ABSTRACT

Title of dissertation: NETWORK FLOW OPTIMIZATION AND DISTRIBUTED CONTROL ALGORITHMS

Huigang Chen, Doctor of Philosophy, 2006

Dissertation directed by: Professor John S. Baras

This thesis concerns the problem of designing distributed algorithms for achieving efficient and fair bandwidth allocations in a resource constrained network. This problem is fundamental to the design of transmission protocols for communication networks, since the fluid models of popular protocols such as TCP and Proportional Fair Controller can be viewed as distributed algorithms which solve the network flow optimization problems corresponding to some fairness criteria. Because of the convexity of the optimization problem as well as its decoupling structure, there exist classical dual algorithm and primal/dual algorithm which are both distributed. However, the main difficulty is the possible instability of the dynamics of these algorithms caused by transmission delays. We use customized Lyapunov-Krasovskii functionals to obtain the stability conditions for these algorithms in networks with heterogeneous time-varying delays. There are two main features of our results. The first is that these stability conditions can be enforced by a small amount of information exchange among relevant users and links. The second is that these stability conditions only depend on the upper bound of delays, not on the rate of delay

variations. We further our discussion on scalable algorithms with minimum information to maintain stability. We present a design methodology for such algorithms and prove the global stability of our scalable controllers by the use of Zames-Falb multipliers. Next we extend this method to design the first scalable and globally stable algorithm for the joint multipath routing and flow optimization problem. We achieve this by adding additional delays to different paths for all users. Lastly we discuss the joint single path routing and flow optimization problem, which is a NP hard problem. We show bounded price of anarchy for combined flow and routing game for simple networks and show for many-user networks, simple Nash algorithm leads to approximate optimum of the problem.

NETWORK FLOW OPTIMIZATIONS AND DISTRIBUTED CONTROL ALGORITHMS

by

Huigang Chen

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2006

Advisory Committee:
Professor John S. Baras, Chair/Advisor
Professor André L. Tits
Professor Richard J. La
Professor Eyad H. Abed
Professor Carlos A. Berenstein

© Copyright by
Huigang Chen
2006

DEDICATION

To Mom and Dad, and my wife Yun

ACKNOWLEDGMENTS

I am greatly indebted to all the people who have contributed to the completion of this thesis.

First and foremost I would like to express my deepest gratitude and appreciation to my advisor Professor John S. Baras for giving me an invaluable opportunity to work on challenging and extremely interesting problems over the course of my graduate study. Without his guidance and support in all these years, the completion of this thesis would not have been possible. His incredible insight in the field as well as his method of constantly exposing me with new techniques have always been a source of inspiration to my research.

I would also like to thank Professor André Tits, Professor Richard La, Professor Eyad Abed, and Professor Carlos Berenstein for kindly joining my thesis committee despite their busy schedules and for providing valuable feedback.

I extend my gratitude to all my colleagues and friends at Maryland who have helped me in many ways: Tao Jiang, Hui Zeng, Yadong Shang, Senni Perumal, Pedram Hovareshti, Punyaslok Purkayastha, Geogios Papageorgiou, George Theodorakopoulos. I am especially grateful to those who used to be at Maryland and whose presense in my graduate experience benefit me greatly: Dr. Chang Zhang, Dr. Jia-Shiang Chou, Dr. Xiaobo Tan, Dr. Nelson Liu, Dr. Maben Rabi, and Dr. Alvaro Cardenas. I also want to thank the administrative staff of ISR for their assistance during my study and life here. Special thanks are due to Althia Kirlew and Kimberly Edwards.

I want to thank my wife, Yun, for her constant love, support, and sacrifice. I also want to thank my parents for their care and encouragement. This thesis is dedicated to them.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Introduction

One of the most distintive features of the Internet is the fact that it operates in a distributed way on agreed protocols which are not administrated by a central authority. This network "agnostic" feature contributes to its fast growth and easy implementation, and at the same time it poses many practical design challenges such as how to achieve efficient transmission rate allocations across the entire network. The prevalent data transmission protocol, TCP, operates on a purely distributed way. The only network information it depends on is the round-trip delay of its own session. It works well in a low bandwidth, small delay network with little help from the intermediate routers. But with the increase of network bandwidths and round-trip delays, more sophisticated router algorithms are needed to prevent TCP flow instability. These router algorithms are no longer network agnostic and they require network specific tune-ups in order to work. In this thesis our main theme is that we can still achieve the distributedness of the protocols and at the same time reach the efficient allocations and keep the network stable. The responsibility to achieving all these falls mainly onto end-to-end protocols, not router algorithms, although in some sense they are dual to each other. The reason is that there is information asymmetry between the users and the routers, for example, the utility information is only available to the users, and the efficiency is defined on user properties. The main part of the thesis consists of 5 chapters and below is an overview of each chapter.

