 is most useful when an imperfect schedule that satisfies (14) can be easily computed for some reasonable value of γ. This is the case under the following node exclusive interference model.

The Node Exclusive Interference Model:

- The data rate of each link is fixed at c_l.

- Each node can only send to or receive from one other node at any time.

This interference model has been used in earlier studies of rate control in multihop wireless networks [12, 13]. Under this model, the *perfect* schedule (according to (11)) at each time slot corresponds to the **Maximum Weighted Matching (MWM)**, where the weight of each link is $q^l c_l$. (A *matching* is a subset of the links such that no two links share the same node. The *weight* of a matching is the total weight over all links belonging to the matching. A *maximum-weighted-matching (MWM)* is the matching with the maximum weight.) An $O(N^3)$-complexity algorithm for MWM can be found in [18], where N is the number of nodes. On the other hand, the following much simpler **Greedy Maximal Matching (GMM)** algorithm can be used to compute an imperfect schedule with $\gamma = 1/2$. Start from an empty schedule. From all possible links $l \in \mathcal{L}$, pick the link with the largest $q^l c_l$. Add this link to the schedule. Remove all links that are incident with either the sending node or the receiving node of link l. Pick the link with the largest $q^l c_l$ from the *remaining* links, and add to the schedule. Continue until there are no links left. The GMM algorithm has only $O(L \log L)$-complexity (where L is the number of links), and is much easier to implement than MWM. Using the technique in Theorem 10 of [19], we can show that the weight of the schedule computed by the GMM algorithm is at least 1/2 of the weight of the *maximum-weighted-matching*. According to Proposition 7, the stability region will be at least $\Lambda/2$ using our cross-layered rate control scheme with the GMM scheduling policy.

For the node-exclusive interference model, a *layered approach* to rate control is also possible, which considers *separately* the dynamics of rate control and scheduling [12, 13]. It has been shown that the optimal capacity region Λ in the node-exclusive interference model is bounded by $\frac{2}{3}\Psi_0 \subseteq \Lambda \subseteq \Psi_0$, where

$$\Psi_0 = \left\{ \vec{x} \,\middle|\, \sum_{l:b(l)=i \text{ or } e(l)=i} \frac{1}{c_l} \sum_{s=1}^{S} H_s^l x_s \leq 1 \text{ for all } i \right\}.$$

and $b(l)$ and $e(l)$ are the sending node and the receiving node, respectively, of link l. The layered approach then chooses the lower bound $\frac{2}{3}\Psi_0$ as the *rate region* for computing the rate allocation [12, 13]. On the other hand, when an imperfect GMM scheduling policy is used, the capacity region can be reduced by half in the worst case (according to Proposition 3). Hence, the layered approach then needs to use $\Psi_0/3 (\subseteq \Lambda/2)$ as the *rate region.* Note that for the layered approach with GMM scheduling, $\Psi_0/3$ is an *upper bound* for its stability region, which is smaller than the *lower bound* of the stability region of the corresponding cross-layered approach (which is $\Lambda/2$

according to Proposition 7). Hence, due to its *conservative* nature, the layered approach always suffers from *worst case* inefficiencies. In Section 7, we will use simulations to show that our cross-layered rate control scheme can in practice substantially outperform the layered approach.

5.3 General Interference Models

Under general interference models, it may still be time-consuming to compute a schedule that satisfies (14) for a given value of γ. We now use Proposition 7 to develop a scheduling policy that can cut down the *frequency* of such computation, and hence effectively reduce the computation overhead. This idea is motivated by the observation that implicit costs, being updated by (17), cannot change abruptly. Hence, there is a high chance that a schedule computed earlier can be *reused* in subsequent time-slots. To see this, assume that we know a schedule $\vec{r}^{\,0}$ that satisfies (14) for an inefficiency factor $\gamma_0 > \gamma$ when the implicit cost vector is $\vec{q}^{\,0}$, i.e.,

$$\sum_{l=1}^{L} r_l^0 q_0^l \geq \gamma_0 \max_{\vec{r}\in\mathcal{R}} \sum_{l=1}^{L} r_l q_0^l. \tag{31}$$

Let the implicit cost vector at the current time slot be $\vec{q}$, and let $\vec{r}^{\,*}$ denote the corresponding (but unknown) perfect schedule. We can normalize $\vec{q}^{\,0}$ and $\vec{q}$ to be of unit length since the corresponding schedules will remain the same. We have,

$$\begin{aligned} \sum_{l=1}^{L} q^l r_l^* &= \sum_{l=1}^{L} (q^l - q_0^l) r_l^* + \sum_{l=1}^{L} q_0^l r_l^* \\ &\leq \sum_{l=1}^{L} [q^l - q_0^l]^+ r_l^{\max} + \frac{\sum_{l=1}^{L} q_0^l r_l^0}{\gamma_0}, \end{aligned}$$

where $r_l^{\max}$ is the maximum rate of link l. Hence, if

$$\sum_{l=1}^{L} q^l r_l^0 \geq \gamma \left\{ \sum_{l=1}^{L} [q^l - q_0^l]^+ r_l^{\max} + \frac{\sum_{l=1}^{L} q_0^l r_l^0}{\gamma_0} \right\},$$

we can still use $\vec{r}^{\,0}$ as the imperfect schedule for $\vec{q}$. This approach is even more powerful when the network can remember multiple schedules from the past. Assume that the schedules $\vec{r}^{\,1}, \vec{r}^{\,2}, ..., \vec{r}^{\,K}$

correspond to $\vec{q}^{\,1}, \vec{q}^{\,2}, ..., \vec{q}^{\,K}$, respectively, and each pair satisfies (31). Then, as long as

$$\max_{k=1,..,K} \sum_{l=1}^{L} q^l r_l^k \qquad (32)$$

$$\geq \min_{k=1,...,K} \gamma \left\{ \sum_{l=1}^{L} [q^l - q_k^l]^+ r_l^{\max} + \frac{\sum_{l=1}^{L} q_k^l r_l^k}{\gamma_0} \right\},$$

we do not need to compute a new schedule. Instead, we can use the schedule that maximizes the left hand side of (32). By Proposition 7, the stability region of the system using the above scheduling policy is no smaller than $\gamma\Lambda$. In Section 7, we will use simulations to show that such a simple policy can perform very well in practice.

6 A Fully Distributed Cross-Layered Rate Control and Scheduling Algorithm

Proposition 7 opens a new avenue for studying cross-layer design for rate control in multihop wireless networks. Instead of restricting our attention to the rate allocation at each snapshot of the system (as we did in Section 4 where the results tend to be weaker), we can now study the entire time horizon by focusing on the stability region of such a cross-layer-designed system. Motivated by Proposition 7, we now present a *fully distributed* cross-layered rate control and scheduling algorithm for the node-exclusive interference model in Section 5.2. (In contrast, the GMM algorithm in Section 5.2 still requires centralized implementation.) This new algorithm can be shown to achieve a stability region no smaller than $\Lambda/2$.

The new algorithm uses **Maximal Matching (MM)** to compute the schedule at each time [20]. A *maximal matching* is a matching such that no more links can be added without violating the node-exclusive interference constraint. To be precise, let q_{ij} denote the implicit cost at link (i, j). (For convenience, in this section we will index a link by a node pair (i, j).) A maximal matching $\mathcal{M}$ is a subset of $\mathcal{L}$ such that $q_{ij} \geq 1$ for all $(i, j) \in \mathcal{M}$, and, for each $(i, j) \in \mathcal{L}$, one of the following holds:

$$
\begin{aligned}
& q_{ij} < 1, \text{ or} && (33)\\
& (i,k) \in \mathcal{M} \text{ for some link } (i,k) \in \mathcal{L}, \text{ or}\\
& (k,i) \in \mathcal{M} \text{ for some link } (k,i) \in \mathcal{L}, \text{ or}\\
& (j,h) \in \mathcal{M} \text{ for some link } (j,h) \in \mathcal{L}, \text{ or}\\
& (h,j) \in \mathcal{M} \text{ for some link } (h,j) \in \mathcal{L}.
\end{aligned}
$$

Note that a maximal matching can be computed in a distributed fashion as follows. When a link (i, j) is added to the matching, we say that both node i and node j are *matched*. For each node i, if it has already been matched, no further action is required. Otherwise, node i scans its neighboring nodes. If there exists a neighboring node j such that node j has not been matched, node i sends a matching request to node j. It is possible that a matching request conflicts with other matching requests. In this case, the nodes involved in the conflict can use some randomization and local coordination to pick any non-conflicting subset of the matching requests. For those nodes whose matching requests are declined, they can repeat the above procedure until every node in the network is either matched or has no neighbors that are not matched.

Let

$$Q_i = \sum_{j:(i,j)\in\mathcal{L}} q_{ij} + \sum_{j:(j,i)\in\mathcal{L}} q_{ji}$$

denote the total cost of the links that either start from, or end at node i. Our new cross-layered rate control and scheduling algorithm then proceeds as follows.

The Distributed Cross-Layered Rate Control Algorithm:

At each time slot $[kT, (k+1)T)$:

- A maximal matching $\mathcal{M}(kT)$ is computed based on the implicit costs $\vec{q}(kT)$.
- The data rate of each user of class s is determined by

$$x_s(t) = x_s(kT)$$

$$= \max\left\{\frac{w_s}{2\sum_{(i,j)\in\mathcal{L}} H_s^{ij}\frac{Q_i(kT)+Q_j(kT)}{c_{ij}}}, M_s\right\} \tag{34}$$

where c_{ij} is the capacity of link (i,j), and H_s^{ij} is defined as H_s^l, i.e., $H_s^{ij}=1$, if users of class s use link (i,j); and $H_s^{ij}=0$, otherwise.

- The implicit costs are updated by:

$$q_{ij}((k+1)T) = \left[q_{ij}(kT) + \alpha\left(\sum_{s=1}^{S} H_s^{ij}\int_{kT}^{(k+1)T}\frac{n_s(t)x_s(kT)}{c_{ij}}dt - T\mathbf{1}_{\{(i,j)\in\mathcal{M}(kT)\}}\right)\right]^+ . \tag{35}$$

This new cross-layered rate control and scheduling algorithm is similar to the algorithms of Section 4 and 5 in many aspects:

- A user reacts to congestion by reducing its data rate when the implicit costs along its path increase.
- The implicit cost at each link (i,j) is updated based on the difference between the offered load and the schedule of the link.

However, there is a critical difference. When the maximal matching is computed, we do not care about the precise value of the implicit costs (see (33)). Hence, the maximal matching typically does not satisfy the requirement of the imperfect scheduling policy S_γ, and Proposition 7 does not apply either. Further, the rate control part (34) is also different from that in the earlier sections. *It has been chosen specifically for the maximal matching scheduling policy.* Nonetheless, using similar techniques as in Section 5, we can show the following result on the stability region of the system. The details are given in Appendix G.

Proposition 8 *If the stepsize α is sufficiently small, then for any offered load $\vec{\rho}$ that resides strictly inside $\Lambda/2$, the system with the above distributed cross-layered rate control algorithm is stable.*

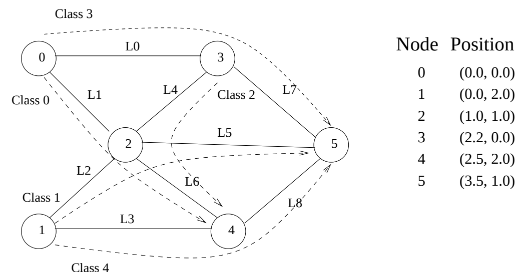

Figure 3: The Network Topology

7 Numerical Results

We now use simulations to verify the results in this paper. We use the network in Fig. 3. There are 5 classes of users, whose paths are shown in Fig. 3. Their utility functions are all given by $U_s(x_s) = \log x_s$. We first use the following interference model. The path loss $G(i,j)$ from a node i to a node j is given by $G(i,j) = d_{ij}^{-4}$ where d_{ij} is the distance from node i to node j (the positions of the nodes are also given in Fig. 3). We assume that the data rate r_{ij} at link $(i,j) \in \mathcal{L}$ is proportional to the SIR, i.e.,

$$r_{ij} = W \frac{G(i,j)P_{ij}}{N_0 + \sum_{(k,h)\in\mathcal{L},(k,h)\neq(i,j)} G(k,j)P_{k,h}},$$

where N_0 is the background noise and W is the bandwidth of the system. This assumption is suitable for CDMA systems with a moderate processing gain [6]. Each node i has a power constraint $P_{i,\max}$, i.e., the power allocation must satisfy $\sum_{j:(i,j)\in\mathcal{L}} P_{ij} \leq P_{i,\max}$ for all i.