1.1 Thesis Outline

In Chapter 2 we start with the overview of network traffic models which all the later chapters will be built upon. Network traffic consists of streams of packets transmitted by sources during discrete time steps by some clocking mechanism of their transmission protocols. A direct study of network flow properties by packet models is in general hard and limited to simple networks such as a single source/link network. Just as fluid dynamics in which the basic objects are flows rather than molecules, the use of fluid model in network traffic studies has been widely adopted. Especially, in the area of network protocol design and dynamic study, simple determinstic continuous time fluid models are predominantly used. We will examine in this chapter the "building blocks" of these fluid models under popular network transmission protocols such as TCP and Proportional Fair Controller. Further, we distinguish the window update scheme and the rate update scheme, since TCP protocols use "windows" to control the sending rate. The derivations of these continuous time fluid models from the original discrete models by direct differentiation method and by many-flow asymptotics are described. It is then observed that the fluid models obtained from direct differentiation method are usually "mean-field" approximations of the ones by many-flow asymptotics. In the last section of the chapter we provide a brief overview of network flow optimization problems and we decribe the dynamics of fluid models as algorithms to solve specific optimization problems.

In Chapter 3 we specifically focus on three main types of fluid dynamics related to general network optimization problems, that is, primal algorithm, dual algorithm, and primal/dual algorithm. These algorithms have been discussed extensively in the literature for their stability conditions, either in delay-free network, or in networks with heterogeneous fixed delays. By customized Lyapunov-Krasovskii functionals, we obtained the delay-dependent stability conditions for networks with heterogeneous time varying delays and showed that due to the special structure of these algorithms, our stability conditions do not depend on the rate of changes of the delays. Another nice feature of our stability conditions is that they are relative easy to enforce in a large network and we give implementation guidelines to set the parameters of these algorithms by coordinations with sources and their relevant links. Stability regions in terms of maximally allowed delay bounds are compared between our conditions and the conditions obtained by Small-Gain type methods and we show that our conditions are better. In the end of the chapter we discuss briefly about the global stability property of scalable algorithm whose local stability condition only depends on simple measurements, such as the round-trip delay and the number of bottleneck links en route. We only focus on a single source/link network and show that the global stability condition is not far from the local condition. But to extend this result to a general network is technically difficult.

In Chapter 4 we continue to explore the protocol design problem to solve the network optimization problem. Since we have seen that it is difficult to match the global stability condition to the local stability condition, even though there exist locally scalable distributed algorithms for general networks, we seek to build a scalable algorithm which also has the global stability property. Firstly we set up our design principles which the target algorithms must abide. Under these principles the algorithm, along with its parameters, must be truly distributed and it does not depend on any global variables. We then establish the structural properties of the algorithms with these principles. Guided by these structural properties, we first design a scalable algorithm which meets the design principles locally. Then by relying on Zames-Falb multiplier method, we successfully prove that this algorithm has a scalable global stability property as well.

In Chapter 5 we extend the design methodology to the combined multipath routing and flow control problem. The problem itself is an extension of the regular network flow optimization problem where all the users can now use a pool of available routes for data transmission. The increase of the network capacity can be substantial since it has been shown in the literature that the upper limit for the flow session stability, the minimum cut condition, can be reached by multipath routing. For the optimization point of view the introduction of multipath routing changes the strict convex programming problem into a nonstrict convex programming problem. This change causes the dual programming to become nonsmooth and the original algorithms, based on dual gradient method, fail to achieve rate convergence even in a delay-free network. Based on our scalable algorithm of the last chapter, we build the first stable scalable multipath algorithm for general networks with heterogeneous delays.

In Chapter 6 we consider the problem of the combined single path routing and flow control problem. This problem is drastically different from all the optimization problems discussed before, since now the problem is a nonconvex problem and it has been proved that it is NP-hard to solve. Therefore we shift our focus from the protocol design to the properties of the optimal solution, but still our study has an implication of distributed solution algorithm. We notice that in reality most large networks have far more number of users than that of bottleneck links, and it can be shown that by allowing a proportionally small number of users to use multipath routing, while keeping the rest majority to use single path routing, the resulting solution achieves mulitpath optimality. Therefore it is conceptually plausible that in many-user region local algorithm can achieve solutions arbitrarily close to the optimal solution. To show this is indeed correct, we focus on the solutions brought out by the simplest local algorithm, the Nash algorithm. We first examine a special type of network, which is the one used to prove the NP-hardness of the problem. We showed that the Nash equilibrium exists and the Nash algorithm always converges. It is then shown that the price of anarchy, that is the gap between the worst Nash equilibrium and the social optimal, is bounded when the number of users goes to infinity. For general networks, it is not known whether there exists Nash equilibrium. But we introduce the concept of approximate Nash equilibrium and we show its existence given sufficiently large number of users. Then we prove that approximate Nash equilibrium will be arbitrary close to the social optimal when the number of users is sufficiently large.