We first simulate the case when there is one user for each class. The left figure in Fig. 4 shows the evolution of the data rates for all five users when the network computes the perfect schedule according to (11) at every time slot. We have chosen $W = 10$, $N_0 = 1.0$, $P_{i,\max} = 1.0$ for all node i and $\alpha_l = 0.1$ for all link l. Note that the scheduling subproblem (11) for this interference model is a complex non-convex global optimization problem. In [8], we have given an $O(2^N)$ algorithm for solving the perfect schedule, where N is the number of nodes. Executing such an algorithm at every time-slot is extremely time-consuming.

We then simulate the imperfect scheduling policy outlined in Section 5.3 for general interference

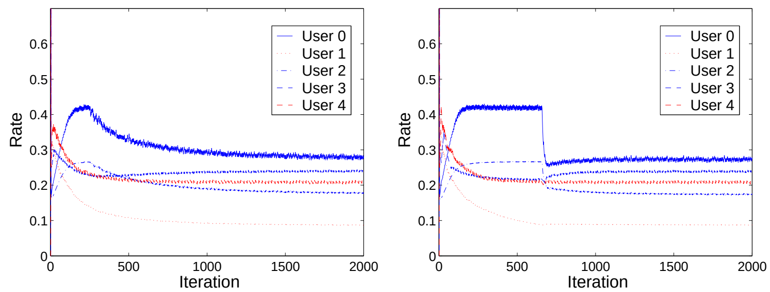

Figure 4: The evolution of the data rates for all users with perfect scheduling (left) and with imperfect scheduling (right, $\gamma = 0.5$).

models. Such an imperfect scheduling policy attempts to *reuse* schedules that have already been computed in the past. In our simulation, we have chosen $\gamma_0 = 1.0$ in (31), i.e., each of these past schedules are perfect schedules. The computational complexity could have been further reduced if we had chosen $\gamma_0 < 1$. However, we leave this for future work. Instead, in this paper we focus on how the imperfect scheduling policy can *reduce the number of times that new perfect schedules have to be computed*. The system that we simulate can store at most 10 past schedules. If there are already 10 past schedules and a new perfect schedule is computed, the new schedule will replace the old one that has the smallest weighted-sum $\sum_{l=1}^{L} q^l r_l$. In the right figure of Fig. 4, we show the evolution of the data rates when $\gamma = 0.5$. Note that the rate allocation eventually converges to values close to that with perfect scheduling. We also record the number of times that perfect schedules are computed. When $\gamma = 0.5$, perfect schedules are computed in only 7 iterations among the entire 2000 iterations of the simulation, and most of these perfect schedules are computed at the initial stage of the simulation. We have simulated other values of γ and find similar results. In fact, by just reducing γ from 1.0 to 0.9, the number of times that perfect schedules have to be computed is reduced to 34 (over 2000 iterations of simulation). These results indicate that our cross-layered rate control scheme with the imperfect scheduling policy in Section 5.3 can substantially reduce the computation overhead and still maintain good performance.

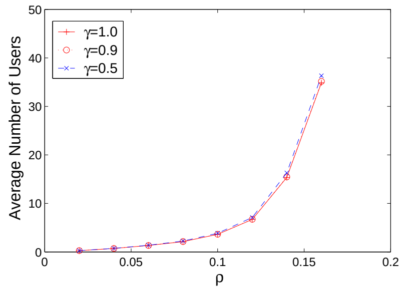

Figure 5: The average number of users in the system versus load.

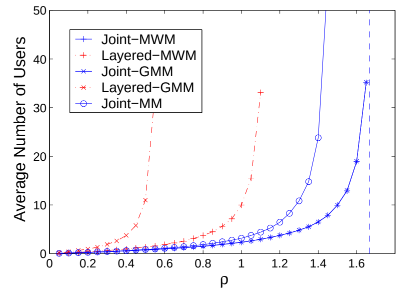

Figure 6: The average number of users in the system versus load: the node-exclusive interference model

We then simulate the case when there are dynamic arrivals and departures of the users as in Section 5. Users of each class arrive to the network according to a Poisson process with rate λ. Each user brings with it a file to transfer whose size is exponentially distributed with mean $1/\mu = 100$ unit. We vary the arrival rate λ (and hence the load $\rho = \lambda/\mu$) and record in Fig. 5 the average number of users in the system at any time for different choice of γ. Given γ, the average number of users in the system will increase to infinity as the offered load ρ approaches a certain limit. This limit can then be viewed as the capacity of the system. From Fig. 5, we observe that the capacity of the system is not significantly affected when γ is reduced from 1.0 to 0.5. On the other hand, the number of time-slots that new perfect schedules have to be computed is reduced to less than 1% of the total number of time-slots when $\gamma = 0.9$, and to less than 0.05% when $\gamma = 0.5$. These results confirm again the effectiveness of our cross-layered rate control scheme with the imperfect scheduling policy in Section 5.3, in reducing the computation overhead and achieving good overall performance.

We next turn to the node-exclusive interference model in Section 5.2, where we can draw a comparison with the layered approach to rate control [12, 13]. We still use the network topology in Fig. 3. The capacity of each link is now fixed at 10 units. Due to space constraints, we only report the result for the case when there are dynamic arrivals and departures of the users. Fig. 6 demonstrates the average number of users in the system versus load with different rate control and scheduling schemes. We label each curve with the rate control scheme (we use "Joint" to denote the cross-layered rate control scheme and use "Layered" to denote the layered approach in [13]), followed by the scheduling policy. (Note that the curve for the cross-layered rate control scheme with GMM scheduling, labeled as "Joint-GMM," in fact overlaps with the curve for the optimal cross-layered rate control scheme with perfect MWM scheduling, which is the right most curve labeled as "Joint-MWM.") From Fig. 6, we observe that, regardless of the scheduling policy used (either MWM, GMM, or MM), the layered approach always performs *much poorer* than the corresponding cross-layered approach. The performance gap widens even more when an imperfect scheduling policy (such as GMM) is used. In particular, the fully distributed joint rate control and scheduling algorithm in Section 6 (with *imperfect* maximal matching scheduling, labeled "Joint-MM"), actually performs even better than the layered approach with the *perfect* (and more complex) MWM scheduling (labeled "Layered-MWM"). These results demonstrate that the conservative nature of the layered approach indeed hurts the overall performance of the system, and an appropriately designed cross-layered rate control scheme can perform very well in practice even with imperfect scheduling.

8 Conclusion

In this paper, we study how the performance of cross-layer rate control will be impacted if the network can only use an imperfect (and potentially distributed) scheduling component. When the number of users in the system is fixed, we are able to show some desirable, but weak, results on the fairness and convergence properties of the system. We then turn to the case with dynamic arrivals and departures of the users, and establish stronger results bounding the stability region of the system. Compared with a layered approach that does not design rate control and scheduling together, the cross-layered approach has provably better performance bounds, and usually substantially outperforms the layered approach. Hence, the cross-layered approach is much more robust to imperfect scheduling than the layered approach. The insights drawn from our analyses also enable us to design a *fully distributed* and high-performance cross-layered rate control and scheduling algorithm for the node-exclusive interference model.

These results constitute an important step towards designing fully distributed cross-layered rate control schemes for multihop wireless networks. Several directions for future work are possible. For example, Proposition 7 may be combined with a clustering scheme to design distributed cross-layered rate-control solutions for large networks. We can also use similar techniques as in [8] to combine cross-layered rate control with multipath routing. Our main result (Proposition 7) can also be extended to the case with random fading. It would also be important to study the impact of feedback delays, and to extend our results to hybrid wireless-wireline networks.

Appendix

A Proof of Proposition 1

The proof of part a) is quite standard (see, for example, Theorem 3.2.8 in [21, p44]). In fact, let $\vec{x}^*$ denote the optimal solution to the primal problem (2), and let $\vec{r}^{\,*}$ denote the corresponding vector of link rates that satisfies (3). It is easy to verify that

$$\max_{0\le x_s\le M_s, \vec{r}\in \mathrm{Co}(\mathcal{R})} L(\vec{x},\vec{r},\vec{q}) \ge L(\vec{x}^*,\vec{r}^{\,*},\vec{q}) \ge \sum_{s=1}^{S} U_s(x_s^*) \text{ for all } \vec{q}\ge 0 \ .$$

To prove part a), we only need to find Lagrange multipliers $\vec{q}\ge 0$ such that

$$\max_{0\le x_s\le M_s, \vec{r}\in \mathrm{Co}(\mathcal{R})} L(\vec{x},\vec{r},\vec{q}) = \sum_{s=1}^{S} U_s(x_s^*).$$

Towards this end, let $\vec{b}=[b_l, l\in\mathcal{L}]$ and let

$$\begin{aligned} G(\vec{b}) = \max_{0\le x_s\le M_s} \quad & \sum_{s=1}^{S} U_s(x_s) \\ \text{subject to} \quad & \sum_{s=1}^{S} H_s^l x_s \le r_l + b_l \text{ for all } l\in\mathcal{L} \\ \text{and} \quad & [r_l]\in \mathrm{Co}(\mathcal{R}). \end{aligned} \tag{36}$$

The original problem (2) corresponds to $\vec{b} = 0$ and hence $G(0) = \sum_{s=1}^{S} U_s(x_s^*)$. It is easy to show that $G(\vec{b})$ is a concave function of $\vec{b}$. Hence, by Theorem 3.1.8 of [21, p36], there exists a subgradient $\vec{q}_0$ of $G(\vec{b})$ at $\vec{b} = 0$. We now show that $\vec{q}_0$ is the desired Lagrange multipliers. For any $\vec{b} \geq 0$, by the concavity of $G(\vec{b})$, we have

$$G(\vec{b}) \leq G(0) + \vec{q}_0^{\,\mathrm{tr}} \vec{b},$$

where $[\cdot]^{\mathrm{tr}}$ denotes the transpose. Further, by the definition of $G(\vec{b})$ in (36),

$$G(0) \leq G(\vec{b}) \text{ for all } \vec{b} \geq 0.$$

Hence, for any $\vec{b} \geq 0$,

$$G(0) \geq G(\vec{b}) - \vec{q}_0^{\,\mathrm{tr}} \vec{b} \geq G(0) - \vec{q}_0^{\,\mathrm{tr}} \vec{b},$$

and we have,

$$\vec{q}_0^{\,\mathrm{tr}} \vec{b} \geq 0 \text{ for all } \vec{b} \geq 0.$$

This implies that $\vec{q}_0 \geq 0$. Next, for any $\vec{x}$ such that $x_s \leq M_s$ for all s, and for any $\vec{r} \in \mathrm{Co}(\mathcal{R})$, if we let $\vec{g}(\vec{x}, \vec{r}) = \left[\sum_{s=1}^{S} H_s^l x_s - r_l, l \in \mathcal{L} \right]$, then $(\vec{x}, \vec{r})$ is a feasible point in the problem (36) with $\vec{b} = \vec{g}(\vec{x}, \vec{r})$. Hence, using the concavity of $G(\vec{b})$ again, we have

$$\begin{aligned} \sum_{s=1}^{S} U_s(x_s) &\leq G(\vec{g}(\vec{x}, \vec{r})) \\ &\leq G(0) + \vec{q}_0^{\,\mathrm{tr}} \vec{g}(\vec{x}, \vec{r}) \\ &= \sum_{s=1}^{S} U_s(x_s^*) + \vec{q}_0^{\,\mathrm{tr}} \vec{g}(\vec{x}, \vec{r}). \end{aligned} \tag{37}$$

Choosing $\vec{x} = \vec{x}^*$ and $\vec{r} = \vec{r}^{\,*}$, we have

$$\sum_{s=1}^{S} U_s(x_s^*) \leq \sum_{s=1}^{S} U_s(x_s^*) + \vec{q}_0^{\,\mathrm{tr}} \vec{g}(\vec{x}^*, \vec{r}^{\,*}),$$

i.e.,

$$\vec{q}_0^{\,\mathrm{tr}} \vec{g}(\vec{x}^*, \vec{r}^{\,*}) \geq 0.$$

However, since $\vec{g}(\vec{x}^*, \vec{r}^*) \leq 0$ and $\vec{q}_0 \geq 0$, we must have

$$\vec{q}_0^{\text{tr}} \vec{g}(\vec{x}^*, \vec{r}^*) = 0.$$

Finally, using (37) again, we obtain

$$\begin{aligned} L(\vec{x}, \vec{r}, \vec{q}_0) &= \sum_{s=1}^{S} U_s(x_s) - \vec{q}_0^{\text{tr}} \vec{g}(\vec{x}, \vec{r}) \\ &\leq \sum_{s=1}^{S} U_s(x_s^*) = \sum_{s=1}^{S} U_s(x_s^*) - \vec{q}_0^{\text{tr}} \vec{g}(\vec{x}^*, \vec{r}^*) \\ &= L(\vec{x}^*, \vec{r}^*, \vec{q}_0). \end{aligned}$$

Hence

$$\max_{0 \leq x_s \leq M_s, \vec{r} \in \text{Co}(\mathcal{R})} L(\vec{x}, \vec{r}, \vec{q}_0) = L(\vec{x}^*, \vec{r}^*, \vec{q}_0) = \sum_{s=1}^{S} U_s(x_s^*),$$

i.e., there is no duality gap.

Proof of part b): Let $\vec{x}^*$ denote the optimal solution to the primal problem (2), and let $\vec{r}^*$ denote the corresponding vector of link rates that satisfies (3). Note that $\vec{x}^*$ is unique due to the strict concavity of $U_s(x_s)$. We shall show that, for any $\vec{q} \in \Phi$, the following holds,

$$\max_{0 \leq x_s \leq M_s, \vec{r} \in \text{Co}(\mathcal{R})} L(\vec{x}, \vec{r}, \vec{q}) = L(\vec{x}^*, \vec{r}^*, \vec{q}).$$

(Note that $\vec{q}$ may be different from $\vec{q}_0$ in the proof of part (a).) In fact, since $\vec{q} \in \Phi$, we have,

$$\begin{aligned} \sum_{s=1}^{S} U_s(x_s^*) &= D(\vec{q}) = \max_{0 \leq x_s \leq M_s, \vec{r} \in \text{Co}(\mathcal{R})} L(\vec{x}, \vec{r}, \vec{q}) \\ &\geq \sum_{s=1}^{S} U_s(x_s^*) - \sum_{l=1}^{L} q^l \left(\sum_{s=1}^{S} H_s^l x_s^* - r_l^* \right). \end{aligned} \tag{38}$$

Hence,

$$\sum_{l=1}^{L} q^l \left(\sum_{s=1}^{S} H_s^l x_s^* - r_l^* \right) \geq 0.$$

However, since $\sum_{s=1}^{S} H_s^l x_s^* - r_l^* \leq 0$ for all l and $\vec{q} \geq 0$, we must have

$$\sum_{l=1}^{L} q^l \left(\sum_{s=1}^{S} H_s^l x_s^* - r_l^* \right) = 0.$$

Substituting into (38), we have,

$$\max_{0 \le x_s \le M_s, \vec{r} \in \mathrm{Co}(\mathcal{R})} L(\vec{x}, \vec{r}, \vec{q}) = \sum_{s=1}^{S} U_s(x_s^*) = L(\vec{x}^*, \vec{r}^{\,*}, \vec{q}).$$

However, given $\vec{q}$, the point $(\vec{x}, \vec{r})$ that maximizes

$$L(\vec{x}, \vec{r}, \vec{q}) = \sum_{s=1}^{S} U_s(x_s) - \sum_{l=1}^{L} q^l \left(\sum_{s=1}^{S} H_s^l x_s^* - r_l^* \right)$$

is unique due to the strict concavity of $U_s(x_s)$. Therefore, the maximizer must be equal to $\vec{x}^*$.

Proof of part c): Let A denote the $S \times S$ diagonal matrix whose l-th diagonal element is α_l^0. Let H denote the $L \times S$ matrix whose (l, s)-element is H_s^l. Then $||\vec{q}||_A = \vec{q}^{\,\mathrm{tr}} A^{-1} \vec{q}$. For any $\vec{q}^{\,*,0} \in \Phi$, by (12), we have

$$\begin{aligned} ||\vec{q}(t+1) - \vec{q}^{\,*,0}||_A \;&\le\; ||\vec{q}(t) - \vec{q}^{\,*,0}||_A + 2h[H\vec{x}(t) - \vec{r}(t)]^{\,\mathrm{tr}}[\vec{q}(t) - \vec{q}^{\,*,0}] \\ &\quad + h^2[H\vec{x}(t) - \vec{r}(t)]^{\,\mathrm{tr}} A [H\vec{x}(t) - \vec{r}(t)]. \end{aligned} \tag{39}$$

Note that

$$D(\vec{q}(t)) \;=\; \sum_{s=1}^{S} U_s(x_s(t)) - [H^{\,\mathrm{tr}} \vec{q}(t)]^{\,\mathrm{tr}} \vec{x}(t) + \vec{r}^{\,\mathrm{tr}}(t) \vec{q}(t),$$

and

$$\begin{aligned} D(\vec{q}^{\,*,0}) \;&=\; \max_{0 \le x_s \le M_s} \left\{ \sum_{s=1}^{S} U_s(x_s) - (H^{\,\mathrm{tr}} \vec{q}^{\,*,0})^{\,\mathrm{tr}} \vec{x} \right\} + \max_{\vec{r} \in \mathrm{Co}(\mathcal{R})} \vec{r}^{\,\mathrm{tr}} \vec{q}^{\,*,0} \\ &\ge\; \sum_{s=1}^{S} U_s(x_s(t)) - (H^{\,\mathrm{tr}} \vec{q}^{\,*,0})^{\,\mathrm{tr}} \vec{x}(t) + \vec{r}^{\,\mathrm{tr}}(t) \vec{q}^{\,*,0}. \end{aligned}$$

Hence,

$$D(\vec{q}^{\,*,0}) - D(\vec{q}(t)) \;\ge\; [H\vec{x}(t) - \vec{r}(t)]^{\,\mathrm{tr}} [\vec{q}(t) - \vec{q}^{\,*,0}].$$

Substituing into (39), we have

$$\begin{aligned} &||\vec{q}(t+1) - \vec{q}^{\,*,0}||_A \\ \le\;& ||\vec{q}(t) - \vec{q}^{\,*,0}||_A + 2h[D(\vec{q}^{\,*,0}) - D(\vec{q}(t))] + h^2[H\vec{x}(t) - \vec{r}(t)]^{\,\mathrm{tr}} A [H\vec{x}(t) - \vec{r}(t)]. \end{aligned}$$

Fix $\eta > 0$. Let

$$\Phi(\eta) = \{\vec{q} | D(\vec{q}) \leq D(\vec{q}^{\,*,0}) + \eta\}. \tag{40}$$

Since both $\vec{x}(t)$ and $\vec{r}(t)$ are bounded, there exists $M < \infty$ such that

$$\max_{0 \leq x_s \leq M_s, \vec{r} \in \mathrm{Co}(\mathcal{R})} (H\vec{x} - \vec{r})^{\,\mathrm{tr}} A (H\vec{x} - \vec{r}) \leq M.$$

If we pick

$$h \leq \eta / M,$$

then as long as $\vec{q}(t) \notin \Phi(\eta)$, we have

$$||\vec{q}(t+1) - \vec{q}^{\,*,0}||_A \leq ||\vec{q}(t) - \vec{q}^{\,*,0}||_A - h\eta.$$

Hence, eventually, $\vec{q}(t)$ will enter the set $\Phi(\eta)$. On the other hand, if we pick

$$h \leq \eta / \sqrt{M},$$

then once $\vec{q}(t) \in \Phi(\eta)$, we have

$$\sqrt{||\vec{q}(t+1) - \vec{q}^{\,*,0}||_A} \leq \sqrt{||\vec{q}(t) - \vec{q}^{\,*,0}||_A} + \sqrt{||\vec{q}(t+1) - \vec{q}(t)||_A} \leq \sqrt{||\vec{q}(t) - \vec{q}^{\,*,0}||_A} + \eta. \tag{41}$$

Since the inequality (41) holds for any $\vec{q}^{\,*,0} \in \Phi \subset \Phi(\eta)$, it implies that

$$d(\vec{q}(t+1), \Phi) \leq d(\vec{q}(t), \Phi) + \eta,$$

where $d(\vec{q}, \Phi) = \min_{\vec{p} \in \Phi} \sqrt{||\vec{q} - \vec{p}||_A}$. Hence, if

$$h \leq \min\{\eta/M, \eta/\sqrt{M}\},$$

then there exists time T_0 such that

$$d(\vec{q}(t), \Phi) \leq \xi(\eta) \text{ for all } t \geq T_0,$$

where

$$\xi(\eta) = \max_{\vec{p} \in \Phi(\eta)} d(\vec{p}, \Phi) + \eta.$$

It is easy to show that, as $\eta \to 0$,

$$\xi(\eta) \to 0.$$

Hence, for any $\epsilon > 0$, we can pick η (and h) sufficiently small such that $\xi(\eta) < \epsilon$, i.e., there exists time T_0 such that

$$d(\vec{q}(t), \Phi) < \epsilon \text{ for all } t \geq T_0.$$

Finally, since the mapping from $\vec{q}(t)$ to $\vec{x}(t)$ is continuous, we can pick η (and h) sufficiently small such that

$$||\vec{x}(t) - \vec{x}^*|| < \epsilon \text{ for all } t \geq T_0.$$

B Proof of Proposition 2

We need the following assumption on the queueing discipline at each link. We assume that, when each link forwards data, data at smaller number of hops away from their source will have priority over data at a large number of hops away from their source. Hence, first-hop data will be forwarded before all second-hop data, then second-hop data will be forwarded before all third-hop data, and so on. One way to achieve such priority queueing is to have each link maintain separate queues for data at different number of hops away from their sources.

The above assumption allows us to study the queue lengths at all links in the network in an inductive manner. We first study all first-hop traffic in isolation because first-hop traffic takes precedence over all other traffic. Once we compute the contribution to the queue lengths by the first-hop traffic, we can then study the second-hop traffic in the network, and so on.

Let $x_s(t,k)$ denote the data from user s injected at time t to the link that is at the k-th hop from the source of user s (let $x_s(t,k) = 0$ if data of user s travels at most $k_0 < k$ hops). Let $H_s^l(k) = 1$, if link l is at the k-th hop from user s, and let $H_s^l(k) = 0$, otherwise. Let $A^l(t,k)$ denote the amount of data injected to link l at time t by all first-hop through k-th hop traffic, i.e.,

$$A^l(t,k) = \sum_{s=1}^{S}\sum_{m=1}^{k} H_s^l(m) x_s(t,m).$$

Let $Q^l(t,k)$ denote the queue length at link l contributed by all first-hop through k-th hop traffic. Applying Loynes' formula, we have

$$Q^l(t,k) = \max_{0\le t'\le t}\left[\sum_{u=t-t'}^{t} A^l(u,k) - \sum_{u=t-t'}^{t} r_l(u)\right].$$

We now use induction to show that the queue lengths at all links are bounded. The induction hypothesis is as follows.

The Induction Hypothesis:

Fix k.

- For each user s, there exists a positive constant $M_s(k)$ such that

$$\sum_{u=t_0}^{t_0+t} x_s(u,k) \le \sum_{u=t_0}^{t_0+t} x_s(u) + M_s(k), \text{ for all } t_0 \text{ and } t. \tag{42}$$

- The queue length $Q^l(t,k)$ at link l contributed by all first-hop through k-th hop traffic is bounded for all t.

We first show that the inequality (42) implies the second part of the induction hypothesis. Assume that $\alpha_l = h\alpha_l^0$. By Proposition 1, there exists some $h_0 > 0$ such that for all $h < h_0$ and any initial implicit costs $\vec{q}(0)$, there exists a time T_0 such that

$$d(\vec{q}(t),\Phi) \le 1 \text{ for all } t \ge T_0.$$

Hence, $\vec{q}(t)$ is bounded for all t. Since

$$q^l(t+1) \ge q^l(t) + \alpha_l\left(\sum_{s=1}^{S} H_s^l x_s(t) - r_l(t)\right).$$

we have,

$$\sum_{u=t_0}^{t_0+t}\sum_{s=1}^{S} H_s^l x_s(u) - \sum_{u=t_0}^{t_0+t} r_l(u) \le \frac{1}{\alpha_l}\left[q^l(t_0+t+1) - q^l(t_0)\right].$$

Hence, the left hand side is bounded from above for all t_0 and t. Let $M^l(0)$ be this upper bound. We then have,

$$
\begin{aligned}
Q^l(t,k) &= \max_{0\le t'\le t}\left[\sum_{u=t-t'}^{t} A^l(u,k) - \sum_{u=t-t'}^{t} r_l(u)\right] \\
&= \max_{0\le t'\le t}\left[\sum_{u=t-t'}^{t}\sum_{m=1}^{k}\sum_{s=1}^{S} H_s^l(m)x_s(u,m) - \sum_{u=t-t'}^{t} r_l(u)\right] \\
&\le \max_{0\le t'\le t}\left[\sum_{u=t-t'}^{t}\sum_{s=1}^{S} H_s^l x_s(u) - \sum_{u=t-t'}^{t} r_l(u) + \sum_{m=1}^{k}\sum_{s=1}^{S} H_s^l(m)M_s(m)\right] \\
&\le M^l(0) + \sum_{m=1}^{k}\sum_{s=1}^{S} H_s^l(m)M_s(m).
\end{aligned}
$$

Hence, $Q^l(t,k)$ is bounded for all t.