Chapter 2

Flow Control Models: From Packet Models to Fluid Models

The basic element of a network flow is a packet. A full description of network flow control models based on packets often involves complex random dynamical structures which are hard to analyze directly in usual cases. Therefore model simplification is a necessary first step for the purpose of studying flow behaviors and designing flow control protocols. Fluid models, which assume flow packets are infinitely divisible, are widely used for its intuitiveness and simplicity. In addition, fluid models can "smooth" out the internal randomness of the flow dynamics in regular settings and give us deterministic model in the end. Therefore we will be only dealing with fluid models in the subsequent chapters. It is then important to understand the assumptions and the consequences of various fluid models as oppose to the real network flows in different flow control protocols. We will focus our discussion on the fluid models of two popular protocols, TCP/Reno and proportional fair controller, and two types updating schemes, packet update and window update. The network setting is a simple single source/link network, although it is straightforward to extend the fluid model to general networks. Also we ignore the effect of random uncontrolled flows, since we do not consider them in this thesis. There are two methods of establishing fluid models, one being direct differentiation and the other many-flows asymptotics. We will present both methods for the derivations.

2.1 Fluid Models by Direct Differentiation

2.1.1 Flow Controlled by TCP/Reno

TCP/Reno is the most widely used TCP congestion control protocol and the core of its congestion avoidance mechanism, additive increase and multiplicative decrease, is used in many other versions of TCP. The source maintains a congestion window during the entire life of the flow and the window size is equal to the number of its packets on the fly in the network. The window is adjusted whenever an acknowledgement (ACK) packet is received from the network for the flow. The ACK packet is sent by the receiver of the flow whenever it receives an packet (or a fixed number of packets). The flow is controlled by the early congestion notification (ECN) bit carried by the flow packet and the ACK packet, which is set by some AQM scheme at intermediate bottleneck routers. The window adjustment scheme in the congestion avoidance phase can be described as

$$W[n+1] = \begin{cases} W[n] + \frac{1}{W[n]}, & \text{when receives ACK without ECN mark,} \\ \frac{W[n]}{2}, & \text{otherwise,} \end{cases} \tag{2.1}$$

where W_n is the window size measured in minimum segment (packet) size at time step n. It should be noted that we only model the congestion avoidance phase here since the stability of the rate of a particular flow makes sense only for long flows and the congestion avoidance phase is predominant in those long flow traffic. It can be observed that the window size increase by 1 for every window of packets sent without ECN bit and decrease by half if there is ECN.

We consider two update schemes. The first is to update the rate immediately when an ACK packet is received, which we call the packet update scheme. The other is to update the rate only after the ACK packets for the whole window are received, which we call the window update scheme. The TCP/Reno protocol uses window update scheme, since the window size for flow control only uses the integer part of $W[n]$ in (2.1) and the flow rate remains constant in the window. Nevertheless fluid models based on packet update scheme is much widely used in the literature due to the simplicity in its final form.

First we derive the fluid model under the packet update scheme. We denote the rate of the flow at time t by $x(t)$. From Little's Law we can approximate the relation between $w(t)$ and $x(t)$ by

$$w(t) = \tau x(t).$$

Therefore we can substitute $w(t)$ by $\tau x(t)$. From fluid approximation, there are $x(t-\tau)\Delta$ feedback packets returning to the sender from time $t-\Delta$ to t, where $x(t-\tau)$ is the sending rate at $t-\tau$ and τ is the fixed round-trip time. The assumption for the round-trip time being fixed is approximately true when the size of the bottleneck buffer is small compared to the delay-throughput product. Among these feedback packets, we assume the dropping probability is p, then there are $px(t-\tau)\Delta$ packets that have been lost. So based on (2.1) the rate $x(t)$ becomes,

$$\tau x(t) = \tau x(t-\Delta) + \frac{1-p}{x(t)\tau} x(t-\tau)\Delta - \frac{p}{2}\tau x(t) x(t-\tau)\Delta.$$

Let $\Delta \to 0$, the above equation converges to the following differential form,

$$\dot{x}(t) = \frac{x(t-\tau)}{x(t)} \left(\frac{1-p}{\tau^2} - \frac{1}{2} x(t)^2 p \right). \tag{2.2}$$

Thus we obtain the basic dynamic equation description of TCP Reno. The queue dynamics with fluid approximation can be expressed as

$$\dot{b}(t) = \left[x(t - \tau^f) - C\right]^+_{b(t)}, \tag{2.3}$$

where $b(t)$ is the bottleneck queue size, τ^f is the forward propagation delay and C is the bandwidth of the bottleneck. The form $[u]^+_v$ takes the value u if $v > 0$ and $\max\{u, 0\}$ if $v = 0$. For a general AQM based on queue size, the marking probability can be described as

$$p(t) = h(b(t - \tau^b)) \tag{2.4}$$

where $h : \mathbb{R}_+ \to [0, 1]$ and τ^b is the backward propagation delay (so $\tau = \tau^f + \tau^b$). Equations (2.2), (2.3), and (2.4) constitute the fluid approximation of a single source/link TCP Reno system. This is a functional differential equation (FDE) model and we will base all of our future discussions upon this kind of models.