We now use induction to show that the inequality (42) holds for all k. We first consider the case $k=1$, i.e., the first-hop traffic only. Since

$$x_s(t,1) = x_s(t),$$

the inequality (42) trivially holds. The second part of the induction hypothesis then follows (for $k=1$).

Assume that the induction hypothesis holds for $1,2,...,k-1$. Let $M(k-1)$ be the upper bound for $Q^l(t,k-1)$ for all t and l, i.e.,

$$M(k-1) = \sup_t \max_l Q^l(t,k-1).$$

We now consider the contribution by the k-th hop traffic. Note that

$$\sum_{u=t_0}^{t_0+t} x_s(u,k) \le \sum_{u=t_0}^{t_0+t} x_s(u,k-1) + M(k-1),$$

where the first term on the right hand side correponds to contribution from $(k-1)$-th hop traffic of user s, and the second term corresponds to the maximum amount of backlog at time t_0. Since the inequality (42) holds for $(k-1)$, we have,

$$\sum_{u=t_0}^{t_0+t} x_s(u,k) \le \sum_{u=t_0}^{t_0+t} x_s(u) + M_s(k-1) + M(k-1),$$

and hence the inequality (42) now holds for k. Again, by the discussion above, the second part of the induction hypothesis also holds for k.

Finally, let $\bar{L}$ denote the maximum number of hops of any users. Note that the overall queue length $Q^l(t)$ at link l is equal to $Q^l(t, \bar{L})$. Hence,

$$\sup_t Q^l(t) < +\infty \text{ for all } l \in \mathcal{L}.$$

C Proof of Proposition 3

Define

$$V_q(\vec{q}) = \sum_{l=1}^{L} \frac{(q^l)^2}{2\alpha_l}.$$

We now show that $V_q(\cdot)$ is a Lyapunov function of the system. In fact, using (17), we have,

$$V_q(\vec{q}(t+1)) - V_q(\vec{q}(t)) \leq \sum_{l=1}^{L} q^l(t) \left[\sum_{s=1}^{S} H_s^l x_s - r_l(t) \right] + E_1(t),$$

where

$$E_1(t) = \frac{1}{2} \sum_{l=1}^{L} \alpha_l \left[\sum_{s=1}^{S} H_s^l x_s - r_l(t) \right]^2 .$$

Since both x_s and $r_l(t)$ are bounded, $E_1(t)$ is bounded for all t. Hence,

$$V_q(\vec{q}(t+1)) - V_q(\vec{q}(t)) \leq \sum_{l=1}^{L} q^l(t) \left[\sum_{s=1}^{S} H_s^l x_s - r_l(t) \right] + E_1^0, \tag{43}$$

for some positive constant E_1^0. By assumption, $\vec{x}$ lies strictly inside $\gamma\Lambda$. Hence, there exists some $\epsilon \geq 0$ such that

$$(1+\epsilon)\vec{x} \in \gamma\Lambda,$$

i.e.,

$$\left[(1+\epsilon) \sum_{l=1}^{L} H_s^l x_s \,, \quad l \in \mathcal{L} \right] \in \gamma \mathrm{Co}(\mathcal{R}).$$

By the definition of the imperfect scheduling policy S_γ,

$$\sum_{l=1}^{L} q^l(t) r_l(t) \geq \gamma \max_{\vec{r} \geq 0, \vec{r} \in \mathcal{R}} \sum_{l=1}^{L} r_l q^l(t) = \gamma \max_{\vec{r} \geq 0, \vec{r} \in \mathrm{Co}(\mathcal{R})} \sum_{l=1}^{L} r_l q^l(t) \geq (1+\epsilon) \sum_{l=1}^{L} q^l(t) H_s^l x_s.$$

Substituting into (43), we have,

$$V_q(\vec{q}(t+1)) - V_q(\vec{q}(t)) \quad \leq \quad -\epsilon \sum_{l=1}^{L} q^l(t) H_s^l x_s + E_1^0.$$

By Theorem 2 of [17], the system in stable. *Q.E.D.*

D Proof of Proposition 5

Since the utility function is logarithmic, we have,

$$x_s(t) = \frac{w_s}{\sum\limits_{l=1}^{L} H_s^l q^l(t)} \text{ for all } t.$$

Taking limits as $t \to \infty$, we have,

$$x_s^{*,I} = \frac{w_s}{\sum\limits_{l=1}^{L} H_s^l q_I^{l,*}}. \tag{44}$$

Fix a positive number ϵ such that

$$\epsilon < \min_{l: q_I^{l,*} \neq 0} q_I^{l,*}.$$

Then there exists a time slot T_0 such that for all $t \geq T_0$

$$|x_s(t) - x_s^{*,I}| \quad \leq \quad \epsilon, \text{ and} \tag{45}$$

$$|q^l(t) - q_I^{l,*}| \quad \leq \quad \epsilon. \tag{46}$$

For each link $l \in \mathcal{L}$, there are two cases:

Case 1: If $q_I^{l,*} > 0$, then (46) implies that $q^l(t) > 0$ for all $t \geq T_0$. Hence, using (17), we have

$$\alpha_l \left(\sum_{s=1}^{S} H_s^l x_s(t) - r_l(t) \right) = q^l(t+1) - q^l(t) \text{ for all } t \geq T_0. \tag{47}$$

For any $T > 0$, summing (47) over $t = T_0, T_0+1, ..., T_0+T$, we have,

$$\alpha_l \left| \sum_{t=T_0}^{T_0+T} \sum_{s=1}^{S} H_s^l x_s(t) - \sum_{t=T_0}^{T_0+T} r_l(t) \right| = \left| q^l(T_0+T+1) - q^l(T_0) \right| \leq q_I^{l,*} + \epsilon.$$

We can thus pick T large enough such that

$$\left|\frac{1}{T}\sum_{t=T_0}^{T_0+T}\sum_{s=1}^{S}H_s^l x_s(t)-\frac{1}{T}\sum_{t=T_0}^{T_0+T}r_l(t)\right|<\epsilon.$$

Using (45), we have

$$\frac{1}{T}\sum_{t=T_0}^{T_0+T}r_l(t)\leq\sum_{s=1}^{S}H_s^l x_s^{*,I}+O(\epsilon), \tag{48}$$

where we have used $O(\epsilon)$ to denote the class of functions $f(\epsilon)$ such that $\limsup_{\epsilon\to 0} f(\epsilon)/\epsilon<+\infty$. Multiplying both side of (48) by $q_I^{l,*}$ and using (46) again, we have

$$q_I^{l,*}\frac{1}{T}\sum_{t=T_0}^{T_0+T}r_l(t) \quad\leq\quad q_I^{l,*}\sum_{s=1}^{S}H_s^l x_s^{*,I}+O(\epsilon). \tag{49}$$

Case 2: If $q_I^{l,*}=0$, the inequality (49) also holds trivially.

Summing (49) over all $l\in\mathcal{L}$ and using (44), we have

$$\begin{aligned}
&\sum_{l=1}^{L}q_I^{l,*}\frac{1}{T}\sum_{t=T_0}^{T_0+T}r_l(t)\\
\leq\ &\sum_{l=1}^{L}q_I^{l,*}\sum_{s=1}^{S}H_s^l x_s^{*,I}+O(\epsilon)\\
=\ &\sum_{s=1}^{S}x_s^{*,I}(\sum_{l=1}^{L}H_s^l q_I^{l,*})+O(\epsilon)\\
=\ &\sum_{s=1}^{S}w_s+O(\epsilon).
\end{aligned} \tag{50}$$

Let $\vec{x}^{\,*,\gamma}$ denote the solution to the γ-reduced problem. Then, $\frac{\vec{x}^{\,*,\gamma}}{\gamma}\in\Lambda$ by definition. Hence, by the definition of the imperfect schedule policy S_γ, $r_l(t)$ must satisfies

$$\begin{aligned}
\sum_{l=1}^{L}q^l(t)r_l(t) \quad&\geq\quad \gamma\sum_{l=1}^{L}q^l(t)\frac{\sum_{s=1}^{S}H_s^l x_s^{*,\gamma}}{\gamma}\\
&=\quad \sum_{l=1}^{L}q^l(t)\sum_{s=1}^{S}H_s^l x_s^{*,\gamma}\text{ for all } t.
\end{aligned}$$

Using (46) again, we obtain,

$$\sum_{l=1}^{L}q_I^{l,*}r_l(t)\geq\sum_{l=1}^{L}q_I^{l,*}\sum_{s=1}^{S}H_s^l x_s^{*,\gamma}-O(\epsilon)\text{ for all } t\geq T_0.$$

Substituting into (50), we have,

$$\begin{aligned} & \sum_{l=1}^{L} q_I^{l,*} \sum_{s=1}^{S} H_s^l x_s^{*,\gamma} \\ \leq & \sum_{l=1}^{L} q_I^{l,*} \frac{1}{T} \sum_{t=T_0}^{T_0+T} r_l(t) + O(\epsilon) \\ \leq & \sum_{s=1}^{S} w_s + O(\epsilon). \end{aligned}$$

Noting that

$$\sum_{l=1}^{L} q_I^{l,*} \sum_{s=1}^{S} H_s^l x_s^{*,\gamma} = \sum_{s=1}^{S} x_s^{*,\gamma} \sum_{l=1}^{L} H_s^l q_I^{l,*} = \sum_{s=1}^{S} \frac{w_s x_s^{*,\gamma}}{x_s^{*,I}},$$

we have,

$$\sum_{s=1}^{S} \frac{w_s x_s^{*,\gamma}}{x_s^{*,I}} \leq \sum_{s=1}^{S} w_s + O(\epsilon).$$

Finally, let $\epsilon \rightarrow 0$. The result then follows. *Q.E.D.*

E Proof of Proposition 6

Let A denote the $S \times S$ diagonal matrix whose l-th diagonal element is α_l^0. Let H denote the $L \times S$ matrix whose (l,s)-element is H_s^l. Then $||\vec{q}||_A = \vec{q}^{\,\mathrm{tr}} A^{-1} \vec{q}$. By (17), we have

$$\begin{aligned} ||\vec{q}(t+1) - \vec{q}^{\,*,0}||_A \ \leq \ & ||\vec{q}(t) - \vec{q}^{\,*,0}||_A + 2h[H\vec{x}(t) - \vec{r}(t)]^{\,\mathrm{tr}}[\vec{q}(t) - \vec{q}^{\,*,0}] \\ & + h^2 [H\vec{x}(t) - \vec{r}(t)]^{\,\mathrm{tr}} A [H\vec{x}(t) - \vec{r}(t)]. \end{aligned} \quad (51)$$

Note that

$$\begin{aligned} D_\gamma(\vec{q}(t)) \ = \ & \sum_{s=1}^{S} U_s(x_s(t)) - [H^{\,\mathrm{tr}} \vec{q}(t)]^{\,\mathrm{tr}} \vec{x}(t) + \gamma \max_{\vec{r} \in \mathrm{Co}(\mathcal{R})} \vec{r}^{\,\mathrm{tr}} \vec{q}(t) \\ \leq \ & \sum_{s=1}^{S} U_s(x_s(t)) - [H^{\,\mathrm{tr}} \vec{q}(t)]^{\,\mathrm{tr}} \vec{x}(t) + \vec{r}^{\,\mathrm{tr}}(t) \vec{q}(t), \end{aligned}$$

and

$$\begin{aligned} D(\vec{q}^{\,*,0}) &= \max_{0\le x_s\le M_s}\left\{\sum_{s=1}^{S} U_s(x_s) - [H^{\,\mathrm{tr}}\vec{q}^{\,*,0}]^{\,\mathrm{tr}}\vec{x}\right\} + \max_{\vec{r}\in \mathrm{Co}(\mathcal{R})} \vec{r}^{\,\mathrm{tr}}\vec{q}^{\,*,0} \\ &\ge \sum_{s=1}^{S} U_s(x_s(t)) - [H^{\,\mathrm{tr}}\vec{q}^{\,*,0}]^{\,\mathrm{tr}}\vec{x}(t) + \vec{r}^{\,\mathrm{tr}}(t)\vec{q}^{\,*,0}. \end{aligned}$$