For the window update scheme, the source receives the ACKs for the whole window of sent packets every round-trip time τ, since the window size is exactly the number of sent packets that have not been ACKed. Therefore the time instances at which flow rate is updated constitutes a discrete sequence $\{\cdots, t - 2\tau, t - \tau, t, t + \tau, t + 2\tau, \cdots\}$ for some t. The rate process then is a piecewise constance cadlag process, which remains constant between the updating instances. According to the AIMD rule (2.1), the mean rate process $x_t(\cdot)$, with the subscription denoting the time instance to update, can be described as

$$x_t(T) = \begin{cases} x_t(T^-) + \frac{1}{\tau}(1 - p(T)) - \frac{1}{2}x_t(T^-)p_t(T), & T = t + k\tau, \\ x_t(t + k\tau), & t + k\tau < T < t + (k+1)\tau. \end{cases} \tag{2.5}$$

The marking probability $p_t(T)$ is different from the instantaneous marking probability in the packet update scheme. This is the accumulative probability for the flow sent during the time interval $[T - 2\tau, T - \tau)$. Therefore it is a function of round-trip time τ and let us denote it by $p_t(T, \tau)$. From the fluid assumption, the probability of marking is increased by $x_t(T - 2\tau)h(b(T - \tau^b))(1 - p_t(T, \tau))\Delta$ if the round-trip time is increased by Δ. That is

$$p_t(T, \tau + \Delta) = p_t(T, \tau) + x_t(T - 2\tau)h(b(T - \tau^b))(1 - p_t(T, \tau))\Delta.$$

By taking $\Delta \to 0$, we have

$$\frac{\partial p_t(T, \tau)}{\partial \tau} = x_t(T - 2\tau)h(b(T - \tau^b))(1 - p_t(T, \tau)).$$

Notice $x_t(T - 2\tau)$ is constant and $p_t(T, 0) = 0$, we can solve the above equation as

$$p_t(T, \tau) = 1 - \exp\left(-x_t(T - 2\tau)\int_{T-\tau-\tau^b}^{T-\tau^b} h(b(s))ds\right)$$

The mean flow rate process in (2.5) depends on t, the time instance for updating. We can treat flow streams with the same updating sequence as streams in the same class. It is beneficial to improve on this model by considering fluid model for aggregated flow streams of the same source-destination pair and taking into account that the flow streams usually have different updating instances. In addition, the previous single stream model (2.5) can be considered as a special case of the general model. Let there be a continuum of streams with different updating instances, uniformly distributed over the interval $[0, \tau)$. That is, for a fixed time origin 0, there are Δ/τ streams whose updating instance lies within the interval $[t, t+\Delta)$. We can rewrite (2.5) to represent flow rates of streams in different classes as follows,

$$x(t) = x(t-\tau) + \frac{1}{\tau} - \left(\frac{1}{\tau} + \frac{1}{2}x(t-\tau)\right) p(t), \tag{2.6}$$

where the marking probability $p(t)$ is

$$p(t) = 1 - \exp\left(-x(t-2\tau)\int_{t-\tau-\tau^b}^{t-\tau^b} h(b(s))ds\right). \tag{2.7}$$

Note in (2.6) $x(t)$ is equivalent to $x_\theta(t)$, the rate of class θ flow with $t = \theta + k\tau$ for some k. Recall that the flow rate for each class of flows keeps constant between updates, then the aggregate flow rate $\bar{x}(t)$ can be expressed by

$$\bar{x}(t) = \frac{1}{\tau}\int_{t-\tau}^{t} x(s)ds. \tag{2.8}$$

Or in the differential equation form,

$$\dot{\bar{x}}(t) = \frac{1-p(t)}{\tau^2} - \frac{1}{2\tau}x(t-\tau)p(t). \tag{2.9}$$

The single flow case, or equivalently synchronized flows case, can be viewed by setting the initial condition of $x(t)$ in (2.6) by

$$x(t) = \frac{\delta(t)}{\tau}, \quad \forall t < \tau.$$

Or if represented by $\bar{x}(t)$

$$\bar{x}(t) = \begin{cases} 0, & t < 0, \\ \frac{1}{\tau^2}, & t \in [0, \tau). \end{cases}$$

The resulting $\bar{x}(t)$ is again a piecewise constant cadlag process.

As we can see now, although the fluid equation (2.9) for the window update scheme is similar to the equation (2.2) for the packet update scheme, there are two marked differences. First, the fluid rate for the window update scheme takes an explicit form of "averaging" of a window of individual rates, which to some extent resembles the definition of the window size in TCP/Reno protocol. Second, the marking probability which controls the flow rates is accumulative over a round-trip time in the window update scheme, rather than instantaneous as in the packet update scheme. This results in a more complex but more representative fluid model for the TCP/Reno dynamics.