Hence,

$$D(\vec{q}^{\,*,0}) - D_\gamma(\vec{q}(t)) \;\ge\; [H\vec{x}(t) - \vec{r}(t)]^{\,\mathrm{tr}}[\vec{q}(t) - \vec{q}^{\,*,0}].$$

Substituing into (51), we have

$$\begin{aligned} & ||\vec{q}(t+1) - \vec{q}^{\,*,0}||_A \\ \le\; & ||\vec{q}(t) - \vec{q}^{\,*,0}||_A + 2h[D(\vec{q}^{\,*,0}) - D_\gamma(\vec{q}(t))] + h^2[H\vec{x}(t) - \vec{r}(t)]^{\,\mathrm{tr}}A[H\vec{x}(t) - \vec{r}(t)] \end{aligned}$$

Fix $\eta > 0$. Let

$$\Phi_\gamma(\eta) = \{\vec{q}|D_\gamma(\vec{q}) \le D(\vec{q}^{\,*,0}) + \eta\}. \tag{52}$$

Since both $\vec{x}(t)$ and $\vec{r}(t)$ are bounded, there exists $M < +\infty$ such that

$$\max_{0\le x_s\le M_s, \vec{r}\in\mathrm{Co}(\mathcal{R})} (H\vec{x} - \vec{r})^{\,\mathrm{tr}}A(H\vec{x} - \vec{r}) \le M.$$

If we pick

$$h \le \eta/M,$$

then as long as $\vec{q}(t) \notin \Phi_\gamma(\eta)$, we have

$$||\vec{q}(t+1) - \vec{q}^{\,*,0}||_A \le ||\vec{q}(t) - \vec{q}^{\,*,0}||_A - h\eta.$$

Hence, eventually, $\vec{q}(t)$ will enter the set $\Phi_\gamma(\eta)$. On the other hand, if we pick

$$h \le \eta/\sqrt{M},$$

then once $\vec{q}(t) \in \Phi_\gamma(\eta)$, we have

$$\sqrt{||\vec{q}(t+1) - \vec{q}^{\,*,0}||_A} \le \sqrt{||\vec{q}(t) - \vec{q}^{\,*,0}||_A} + \sqrt{||\vec{q}(t+1) - \vec{q}(t)||_A} \le \sqrt{||\vec{q}(t) - \vec{q}^{\,*,0}||_A} + \eta.$$

Hence, if

$$h \leq \min\{\eta/M, \eta/\sqrt{M}\}.$$

then there exists a time T_0 such that

$$\sqrt{||\vec{q}(t) - \vec{q}^{\,*,0}||_A} \leq \xi(\eta) \text{ for all } t \geq T_0,$$

where

$$\xi(\eta) = \max_{\vec{p} \in \Phi_\gamma(\eta)} \sqrt{||\vec{p} - \vec{q}^{\,*,0}||_A} + \eta.$$

It is easy to show that, as $\eta \to 0$,

$$\xi(\eta) \to \max_{\vec{p} \in \Phi_\gamma} \sqrt{||\vec{p} - \vec{q}^{\,*,0}||_A}.$$

The result then follows.

Q.E.D.

F Proof of Proposition 7

Define

$$\mathcal{V}(\vec{n}, \vec{q}) = (1+\epsilon)V_n(\vec{n}) + V_q(\vec{q}),$$

where

$$V_n(\vec{n}) = \sum_{s=1}^{S} \frac{w_s n_s^2}{2\lambda_s}, \quad V_q(\vec{q}) = \sum_{l=1}^{L} \frac{(q^l)^2}{2\alpha_l},$$

and ϵ is a positive constant to be chosen later. We shall show that $\mathcal{V}(\cdot,\cdot)$ is a Lyapunov function of the system. We begin with a few lemmas. The first two lemmas bound the changes in $V_n(\cdot)$.

Lemma 9

$$\mathbf{E}[V_n(\vec{n}((k+1)T) - V_n(\vec{n}(kT))|\vec{n}(kT), \vec{q}(kT)]$$

$$
\leq \sum_{s=1}^{S} \left\{ \left[\sum_{l=1}^{L} H_s^l q^l(kT) \right] \left[\rho_s T - \int_{kT}^{(k+1)T} \mathbf{E}[n_s(t)x_s(t)|\vec{n}(kT), \vec{q}(kT)]dt \right] \right.
$$
$$
\left. - \frac{3w_s}{8\rho_s M_s} \int_{kT}^{(k+1)T} \mathbf{E}[n_s^2(t)x_s^2(t)|\vec{n}(kT), \vec{q}(kT)]dt \right\} + E_1 \tag{53}
$$

where E_1 is a finite positive constant.

Proof: Over a small time interval δt, we have

$$
\begin{aligned}
& \mathbf{E}\left[\frac{w_s}{2\lambda_s} \left[n_s^2(t+\delta t) - n_s^2(t) \right] |\vec{n}(t), \vec{q}(t) \right] \\
= & \frac{w_s}{2\lambda_s} \left\{ [(n_s(t)+1)^2 - n_s^2(t)]\lambda_s \delta t + [(n_s(t)-1)^2 - n_s^2(t)]\mu_s n_s(t)x_s(t)\delta t \right\} + o(\delta t) \\
= & \frac{w_s}{\lambda_s} [n_s(t)\lambda_s \delta t - n_s(t)\mu_s n_s(t)x_s(t)\delta t] + \frac{w_s}{2\lambda_s} [\lambda_s \delta t + \mu_s n_s(t)x_s(t)\delta t] + o(\delta t)
\end{aligned}
$$

Let $\rho_s = \lambda_s/\mu_s$. We have,

$$
\begin{aligned}
& \frac{\mathbf{E}[V_n(\vec{n}(t+\delta t)) - V_n(\vec{n}(t))|\vec{n}(t), \vec{q}(t)]}{\delta t} \\
\leq & \sum_{s=1}^{S} \left\{ \frac{w_s n_s(t)}{\lambda_s}[\lambda_s - \mu_s n_s(t)x_s(t)] + \frac{w_s}{2\lambda_s}[\lambda_s + \mu_s n_s(t)x_s(t)] \right\} + o(1) \\
= & \sum_{s=1}^{S} \left\{ \frac{w_s n_s(t)}{\rho_s}[\rho_s - n_s(t)x_s(t)] + \frac{w_s}{2}(1 + \frac{n_s(t)x_s(t)}{\rho_s}) \right\} + o(1) \qquad (54) \\
\leq & \sum_{s=1}^{S} \left\{ \left[\sum_{l=1}^{L} H_s^l q^l(t) \right] [\rho_s - n_s(t)x_s(t)] \right. \\
& + \left[\frac{w_s}{x_s(t)} - \sum_{l=1}^{L} H_s^l q^l(t) \right] [\rho_s - n_s(t)x_s(t)] \qquad (55) \\
& + w_s \left[\frac{n_s(t)}{\rho_s} - \frac{1}{x_s(t)} \right] [\rho_s - n_s(t)x_s(t)] \qquad (56) \\
& \left. + \frac{w_s}{2}(1 + \frac{n_s(t)x_s(t)}{\rho_s}) \right\} \qquad (57) \\
& + o(1),
\end{aligned}
$$

where $\rho_s = \lambda_s/\mu_s$. We shall bound the three terms (55-57). By (24),

$$
\frac{w_s}{x_s(t)} = \max\left\{ \sum_{l=1}^{L} H_s^l q^l(t), \frac{w_s}{M_s} \right\}.
$$

Hence, the term (55) can be bounded by

$$\begin{aligned}
& \left[\frac{w_s}{x_s(t)} - \sum_{l=1}^{L} H_s^l q^l(t)\right][\rho_s - n_s(t)x_s(t)] \\
\leq & \left[\frac{w_s}{x_s(t)} - \sum_{l=1}^{L} H_s^l q^l(t)\right]\rho_s \\
\leq & \left[\frac{w_s}{M_s} - \sum_{l=1}^{L} H_s^l q^l(t)\right]^+ \rho_s \\
\leq & \frac{w_s \rho_s}{M_s}. \qquad (58)
\end{aligned}$$

For the term (56), note that

$$\begin{aligned}
& \left[\frac{n_s(t)}{\rho_s} - \frac{1}{x_s(t)}\right][\rho_s - n_s(t)x_s(t)] \\
= & -\frac{[\rho_s - n_s(t)x_s(t)]^2}{\rho_s x_s(t)} \\
\leq & -\frac{[\rho_s - n_s(t)x_s(t)]^2}{\rho_s M_s}.
\end{aligned}$$

Using

$$[\rho_s - n_s(t)x_s(t)]^2 + \rho_s^2 \geq \frac{n_s^2(t)x_s^2(t)}{2},$$

we have,

$$\begin{aligned}
& \left[\frac{n_s(t)}{\rho_s} - \frac{1}{x_s(t)}\right][\rho_s - n_s(t)x_s(t)] \\
\leq & -\frac{1}{\rho_s M_s}\left[\frac{n_s^2(t)x_s^2(t)}{2} - \rho_s^2\right] \\
= & -\left[\frac{n_s^2(t)x_s^2(t)}{2\rho_s M_s} - \frac{\rho_s}{M_s}\right]. \qquad (59)
\end{aligned}$$

Finally, for the last term (57), we have

$$\frac{n_s(t)x_s(t)}{2\rho_s} \leq \frac{n_s^2(t)x_s^2(t)}{8\rho_s M_s} + \frac{M_s}{2\rho_s}. \qquad (60)$$

Substituting (58-60) back to (55-57), we have,

$$\frac{\mathbf{E}[V_n(\vec{n}(t+\delta t)) - V_n(\vec{n}(t))|\vec{n}(t), \vec{q}(t)]}{\delta t}$$

$$
\begin{aligned}
&\le \sum_{s=1}^{S}\left\{\left[\sum_{l=1}^{L} H_s^l q^l(t)\right][\rho_s - n_s(t)x_s(t)]\right. \\
&\quad \left. + \frac{w_s\rho_s}{M_s} - w_s\left[\frac{n_s^2(t)x_s^2(t)}{2\rho_s M_s} - \frac{\rho_s}{M_s}\right] + w_s\left[\frac{1}{2} + \frac{n_s^2(t)x_s^2(t)}{8\rho_s M_s} + \frac{M_s}{2\rho_s}\right]\right\} + o(1) \\
&= \sum_{s=1}^{S}\left\{\left[\sum_{l=1}^{L} H_s^l q^l(t)\right][\rho_s - n_s(t)x_s(t)] - \frac{3w_s n_s^2(t)x_s^2(t)}{8\rho_s M_s}\right. \\
&\quad \left. + w_s\left[\frac{1}{2} + \frac{M_s}{2\rho_s} + \frac{2\rho_s}{M_s}\right]\right\} + o(1).
\end{aligned} \tag{61}
$$

Integrating over $[kT, (k+1)T]$, and letting

$$
E_1 = \sum_{s=1}^{S} w_s T\left[\frac{1}{2} + \frac{M_s}{2\rho_s} + \frac{2\rho_s}{M_s}\right],
$$

the result (53) follows. *Q.E.D.*

Lemma 10

$$
\begin{aligned}
&\mathbf{E}[V_n(\vec{n}((k+1)T)) - V_n(\vec{n}(kT))|\vec{n}(kT), \vec{q}(kT)] \\
&\le \sum_{s=1}^{S}\left\{\rho_s T\left[\sum_{l=1}^{L} H_s^l q^l(kT)\right] - w_s\int_{kT}^{(k+1)T}\mathbf{E}[n_s(t)|\vec{n}(kT), \vec{q}(kT)]dt\right. \\
&\qquad \left. + \frac{w_s}{8\rho_s M_s}\int_{kT}^{(k+1)T}\mathbf{E}[n_s^2(t)x_s^2(t)|\vec{n}(kT), \vec{q}(kT)]dt\right\} + E_2,
\end{aligned} \tag{62}
$$

where E_2 is a finite positive constant.