2.1.2 Flow Controlled by Proportional Fairness Controller

The Proportional Fairness (PF) flow control protocol achieves the weighted proportional fair allocation of equilibrium flow rates in a network [1]. The definition of weighted proportional fairness is deferred to the next section. Here we only focus on the fluid model of flow rates controller by the PF protocol. Like in the TCP/Reno case, the sending rate is controlled by ECN bits in the received ACK packets. The source increments the sending rate by the same amount $w\Delta$ every fixed interval Δ and decrease the rate proportional to the number of ACK packets with ECN bit marked.

In the packet update scheme, the rate is updated once there is a new ACK packet coming in. For the purpose of fluid approximation, we assume that the rate increase is $w\Delta$ for any arbitrary small Δ. Again there are $x(t-\tau)\Delta$ feedback ACK packets from time $t-\Delta$ to t, among which there are $x(t-\tau)p(t)\Delta$ packets whose ECN is marked. Denote the proportional factor of rate decrease caused by ECN marking by β. By taking $\Delta \to 0$, we obtain directly the fluid model for PF controlled flow with packet update scheme,

$$\dot{x}(t) = w - \beta x(t-\tau)p(t). \tag{2.10}$$

Here the instantaneous marking probability is the same as that in (2.4) and the bottleneck queue dynamics in (2.3) still applies here.

For the window update scheme, suppose as in TCP/Reno there is a sliding window controlling the packet sending schedule for the PF controller. Then the source only updates its flow rate every round-trip time τ. But it is important to note that TCP/Reno and PF take the feedback signal, viz. ECN marking, in a very different way: TCP/Reno takes the probability of the occurrence of ECN marking while PF takes the number of ECN markings. For the flow stream in class t, the number of ECN markings during $[T-\tau, T)$ can be approximated in the fluid model by

$$x_t(T-2\tau)\int_{T-\tau-\tau^b}^{T-\tau^b} h(b(s))ds.$$

Therefore the rate process of each class of flow stream can be described as (compared to (2.6)

$$x(t) = x(t-\tau) + w\tau - \beta x(t-2\tau)\int_{t-\tau-\tau^b}^{t-\tau^b} h(b(s))ds. \tag{2.11}$$

Using the same idea as in the case of TCP/Reno, we establish the fluid model with rate process $\bar{x}(\cdot)$ (2.8) for the aggregate flow consisted of a continuum of different classes of flows. Its differential equation form can be expressed by

$$\dot{\bar{x}}(t) = w - \beta \frac{x(t-2\tau)}{\tau} \int_{t-\tau-\tau^b}^{t-\tau^b} h(b(s))ds. \tag{2.12}$$

2.2 Fluid Model by Many-Flows Asymptotics

In this section we frequently use the concept of weak convergence [2, 3]. Let B denote a metric space, which can be either a Euclidean space $\mathbb{R}^k$ or a functional space for example continuous function space $C[0,T]$ and cadlag function space $D[0,T]$, and denote $\mathcal{B}$ as the σ-algebra on B induced by this metric. For a sequence of probability measures $\{P_n\}$ defined on $(B,\mathcal{B})$ to converge weakly to another probability measure P defined on $(B,\mathcal{B})$, written as $P_n \Rightarrow P$, the following limit has to hold,

$$\int_B f(x)dP_n(x) \to \int_B f(x)dP(x), \quad n \to \infty$$

for any bounded and continuous real-valued function $f(\cdot)$ on B. Let X_n and X be the random variables (random processes) associated with P_n and P. We can also say X_n converge weakly to X in this case and express this as $X_n \Rightarrow X$. An important concept in proving the weak convergence of probability measures is the tightness [2]. A probability measure sequence $\{P_n\}$ on $(B,\mathcal{B})$ is tight if for every positive ϵ there exists a compact set $K \subset B$ such that $P_n(K) > 1-\epsilon$ for all n. By the Prohorov Theorem, if $\{P_n\}$ on $(B,\mathcal{B})$ is tight, then each subsequence of $\{P_n\}$ contains a further subsequence that converges weakly to some probability measure on $(B,\mathcal{B})$.

2.2.1 Flow Controlled by TCP/Reno

The deduction of fluid models in the previous section is based on the assumption that the network flow is continuous in time and thus infinitely divisible. In this subsection we are interested in a more constructive way to derive fluid models, based on a macroscopic view of discrete flows aggregation. In other words we offer a more careful argument for a continuous time fluid model of a single bottleneck TCP/AQM system without assuming flow continuity *a priori*, in which the random nature of packet marking at the bottleneck queue is considered. We try to establish the link between the many-flows region asymptotics of TCP/AQM system and deterministic FDE model. This is similar to the works by [4, 5, 6, 7]. In [4] an ODE limit is obtained through a stochastic approximation based model while the round-trip delay is unaccounted for in the final ODE, while in [5] only deterministic queue based marking is considered. More accurate models for TCP/AQM are considered in [6] and [7] but in their models rates are updated synchronously among all sources. Therefore their final asymptotic fluid models are discrete at update intervals.