Proof: From (54),

$$
\begin{aligned}
&\frac{\mathbf{E}[V_n(\vec{n}(t+\delta t)) - V_n(\vec{n}(t))|\vec{n}(t), \vec{q}(t)]}{\delta t} \\
&\le \sum_{s=1}^{S}\left\{\frac{w_s n_s(t)}{\rho_s}[\rho_s - n_s(t)x_s(t)] + \frac{w_s}{2}(1 + \frac{n_s(t)x_s(t)}{\rho_s})\right\} + o(1) \\
&\le \sum_{s=1}^{S}\left\{\frac{w_s}{x_s(t)}[\rho_s - n_s(t)x_s(t)] + \frac{w_s}{2}(1 + \frac{n_s(t)x_s(t)}{\rho_s})\right\} + o(1) \\
&= \sum_{s=1}^{S}\left\{\left[\frac{w_s\rho_s}{x_s(t)} - w_s n_s(t)\right] + \frac{w_s}{2}(1 + \frac{n_s(t)x_s(t)}{\rho_s})\right\} + o(1)
\end{aligned}
$$

By (24),

$$\frac{w_s\rho_s}{x_s(t)} = \rho_s \max\left\{\sum_{l=1}^{L} H_s^l q^l(t), \frac{w_s}{M_s}\right\} \le \rho_s\left(\sum_{l=1}^{L} H_s^l q^l(t) + \frac{w_s}{M_s}\right).$$

Combining with (60), we have,

$$\begin{aligned} &\frac{\mathbf{E}[V_n(\vec{n}(t+\delta t)) - V_n(\vec{n}(t))|\vec{n}(t), \vec{q}(t)]}{\delta t} \\ \le\ & \sum_{s=1}^{S}\left\{\left[\rho_s \sum_{l=1}^{L} H_s^l q^l(t) - w_s n_s(t)\right] + \frac{w_s n_s^2(t) x_s^2(t)}{8\rho_s M_s} + w_s\left[\frac{1}{2} + \frac{\rho_s}{M_s} + \frac{M_s}{2\rho_s}\right]\right\} + o(1). \quad (63)\end{aligned}$$

Integrating over $[kT, (k+1)T]$, and letting

$$E_2 = \sum_{s=1}^{S} w_s T\left[\frac{1}{2} + \frac{\rho_s}{M_s} + \frac{M_s}{2\rho_s}\right],$$

the result (62) then follows. *Q.E.D.*

The next lemma bounds the change in $V_q(\cdot)$. For simplicity, we use the following matrix notation. Let A denote the $L \times L$ diagonal matrix whose l-th diagonal element is α_l. Let H denote the $L \times S$ matrix whose (l, s)-element is H_s^l. Further, let $X_s(t) = n_s(t)x_s(t)$ and let $\vec{X}(t) = [X_1(t), ..., X_S(t)]$. Then

$$V_q(\vec{q}) = \frac{\vec{q}^{\,\mathrm{tr}} A^{-1} \vec{q}}{2},$$

and the update on the implicit costs (25) can be written as

$$\vec{q}((k+1)T) = \left[\vec{q}(kT) + A\left(H\int_{kT}^{(k+1)T} \vec{X}(t)dt - \vec{r}(kT)T\right)\right]^+. \quad (64)$$

Lemma 11

$$\begin{aligned} &\mathbf{E}[V_q(\vec{q}((k+1)T) - V_q(\vec{q}(kT))|\vec{n}(kT), \vec{q}(kT)] \\ \le\ & \vec{q}^{\,\mathrm{tr}}(kT)\left[H\int_{kT}^{(k+1)T} \mathbf{E}[\vec{X}(t)|\vec{n}(kT), \vec{q}(kT)]dt - \vec{r}(kT)T\right] \\ & + T\alpha_{\max}\bar{S}\bar{L}\sum_{s=1}^{S}\left[\int_{kT}^{(k+1)T} \mathbf{E}[n_s^2(t)x_s^2(t)|\vec{n}(kT), \vec{q}(kT)]dt\right] + E_3, \quad (65)\end{aligned}$$

where

$$\alpha_{\max} = \max_{l\in\mathcal{L}} \alpha^l, \bar{L} = \max_s \sum_{l=1}^{L} H_s^l, \bar{S} = \max_l \sum_{s=1}^{S} H_s^l,$$

and E_3 is a finite positive constant, .

Proof: By (64),

$$\begin{aligned}
& V_q(\vec{q}((k+1)T) - V_q(\vec{q}(kT)) \\
\leq\ & \vec{q}^{\,\mathrm{tr}}(kT)\left[H\int_{kT}^{(k+1)T}\vec{X}(t)dt - \vec{r}(kT)T\right] \\
& + \frac{1}{2}\left[H\int_{kT}^{(k+1)T}\vec{X}(t)dt - \vec{r}(kT)T\right]^{\mathrm{tr}} A \left[H\int_{kT}^{(k+1)T}\vec{X}(t)dt - \vec{r}(kT)T\right] \\
\leq\ & \vec{q}^{\,\mathrm{tr}}(kT)\left[H\int_{kT}^{(k+1)T}\vec{X}(t)dt - \vec{r}(kT)T\right] \\
& + \left[H\int_{kT}^{(k+1)T}\vec{X}(t)dt\right]^{\mathrm{tr}} A \left[H\int_{kT}^{(k+1)T}\vec{X}(t)dt\right] + T^2[\vec{r}(kT)]^{\mathrm{tr}} A\vec{r}(kT)
\end{aligned}$$

where $[\cdot]^{\mathrm{tr}}$ denotes the transpose. Let

$$\alpha_{\max} = \max_{l\in\mathcal{L}}\alpha^l, \bar{L} = \max_s \sum_{l=1}^{L} H_s^l, \bar{S} = \max_l \sum_{s=1}^{S} H_s^l.$$

Then, we have,

$$\begin{aligned}
& \left[H\int_{kT}^{(k+1)T}\vec{X}(t)dt\right]^{\mathrm{tr}} A \left[H\int_{kT}^{(k+1)T}\vec{X}(t)dt\right] \\
=\ & \sum_{l=1}^{L}\alpha_l\left[\sum_{s=1}^{S}H_s^l\int_{kT}^{(k+1)T} n_s(t)x_s(t)dt\right]^2 \\
\leq\ & \sum_{l=1}^{L}\alpha_l\left[\sum_{s=1}^{S}H_s^l\right]\left[\sum_{s=1}^{S}H_s^l\left(\int_{kT}^{(k+1)T} n_s(t)x_s(t)dt\right)^2\right] \\
\leq\ & \bar{S}\sum_{l=1}^{L}\alpha_l\left[\sum_{s=1}^{S}H_s^l\left(\int_{kT}^{(k+1)T} n_s(t)x_s(t)dt\right)^2\right] \\
\leq\ & T\bar{S}\sum_{l=1}^{L}\alpha_l\sum_{s=1}^{S}H_s^l\int_{kT}^{(k+1)T} n_s^2(t)x_s^2(t)dt \\
=\ & T\bar{S}\sum_{s=1}^{S}\left[\int_{kT}^{(k+1)T} n_s^2(t)x_s^2(t)dt\right]\left[\sum_{l=1}^{L}\alpha_l H_s^l\right] \\
\leq\ & T\alpha_{\max}\bar{S}\bar{L}\sum_{s=1}^{S}\left[\int_{kT}^{(k+1)T} n_s^2(t)x_s^2(t)dt\right].
\end{aligned}$$

Letting

$$E_3 = \max_{\vec{r}\in\mathrm{Co}(\mathcal{R})} T^2\vec{r}^{\,\mathrm{tr}}A\vec{r},$$

the result (65) then follows. *Q.E.D.*

Proof of Proposition 7 : Multiply (62) by $\epsilon < 1$ and add to (53). We have

$$\begin{aligned}
& (1+\epsilon)\mathbf{E}[V_n(\vec{n}((k+1)T)) - V_n(\vec{n}(kT))|\vec{n}(kT),\vec{q}(kT)] \\
\leq & \sum_{s=1}^{S}\Bigg\{\left[\sum_{l=1}^{L} H_s^l q^l(kT)\right]\left[(1+\epsilon)\rho_s T - \int_{kT}^{(k+1)T}\mathbf{E}[n_s(t)x_s(t)|\vec{n}(kT),\vec{q}(kT)]dt\right] \\
& \quad -\epsilon w_s \int_{kT}^{(k+1)T}\mathbf{E}[n_s(t)|\vec{n}(kT),\vec{q}(kT)]dt \\
& \quad -\frac{w_s}{4\rho_s M_s}\int_{kT}^{(k+1)T}\mathbf{E}[n_s^2(t)x_s^2(t)|\vec{n}(kT),\vec{q}(kT)]dt\Bigg\} + E_1 + E_2 \qquad (66)
\end{aligned}$$

Adding (65) to (66), and noting that

$$\begin{aligned}
& \sum_{s=1}^{S}\left\{\left[\sum_{l=1}^{L} H_s^l q^l(kT)\right]\int_{kT}^{(k+1)T}\mathbf{E}[n_s(t)x_s(t)|\vec{n}(kT),\vec{q}(kT)]dt\right\} \\
= & \sum_{l=1}^{L} q^l(kT)\sum_{s=1}^{S} H_s^l \int_{kT}^{(k+1)T}\mathbf{E}[n_s(t)x_s(t)|\vec{n}(kT),\vec{q}(kT)]dt \\
= & \vec{q}^{\,\mathrm{tr}}(kT)\left[H\int_{kT}^{(k+1)T}\mathbf{E}[\vec{X}(t)|\vec{n}(kT),\vec{q}(kT)]dt\right],
\end{aligned}$$

we have

$$\begin{aligned}
& \mathbf{E}[\mathcal{V}(\vec{n}((k+1)T),\vec{q}((k+1)T)) - \mathcal{V}(\vec{n}(kT),\vec{q}(kT))|\vec{n}(kT),\vec{q}(kT)] \\
\leq & \sum_{s=1}^{S}\left[\sum_{l=1}^{L} H_s^l q^l(kT)\right](1+\epsilon)\rho_s T - \vec{q}^{\,\mathrm{tr}}(kT)\vec{r}(kT)T \\
& -\epsilon\sum_{s=1}^{S} w_s\int_{kT}^{(k+1)T}\mathbf{E}[n_s(t)|\vec{n}(kT),\vec{q}(kT)]dt \\
& -\sum_{s=1}^{S}\left[\frac{w_s}{4\rho_s M_s} - T\alpha_{\max}\bar{S}\bar{L}\right]\int_{kT}^{(k+1)T}\mathbf{E}[n_s^2(t)x_s^2(t)|\vec{n}(kT),\vec{q}(kT)]dt \qquad (67) \\
& +E_0,
\end{aligned}$$

where $E_0 = E_1 + E_2 + E_3$. If (26) is satisfied, then

$$T\alpha_{\max}\bar{S}\bar{L} \leq \frac{w_s}{4\rho_s M_s} \text{ for all } s.$$

Hence, the term in (67) is negative. By a rearrangement of the order of the summation, we have,

$$\begin{aligned}
& \mathbf{E}[\mathcal{V}(\vec{n}((k+1)T), \vec{q}((k+1)T)) - \mathcal{V}(\vec{n}(kT), \vec{q}(kT)) | \vec{n}(kT), \vec{q}(kT)] \\
\leq \; & T \vec{q}^{\,\mathrm{tr}}(kT) \left[(1+\epsilon) H \vec{\rho} - \vec{r}(kT) \right] \\
& - \epsilon \sum_{s=1}^{S} w_s \int_{kT}^{(k+1)T} \mathbf{E}[n_s(t) | \vec{n}(kT), \vec{q}(kT)] dt + E_0,
\end{aligned}$$

where $\vec{\rho} = [\rho_1, ..., \rho_s]$. By assumption, $\vec{\rho}$ lies strictly inside $\gamma\Lambda$. Hence, there exists some $\epsilon > 0$ such that $(1+2\epsilon)\vec{\rho} \in \gamma\Lambda$, i.e.,

$$(1+2\epsilon) H \vec{\rho} \in \gamma \mathrm{Co}(\mathcal{R}).$$

Use this value of ϵ in the definition of $\mathcal{V}(\cdot,\cdot)$. Further, by the definition of the imperfect scheduling policy S_γ,

$$\vec{q}^{\,\mathrm{tr}}(kT) \vec{r}(kT) \geq \gamma \max_{\vec{r} \in \Lambda} \vec{q}^{\,\mathrm{tr}}(kT) \vec{r} \geq (1+2\epsilon) \vec{q}^{\,\mathrm{tr}}(kT) H \vec{\rho}.$$

Hence,

$$\begin{aligned}
& \mathbf{E}[\mathcal{V}(\vec{n}((k+1)T), \vec{q}((k+1)T)) - \mathcal{V}(\vec{n}(kT), \vec{q}(kT)) | \vec{n}(kT), \vec{q}(kT)] \\
\leq \; & -\epsilon T \vec{q}^{\,\mathrm{tr}}(kT) H \vec{\rho} - \epsilon \sum_{s=1}^{S} w_s \int_{kT}^{(k+1)T} \mathbf{E}[n_s(t) | \vec{n}(kT), \vec{q}(kT)] dt + E_0 \\
\leq \; & -\epsilon T \vec{q}^{\,\mathrm{tr}}(kT) H \vec{\rho} - \epsilon' \sum_{s=1}^{S} w_s n_s(kT) + E_0
\end{aligned}$$

for some $\epsilon' > 0$. By Theorem 2 of [17], the result then follows. *Q.E.D.*