In all the flow control schemes studied in this section we consider a sequence of systems indexed by N, in which the Nth system consists of N identical flows accessing a common bottleneck link. The link capacity in the Nth system is scaled as NC packets per second. The round trip delays for all the systems are the same τ, with the forward delay τ^f and the backward delay τ^b. Without loss of generality we assume they are all integers. This can be seen as N replications of the single flow/link TCP/AQM system. We consider a slotted time system and the time interval for each slot is $1/N$ seconds for the Nth system. For any process $z^{(N)}[\cdot]$ considered in the system, its continuous time version is piecewise constant and defined as

$$z^{(N)}(t) \triangleq z^{(N)}[\lfloor Nt \rfloor].$$

We denote the rate of flow i, the aggregated number of packets received at bottleneck, and the bottleneck queue size of the Nth system by $x_i^{(N)}$, $y^{(N)}$, and $\tilde{b}^{(N)}$, respectively. The packet model for each flow i is as follows. At each time slot l let $\{0,1\}$-valued series $\tilde{x}_i^{(N)}[l]$ denote whether there is a packet sent by flow i, $\tilde{x}_i^{(N)}[l] = 1$, or not, $\tilde{x}_i^{(N)}[l] = 0$. Again we use a model of probabilistic sending scheme, that is,

$$\tilde{x}_i^{(N)}[l] = 1(\eta_{i,l}^{(N)} \leq x_i^{(N)}[l]/N),$$

where $\eta_{i,l}^{(N)}$ is an independent random variable distributed uniformly over $[0,1]$. As usual we use the notation 1_A as indication function of the occurrence of event A. This modeling assumption is made purely from technical reasons. Although a real TCP flow sends packets in a uniform fashion, the assumption of probabilistic sending nevertheless maintains the same mean flow rate and more importantly it greatly simplifies the subsequent analysis, since the packet arrival process at the bottleneck queue can be regarded as a Poisson process with varying rate.

For the packet update scheme, the rate is updated every time an ACK of the previously sent packet is received at the source. That is at time slot $l+1$, if there is an arrival ACK without its ECN marked, the rate is incremented by $1/(x_i^{(N)}[l]\tau^2)$. Otherwise, the rate reduces to its half. Put these together we have for individual flow the updating rule,

$$x_i^{(N)}[l+1] = x_i^{(N)}[l] + \left(\frac{1 - 1(\xi_{i,l}^{(N)} < p^{(N)}[l])}{x_i^{(N)}[l]\tau^2} - \frac{1}{2} x_i^{(N)}[l] 1(\xi_{i,l}^{(N)} < p^{(N)}[l]) \right) \tilde{x}_i^{(N)}[l - N\tau].$$

Here similar to $\eta_{i,l}^{(N)}$, the random variable $\xi_{i,k}^{(N)}$ is also independent and uniform distributed over $[0,1]$. The value $p^{(N)}[l]$ is the packet marking probability experienced by the packet associated with the ACK received at time l, which can be expressed by as in (2.4)

$$p^{(N)}[l] = h^{(N)}(\tilde{b}^{(N)}[l - N\tau^b]),$$

where the marking function $h^{(N)} : \mathbb{R}_+ \to [0,1]$ of the Nth system satisfies the following scaling

$$h^{(N)}(\tilde{b}^{(N)}) = h(\tilde{b}^{(N)}/N).$$

By the definition the aggregated number of arrival packets at the bottleneck queue is

$$y^{(N)}[l] = \sum_{i=1}^{N} \tilde{x}_i^{(N)}[l - N\tau^f].$$

So the queue dynamics can be expressed by

$$\tilde{b}^{(N)}[l] = \tilde{b}^{(N)}[l-1] + (y^{(N)}[l] - C)^+_{\tilde{b}^{(N)}[l-1]}.$$

The many-flow asymptotics is concerned with the following processes $\bar{x}^{(N)}[\cdot]$ and $\bar{b}^{(N)}[\cdot]$, which are defined as

$$\bar{x}^{(N)}[l] \triangleq \frac{1}{N}\sum_{i=1}^{N} x_i^{(N)}[l],$$

$$\bar{b}^{(N)}[l] \triangleq \frac{1}{N}\tilde{b}^{(N)}.$$

Then it follows from the updating rule and queue dynamics,

$$\begin{aligned}
\bar{x}^{(N)}[l+1] &= \bar{x}^{(N)}[l] + \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1-1(\xi_{i,l}^{(N)} < p^{(N)}[l])}{x_i^{(N)}[l]\tau^2} - \frac{1}{2}x_i^{(N)}[l]1(\xi_{i,l}^{(N)} < p^{(N)}[l])\right) \\
&\quad \times \tilde{x}_i^{(N)}[l-N\tau] \\
&\triangleq \bar{x}^{(N)}[l] + \frac{1}{N}G^{(N)}[l], \\
\bar{b}^{(N)}[l] &= \bar{b}^{(N)}[l-1] + \frac{1}{N}\left(\sum_{i=1}^{N}\tilde{x}_i^{(N)}[l-N\tau^f] - C\right)^{+}_{\bar{b}^{(N)}[l-1]} \\
&\triangleq \bar{b}^{(N)}[l-1] + \frac{1}{N}F^{(N)}[l].
\end{aligned}$$