G Proof of Proposition 8

Recall that

$$Q_i = \sum_{j:(i,j) \in \mathcal{L}} q_{ij} + \sum_{j:(j,i) \in \mathcal{L}} q_{ji}$$

denote the total cost of the links that either start from, or end at node i. Define

$$\mathcal{V}(\vec{n}, \vec{q}) = (1+\epsilon) V_n(\vec{n}) + V_q(\vec{q}),$$

where

$$V_n(\vec{n}) = \sum_{s=1}^{S} \frac{w_s n_s^2}{2\lambda_s}, \quad V_q(\vec{q}) = \frac{\sum_{i=1}^{N} Q_i^2}{\alpha}, \tag{68}$$

and ϵ is a positive constant to be chosen later. (Note that the definition of $V_q(\cdot)$ is different from that in the earlier proofs.) We shall show that $\mathcal{V}(\cdot,\cdot)$ is a Lyapunov function of the system. In fact, analogous to Lemmas 9 and 10, we can show that,

$$\begin{aligned}
&\mathbf{E}[V_n(\vec{n}((k+1)T) - V_n(\vec{n}(kT))|\vec{n}(kT), \vec{q}(kT)] \\
\leq\ & \sum_{s=1}^{S} \left\{ \left[2 \sum_{(i,j)\in\mathcal{L}} H_s^{ij} \frac{Q_i(kT) + Q_j(kT)}{c_{ij}} \right] \left[\rho_s T - \int_{kT}^{(k+1)T} \mathbf{E}[n_s(t)x_s(t)|\vec{n}(kT), \vec{q}(kT)]dt \right] \right. \\
& \left. - \frac{3w_s}{8\rho_s M_s} \int_{kT}^{(k+1)T} \mathbf{E}[n_s^2(t)x_s^2(t)|\vec{n}(kT), \vec{q}(kT)]dt \right\} + E_1
\end{aligned} \tag{69}$$

and

$$\begin{aligned}
&\mathbf{E}[V_n(\vec{n}((k+1)T)) - V_n(\vec{n}(kT))|\vec{n}(kT), \vec{q}(kT)] \\
\leq\ & \sum_{s=1}^{S} \left\{ \rho_s T \left[2 \sum_{(i,j)\in\mathcal{L}} H_s^{ij} \frac{Q_i(kT) + Q_j(kT)}{c_{ij}} \right] - w_s \int_{kT}^{(k+1)T} \mathbf{E}[n_s(t)|\vec{n}(kT), \vec{q}(kT)]dt \right. \\
& \left. + \frac{w_s}{8\rho_s M_s} \int_{kT}^{(k+1)T} \mathbf{E}[n_s^2(t)x_s^2(t)|\vec{n}(kT), \vec{q}(kT)]dt \right\} + E_2,
\end{aligned} \tag{70}$$

where E_1 and E_2 are finite positive constants. Multiply (70) by $\epsilon < 1$ and add to (69). We have

$$\begin{aligned}
&(1+\epsilon)\mathbf{E}[V_n(\vec{n}((k+1)T)) - V_n(\vec{n}(kT))|\vec{n}(kT), \vec{q}(kT)] \\
\leq\ & \sum_{s=1}^{S} \left\{ \left[2 \sum_{(i,j)\in\mathcal{L}} H_s^{ij} \frac{Q_i(kT) + Q_j(kT)}{c_{ij}} \right] \left[(1+\epsilon)\rho_s T - \int_{kT}^{(k+1)T} \mathbf{E}[n_s(t)x_s(t)|\vec{n}(kT), \vec{q}(kT)]dt \right] \right. \\
& - \epsilon w_s \int_{kT}^{(k+1)T} \mathbf{E}[n_s(t)|\vec{n}(kT), \vec{q}(kT)]dt \\
& \left. - \frac{w_s}{4\rho_s M_s} \int_{kT}^{(k+1)T} \mathbf{E}[n_s^2(t)x_s^2(t)|\vec{n}(kT), \vec{q}(kT)]dt \right\} + E_1 + E_2
\end{aligned} \tag{71}$$

As in Lemma 11, we shall prove the following lemma bounding the change in $V_q(\cdot)$.

Lemma 12 *If $\alpha < 1/T$, then*

$$
\begin{aligned}
& \mathbf{E}[V_q(\vec{q}((k+1)T)) - V_q(\vec{q}(kT))|\vec{n}(kT), \vec{q}(kT)] \\
\leq\ & 2 \sum_{(i,j)\in\mathcal{L}} q_{ij}(kT) \Bigg\{ \sum_{m:(i,m)\in\mathcal{L}} \sum_{s=1}^{S} H_s^{im} \int_{kT}^{(k+1)T} \frac{\mathbf{E}[n_s(t)x_s(kT)|\vec{n}(kT), \vec{q}(kT)]}{c_{im}} dt \\
& \quad + \sum_{m:(m,i)\in\mathcal{L}} \sum_{s=1}^{S} H_s^{mi} \int_{kT}^{(k+1)T} \frac{\mathbf{E}[n_s(t)x_s(kT)|\vec{n}(kT), \vec{q}(kT)]}{c_{mi}} dt \\
& \quad + \sum_{h:(j,h)\in\mathcal{L}} \sum_{s=1}^{S} H_s^{jh} \int_{kT}^{(k+1)T} \frac{\mathbf{E}[n_s(t)x_s(kT)|\vec{n}(kT), \vec{q}(kT)]}{c_{jh}} dt \\
& \quad + \sum_{h:(h,j)\in\mathcal{L}} \sum_{s=1}^{S} H_s^{hj} \int_{kT}^{(k+1)T} \frac{\mathbf{E}[n_s(t)x_s(kT)|\vec{n}(kT), \vec{q}(kT)]}{c_{hj}} dt - T \Bigg\} \\
& + \alpha E_3 \sum_{s=1}^{S} \left[\int_{kT}^{(k+1)T} \mathbf{E}[n_s^2(t)x_s^2(t)|\vec{n}(kT), \vec{q}(kT)] dt \right] + E_4,
\end{aligned}
\tag{72}
$$

where E_3 and E_4 are positive constants.

Proof: By the definition of the maximal matching $\mathcal{M}(kT)$, $(i,j) \in \mathcal{M}(kT)$ implies $q_{ij}(kT) \geq 1$. Further, since $\alpha < 1/T$, the projection operator in (35) is not needed. Hence,

$$
\begin{aligned}
& [Q_i((k+1)T)]^2 - [Q_i(kT)]^2 \\
=\ & 2\alpha T \left[\sum_{j:(i,j)\in\mathcal{L}} q_{ij}(kT) + \sum_{j:(j,i)\in\mathcal{L}} q_{ji}(kT) \right] \\
& \quad \times \left[\sum_{j:(i,j)\in\mathcal{L}} \left(\frac{1}{T} \sum_{s=1}^{S} H_s^{ij} \int_{kT}^{(k+1)T} \frac{n_s(t)x_s(kT)}{c_{ij}} dt - \mathbf{1}_{\{(i,j)\in\mathcal{M}(kT)\}} \right) \right. \\
& \qquad \left. + \sum_{j:(j,i)\in\mathcal{L}} \left(\frac{1}{T} \sum_{s=1}^{S} H_s^{ji} \int_{kT}^{(k+1)T} \frac{n_s(t)x_s(kT)}{c_{ji}} dt - \mathbf{1}_{\{(j,i)\in\mathcal{M}(kT)\}} \right) \right] \\
& + \alpha^2 T^2 \left[\sum_{j:(i,j)\in\mathcal{L}} \left(\frac{1}{T} \sum_{s=1}^{S} H_s^{ij} \int_{kT}^{(k+1)T} \frac{n_s(t)x_s(kT)}{c_{ij}} dt - \mathbf{1}_{\{(i,j)\in\mathcal{M}(kT)\}} \right) \right. \\
& \qquad \left. + \sum_{j:(j,i)\in\mathcal{L}} \left(\frac{1}{T} \sum_{s=1}^{S} H_s^{ji} \int_{kT}^{(k+1)T} \frac{n_s(t)x_s(kT)}{c_{ji}} dt - \mathbf{1}_{\{(j,i)\in\mathcal{M}(kT)\}} \right) \right]^2 .
\end{aligned}
$$

Substituting into (68) and rearranging the terms, we have,

$$
\begin{aligned}
& V_q(\vec{q}((k+1)T)) - V_q(\vec{q}(kT)) \\
= \; & 2 \sum_{(i,j)\in\mathcal{L}} q_{ij}(kT) \Bigg[\sum_{m:(i,m)\in\mathcal{L}} \sum_{s=1}^{S} H_s^{im} \int_{kT}^{(k+1)T} \frac{n_s(t)x_s(kT)}{c_{im}} dt \\
& \quad + \sum_{m:(m,i)\in\mathcal{L}} \sum_{s=1}^{S} H_s^{mi} \int_{kT}^{(k+1)T} \frac{n_s(t)x_s(kT)}{c_{mi}} dt \\
& \quad + \sum_{h:(j,h)\in\mathcal{L}} \sum_{s=1}^{S} H_s^{jh} \int_{kT}^{(k+1)T} \frac{n_s(t)x_s(kT)}{c_{jh}} dt \\
& \quad + \sum_{h:(h,j)\in\mathcal{L}} \sum_{s=1}^{S} H_s^{hj} \int_{kT}^{(k+1)T} \frac{n_s(t)x_s(kT)}{c_{hj}} dt \Bigg] \\
& -2T \sum_{(i,j)\in\mathcal{L}} q_{ij}(kT) \Bigg[\sum_{m:(i,m)\in\mathcal{L}} \mathbf{1}_{\{(i,m)\in\mathcal{M}(kT)\}} + \sum_{m:(m,i)\in\mathcal{L}} \mathbf{1}_{\{(m,i)\in\mathcal{M}(kT)\}} \\
& \quad + \sum_{h:(j,h)\in\mathcal{L}} \mathbf{1}_{\{(j,h)\in\mathcal{M}(kT)\}} + \sum_{h:(h,j)\in\mathcal{L}} \mathbf{1}_{\{(h,j)\in\mathcal{M}(kT)\}} \Bigg] + E_4(k),
\end{aligned} \tag{73}
$$

where

$$
\begin{aligned}
E_4(k) \;=\; & \alpha T^2 \sum_{i=1}^{N} \Bigg[\sum_{j:(i,j)\in\mathcal{L}} \left(\frac{1}{T} \sum_{s=1}^{S} H_s^{ij} \int_{kT}^{(k+1)T} \frac{n_s(t)x_s(kT)}{c_{ij}} dt - \mathbf{1}_{\{(i,j)\in\mathcal{M}(kT)\}} \right) \\
& + \sum_{j:(j,i)\in\mathcal{L}} \left(\frac{1}{T} \sum_{s=1}^{S} H_s^{ji} \int_{kT}^{(k+1)T} \frac{n_s(t)x_s(kT)}{c_{ji}} dt - \mathbf{1}_{\{(j,i)\in\mathcal{M}(kT)\}} \right) \Bigg]^2 .
\end{aligned}
$$

Similar to the proof of Lemma 11, we can show that

$$
E_4(k) \le \alpha E_3 \sum_{s=1}^{S} \left[\int_{kT}^{(k+1)T} n_s^2(t) x_s^2(t) dt \right] + E_5, \tag{74}
$$

where E_3 and E_5 are positive constants. (Note that E_3 can be shown to only depend on the topology of the network.) Further, by the definition of the maximal matching $\mathcal{M}(kT)$,

$$
\begin{aligned}
& \sum_{m:(i,m)\in\mathcal{L}} \mathbf{1}_{\{(i,m)\in\mathcal{M}(kT)\}} + \sum_{m:(m,i)\in\mathcal{L}} \mathbf{1}_{\{(m,i)\in\mathcal{M}(kT)\}} \\
& \quad + \sum_{h:(j,h)\in\mathcal{L}} \mathbf{1}_{\{(j,h)\in\mathcal{M}(kT)\}} + \sum_{h:(h,j)\in\mathcal{L}} \mathbf{1}_{\{(h,j)\in\mathcal{M}(kT)\}} \ge 1, \\
& \qquad \text{for all } (i,j) \text{ such that } q_{ij}(kT) \ge 1.
\end{aligned}
$$