Without loss of generality, we assume that at time slot 0 $x_i^{(N)}[l] \equiv x_0$, $-N\tau \leq l \leq 0$ for all N and $i = 1, \cdots, N$. To show the weak convergence of the above discrete time stochastic system to a particular continuous time differential system, the first step is to show that the random sequences $\{x_i^{(N)}[\cdot]\}$ and $\{\bar{b}^{(N)}[\cdot]\}$ are tight. From its discrete time dynamics, it is obvious that the solution value of $x_i^{(N)}[\cdot]$ is positive. We use the similar truncation technique as in [3] section 2.3 such that we may assume $x_i^{(N)}[\cdot] \in [1/K, K]$ for some arbitrary large K in the subsequent deduction. Denote $\mathcal{F}_l^{(N)}$ the σ-algebra measurable by $\{x_i^{(N)}[k], \bar{b}^{(N)}[k], 0 \leq k \leq l\}$. By the update rule, we have

$$P^{(N)}\{|x_i^{(N)}[l+1] - x_i^{(N)}[l]| \neq 0 | \mathcal{F}_l^{(N)}\} = \frac{1}{N}x_i^{(N)}[l-N\tau]$$

where $P^{(N)}$ is the probability measure of the Nth system. In addition by the truncation assumption each rate jump is uniformly bounded. Therefore, by the use of Theorem 15.2 in [2], we reach the conclusion that $\{x_i^{(N)}[\cdot]\}$ and $\{\bar{b}^{(N)}[\cdot]\}$ are tight in $D[0,\infty)$. By the Prohorov Theorem, it is then sufficient to work with an arbitrary weakly convergent subsequence and without loss of generality we also index this subsequence by N. That is, we can suppose that $\{x_i^{(N)}(\cdot), \bar{b}^{(N)}(\cdot)\} \Rightarrow \{x_i(\cdot), \bar{b}(\cdot)\}$.

We will use the martingale method [3] to derive the asymptotic limit. The idea is to get the infinitesimal operator A for the continuous time limit process from averaging and then obtain the limit process as the solution to the martingale problem with operator A. The limit process for $\bar{b}^{(N)}(\cdot)$ can be obtained directly from the fluid limit result by [8]

$$\dot{\bar{b}}(t) = (\bar{x}(t - \tau^f) - C)^+_{\bar{b}(t)}. \tag{2.13}$$

Consequently, the continuous time version of marking probability $p(t)$ can be expressed as

$$p(t) = h(\bar{b}(t - \tau^b)). \tag{2.14}$$

Let us now consider the limit process $x_i(\cdot)$. For each i, consider any bounded continuously differentiable function $f(\cdot)$, one has

$$\begin{aligned} & E^{(N)}[f(x_i^{(N)}[l+1]|\mathcal{F}_l^{(N)}] - f(x_i^{(N)}[l]) \\ =& \frac{1}{N} x_i^{(N)}[l - N\tau] \left(\begin{array}{l} (1 - p^{(N)}[l]) \left(f\left(x_i^{(N)}[l] + (x_i^{(N)}[l]\tau^2)^{-1} \right) - f(x_i^{(N)}[l]) \right) \\ + p^{(N)}[l] \left(f\left(x_i^{(N)}[l]/2 \right) - f(x_i^{(N)}[l]) \right) \end{array} \right) \\ & + o(N^{-1}). \end{aligned}$$

Therefore for arbitrary k, t, s, and $s_1 < \cdots < s_k < t < t + s$ and any bounded and continuous function $g(\cdot)$,

$$
\begin{aligned}
&Eg(x_i^{(N)}(s_j), j \le k) \\
&\times \left[\begin{array}{l} f(x_i^{(N)}(t+s)) - f(x_i^{(N)}(t)) - \frac{1}{N}\sum_{l=tN}^{(t+s)N} x_i^{(N)}[l-N\tau] \\ \times \left(\begin{array}{l} (1-p^{(N)}[l])\left(f\left(x_i^{(N)}[l] + (x_i^{(N)}[l]\tau^2)^{-1}\right) - f(x_i^{(N)}[l])\right) \\ +p^{(N)}[l]\left(f\left(x_i^{(N)}[l]/2\right) - f(x_i^{(N)}[l])\right) \end{array} \right) \end{array} \right] \\
=&o(N^{-1}).
\end{aligned}
$$

Since $\bar{b}^{(N)}(\cdot) \Rightarrow \bar{b}(\cdot)$, $p^{(N)}(\cdot) \Rightarrow p(\cdot)$, then one has the following

$$
\begin{aligned}
&Eg(x_i(s_j), j \le k) \\
&\times \left[\begin{array}{l} f(x_i(t+s)) - f(x_i(t)) \\ -\int_t^{t+s} x_i(r-\tau) \left(\begin{array}{l} (1-p(r))\left(f\left(x_i(r) + (x_i(r)\tau^2)^{-1}\right) - f(x_i(r))\right) \\ +p(r)\left(f(x_i(r)/2) - f(x_i(r))\right) \end{array} \right) dr \end{array} \right] \\
=&0.
\end{aligned}
$$