Hence,

$$-q_{ij}(kT)\left[\sum_{m:(i,m)\in\mathcal{L}}\mathbf{1}_{\{(i,m)\in\mathcal{M}(kT)\}}+\sum_{m:(m,i)\in\mathcal{L}}\mathbf{1}_{\{(m,i)\in\mathcal{M}(kT)\}}\right.$$
$$\left.+\sum_{h:(j,h)\in\mathcal{L}}\mathbf{1}_{\{(j,h)\in\mathcal{M}(kT)\}}+\sum_{h:(h,j)\in\mathcal{L}}\mathbf{1}_{\{(h,j)\in\mathcal{M}(kT)\}}\right]\leq -q_{ij}(kT)+1,$$
$$\text{for all } (i,j)\in\mathcal{L}. \tag{75}$$

Substituting (74) and (75) into (73), the result (72) then follows with $E_4 = E_5 + 2LT$, where L is the total number of links. *Q.E.D.*

Proof of Proposition 8 : Note that for all $a_s, s = 1, ..., S$,

$$\begin{aligned}
&\sum_{(i,j)\in\mathcal{L}} q_{ij}(kT)\sum_{m:(i,m)\in\mathcal{L}}\sum_{s=1}^{S}\frac{H_s^{im}a_s}{c_{im}}\\
=\ &\sum_{s=1}^{S}a_s\sum_{(i,m)\in\mathcal{L}}H_s^{im}\frac{\sum_{j:(i,j)\in\mathcal{L}}q_{ij}(kT)}{c_{im}}\\
=\ &\sum_{s=1}^{S}a_s\sum_{(i,j)\in\mathcal{L}}H_s^{ij}\frac{\sum_{m:(i,m)\in\mathcal{L}}q_{im}(kT)}{c_{ij}}.
\end{aligned}\tag{76}$$

Similarly,

$$\begin{aligned}
\sum_{(i,j)\in\mathcal{L}} q_{ij}(kT)\sum_{m:(m,i)\in\mathcal{L}}\sum_{s=1}^{S}\frac{H_s^{mi}a_s}{c_{mi}} &= \sum_{s=1}^{S}a_s\sum_{(i,j)\in\mathcal{L}}H_s^{ij}\frac{\sum_{m:(m,i)\in\mathcal{L}}q_{mi}(kT)}{c_{ij}}\\
\sum_{(i,j)\in\mathcal{L}} q_{ij}(kT)\sum_{h:(j,h)\in\mathcal{L}}\sum_{s=1}^{S}\frac{H_s^{jh}a_s}{c_{jh}} &= \sum_{s=1}^{S}a_s\sum_{(i,j)\in\mathcal{L}}H_s^{ij}\frac{\sum_{h:(j,h)\in\mathcal{L}}q_{jh}(kT)}{c_{ij}}\\
\sum_{(i,j)\in\mathcal{L}} q_{ij}(kT)\sum_{h:(h,j)\in\mathcal{L}}\sum_{s=1}^{S}\frac{H_s^{hj}a_s}{c_{hj}} &= \sum_{s=1}^{S}a_s\sum_{(i,j)\in\mathcal{L}}H_s^{ij}\frac{\sum_{h:(h,j)\in\mathcal{L}}q_{hj}(kT)}{c_{ij}}.
\end{aligned}\tag{77}$$

Hence, by Lemma 12,

$$\begin{aligned}
&\mathbf{E}[V_q(\vec{q}((k+1)T)-V_q(\vec{q}(kT))|\vec{n}(kT),\vec{q}(kT)]\\
\leq\ & 2\sum_{s=1}^{S}\left[\sum_{(i,j)\in\mathcal{L}}H_s^{ij}\frac{Q_i(kT)+Q_j(kT)}{c_{ij}}\right]\int_{kT}^{(k+1)T}\mathbf{E}[n_s(t)x_s(t)|\vec{n}(kT),\vec{q}(kT)]dt\\
&-2T\sum_{(i,j)\in\mathcal{L}}q_{ij}(kT)+\alpha E_3\sum_{s=1}^{S}\left[\int_{kT}^{(k+1)T}\mathbf{E}[n_s^2(t)x_s^2(t)|\vec{n}(kT),\vec{q}(kT)]dt\right]+E_4.
\end{aligned}\tag{78}$$

Adding (78) to (71), we have

$$\begin{aligned}
& \mathbf{E}[\mathcal{V}(\vec{n}((k+1)T),\vec{q}((k+1)T))-\mathcal{V}(\vec{n}(kT),\vec{q}(kT))|\vec{n}(kT),\vec{q}(kT)] \\
\leq\; & 2T(1+\epsilon)\sum_{s=1}^{S}\rho_s\left[\sum_{(i,j)\in\mathcal{L}}H_s^{ij}\frac{Q_i(kT)+Q_j(kT)}{c_{ij}}\right]-2T\sum_{(i,j)\in\mathcal{L}}q_{ij}(kT) \\
& -\epsilon\sum_{s=1}^{S}w_s\int_{kT}^{(k+1)T}\mathbf{E}[n_s(t)|\vec{n}(kT),\vec{q}(kT)]dt \\
& -\sum_{s=1}^{S}\left[\frac{w_s}{4\rho_s M_s}-\alpha E_3\right]\int_{kT}^{(k+1)T}\mathbf{E}[n_s^2(t)x_s^2(t)|\vec{n}(kT),\vec{q}(kT)]dt \\
& +E_7,
\end{aligned} \tag{79}$$

where $E_7 = E_1 + E_2 + E_4$. When α is sufficiently small, the product term in (79) is negative. Further, by assumption, $[\rho_s]$ lies strictly inside $\frac{\Lambda}{2}$. Hence, there exists some positive number ϵ such that, for all node i,

$$(1+2\epsilon)\left[\sum_{j:(i,j)\in\mathcal{L}}\sum_{s=1}^{S}\frac{H_s^{ij}\rho_s}{c_{ij}}+\sum_{j:(j,i)\in\mathcal{L}}\sum_{s=1}^{S}\frac{H_s^{ji}\rho_s}{c_{ji}}\right]\leq 1/2.$$

Hence, applying (76-77) again on the inequality (79), we have,

$$\begin{aligned}
& \mathbf{E}[\mathcal{V}(\vec{n}((k+1)T),\vec{q}((k+1)T))-\mathcal{V}(\vec{n}(kT),\vec{q}(kT))|\vec{n}(kT),\vec{q}(kT)] \\
\leq\; & 2T\sum_{(i,j)\in\mathcal{L}}q_{ij}(kT)\left[\sum_{m:(i,m)\in\mathcal{L}}\sum_{s=1}^{S}\frac{H_s^{im}(1+\epsilon)\rho_s}{c_{im}}+\sum_{m:(m,i)\in\mathcal{L}}\sum_{s=1}^{S}\frac{H_s^{mi}(1+\epsilon)\rho_s}{c_{mi}}\right. \\
& \qquad\left.+\sum_{h:(j,h)\in\mathcal{L}}\sum_{s=1}^{S}\frac{H_s^{jh}(1+\epsilon)\rho_s}{c_{jh}}+\sum_{h:(h,j)\in\mathcal{L}}\sum_{s=1}^{S}\frac{H_s^{hj}(1+\epsilon)\rho_s}{c_{hj}}-1\right] \\
& -\epsilon\sum_{s=1}^{S}w_s\int_{kT}^{(k+1)T}\mathbf{E}[n_s(t)|\vec{n}(kT),\vec{q}(kT)]dt+E_7 \\
\leq\; & -2T\epsilon\sum_{(i,j)\in\mathcal{L}}q_{ij}(kT)\left[\sum_{m:(i,m)\in\mathcal{L}}\sum_{s=1}^{S}\frac{H_s^{im}\rho_s}{c_{im}}+\sum_{m:(m,i)\in\mathcal{L}}\sum_{s=1}^{S}\frac{H_s^{mi}\rho_s}{c_{mi}}\right. \\
& \qquad\left.+\sum_{h:(j,h)\in\mathcal{L}}\sum_{s=1}^{S}\frac{H_s^{jh}\rho_s}{c_{jh}}+\sum_{h:(h,j)\in\mathcal{L}}\sum_{s=1}^{S}\frac{H_s^{hj}\rho_s}{c_{hj}}\right] \\
& -\epsilon\sum_{s=1}^{S}w_s\int_{kT}^{(k+1)T}\mathbf{E}[n_s(t)|\vec{n}(kT),\vec{q}(kT)]dt+E_7.
\end{aligned}$$

By Theorem 2 of [17], the result then follows. *Q.E.D.*

References

[1] L. Xiao, M. Johansson, and S. Boyd, "Simultaneous Routing and Resource Allocation via Dual Decomposition," in *Proceedings of 4th Asian Control Conference*, Singapore, September 2002, pp. 29–34.

[2] M. Johansson, L. Xiao, and S. Boyd, "Simultaneous Routing and Power Allocation in CDMA Wireless Data Networks," in *Proceedings of IEEE International Conference on Communications*, Anchorage, Alaska, May 2003, pp. 51–55.

[3] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic Power Allocation and Routing for Time Varying Wireless Networks," in *Proceedings of IEEE INFOCOM*, San Francisco, April 2003.

[4] S. Toumpis and A. J. Goldsmith, "Capacity Regions for Wireless Ad Hoc Networks," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 736–748, July 2003.

[5] A. Eryilmaz, R. Srikant, and J. Perkins, "Stable Scheduling Policies for Fading Wireless Channels," *to appear in the IEEE/ACM Transactions on Networking, earlier versions presented at the IEEE International Symposium on Information Theory, 2002, also available at http://comm.csl.uiuc.edu/~srikant/pub.html.*

[6] R. L. Cruz and A. V. Santhanam, "Optimal Routing, Link Scheduling and Power Control in Multi-hop Wireless Networks," in *Proceedings of IEEE INFOCOM*, San Francisco, April 2003.

[7] 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, December 1992.

[8] X. Lin and N. B. Shroff, "Joint Rate Control and Scheduling in Multihop Wireless Networks," in *Proceedings of the IEEE Conference on Decision and Control*, Paradise Island, Bahamas, December 2004.

[9] M. Johansson and L. Xiao, "Scheduling, Routing and Power Allocation for Fairness in Wireless Networks," in *IEEE VTC - Spring*, Milan, Italy, May 2004.

[10] M. Chiang, "To Layer or Not to Layer: Balancing Transport and Physical Layers in Wireless Multihop Networks," in *Proceedings of IEEE INFOCOM*, Hong Kong, March 2004.

[11] S. H. Low and R. Srikant, "A Mathematical Framework for Designing a Low-Loss Low-Delay Internet," *Network and Spatial Economics*, vol. 4, no. 1, pp. 75–102, March 2004.

[12] S. Sarkar and L. Tassiulas, "End-to-end Bandwidth Guarantees Through Fair Local Spectrum Share in Wireless Ad-hoc Networks," in *Proceedings of the IEEE Conference on Decision and Control*, Maui, Hawaii, December 2003.

[13] Y. Yi and S. Shakkottai, "Hop-by-hop Congestion Control over a Wireless Multi-hop Network," in *Proceedings of IEEE INFOCOM*, Hong Kong, March 2004.

[14] Y. Xue, B. Li, and K. Nahrstedt, "Price-based Resource Allocation in Wireless Ad Hoc Networks," in *Proceedings of the Eleventh International Workshop on Quality of Service (IWQoS 2003), also Lecture Notes in Computer Science, ACM Springer-Verlag*, vol. 2707, Monterey, CA, June 2003, pp. 79–96.

[15] T. Bonald and L. Massoulie, "Impact of Fairness on Internet Performance," in *Proceedings of ACM Sigmetrics*, Cambridge, MA, June 2001, pp. 82–91.

[16] N. Z. Shor, *Minimization Methods for Non-Differentiable Functions.* Berlin: Springer-Verlag, 1985.

[17] M. J. Neely, E. Modiano, and C. E. Rohrs, "Power Allocation and Routing in Multibeam Satellites with Time-Varying Channels," *IEEE/ACM Transactions on Networking*, vol. 11, no. 1, pp. 138–152, February 2003.

[18] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity.* Englewood Cliffs, New Jersey: Prentice-Hall, 1982.

[19] E. Leonardi, M. Mellia, F. Neri, and M. A. Marsan, "On the Stability of Input-Queued Switches with Speed-Up," *IEEE/ACM Transactions on Networking*, vol. 9, no. 1, pp. 104–118, February 2001.

[20] J. G. Dai and B. Prabhakar, "The Throughput of Data Switches with and without Speedup," in *Proceedings of IEEE INFOCOM*, Tel Aviv, Israel, March 2000, pp. 556–564.

[21] J. M. Borwein and A. S. Lewis, *Convex Analysis and Nonlinear Optimization: Theory and Examples.* New York: Springer, 2000.