Define the operator A by

$$
Af(x(t)) = x(t-\tau) \left(\begin{array}{l} (1-p(t))\left(f\left(x(t) + (x(t)\tau^2)^{-1}\right) - f(x(t))\right) \\ +p(t)\left(f(x(t)/2) - f(x(t))\right) \end{array} \right),
$$

then by arbitrariness of $f(\cdot)$, $g(\cdot)$, s_i, s, and t we conclude that $x_i(\cdot)$ solves the martingale problem associated with operator A with initial condition $x_i(t) = x_0$, $-\tau \le t \le 0$. It is easy to see that the martingale problem has unique solution. Since we already know that any subsequence of $\{x_i^{(N)}(\cdot)\}$ contains a further subsequence that weakly converges, it follows that $x_i^{(N)}(\cdot) \Rightarrow x_i(\cdot)$ uniquely, where

$$
dx_i(t) = \left(\frac{1 - 1(\xi(t) < p(t))}{x_i(t)\tau^2} - \frac{1}{2}x_i(t)1(\xi(t) < p(t)) \right) dN_{x_i(t-\tau)}(t) \tag{2.15}
$$

where $N_\lambda(t)$ is a Poisson process with rate λ and $\xi(t)$ is an independent random process whose marginal distribution is uniform over $[0,1]$. Notice that the equation (2.15) does not depend on particular i, therefore it is the limit process of all individual flows.

Lastly from the Law of Large Numbers, we obtain

$$G(t) = E_t\left[\left(\frac{1-p(t)}{x(t)\tau^2} - \frac{1}{2}x(t)p(t)\right)x(t-\tau)\right] \tag{2.16}$$

so the limit process $\bar{x}(\cdot)$ solves the following equation

$$\dot{\bar{x}}(t) = G(t) \tag{2.17}$$

or equivalently $\bar{x}(t) = E_t[x(t)]$.

Now let us consider the window update scheme. Recall that we derive the fluid model from taking the aggregate flow from a continuum of flow classes in the previous section. Here we model a finite set of flows with different update times for each system N. Although in reality each flow in the system can be started at random, for the ease of exposition, we suppose that in the Nth system the flow i starts at $i\tau/N$ seconds so all the flows are evenly distributed. From the previous description, each flow only updates its rate every round-trip time τ. For notational simplicity we denote the kth update for flow i takes place at time step $t_i^{(N)}[k] \triangleq (i+kN)\tau$. Therefore we have for individual flow the updating rule,

$$x_i^{(N)}[l] = \begin{cases} \quad x_i^{(N)}[l-N\tau] + \frac{1}{\tau}1(\xi_{i,k}^{(N)} > p_i^{(N)}[k]) & \\ -\frac{1}{2}x_i^{(N)}[l-N\tau]1(\xi_{i,k}^{(N)} \le p_i^{(N)}[k]), & l = t_i^{(N)}[k], \\ x_i^{(N)}[l-1], & \text{otherwise}, \end{cases}$$

where $\xi_{i,k}^{(N)}$ is an independent random variable which is uniformly distributed over $[0,1]$, and $p_i^{(N)}[k]$ is the packets marking probability experienced by flow i for the packets sent during the time steps from $t_i^{(N)}[k-2]$ to $t_i^{(N)}[k-1]-1$, similar to $\xi_{i,k}^{(N)}$,

The packet marking event $d^{(N)}[l]$ can be defined as

$$d^{(N)}[l] = 1(\zeta_l^{(N)} \leq h(\tilde{b}^{(N)}[l]/N))$$

where $\zeta_l^{(N)}$ is a $[0,1]$ uniformly distributed random variable, similar to $\xi_{i,k}^{(N)}$ and $\eta_{i,l}^{(N)}$. Therefore the packets marking probability $p_i^{(N)}[k]$ can then be expressed as

$$\begin{aligned} p_i^{(N)}[k] &= \Pr\left(\bigcup_{j=0}^{N\tau-1}\left\{\tilde{x}_i^{(N)}[t_i^{(N)}[k-2]+j]d^{(N)}[t_i^{(N)}[k-2]+N\tau^f+j]=1\right\}\right) \\ &= 1-\Pr\left(\bigcap_{j=0}^{N\tau-1}\left\{\tilde{x}_i^{(N)}[t_i^{(N)}[k-2]+j]d^{(N)}[t_i^{(N)}[k-2]+N\tau^f+j]=0\right\}\right) \end{aligned}$$

From the queueing dynamics the packet sent event at time l $\tilde{x}_i^{(N)}[l]$ affects the bottleneck queue length $\tilde{b}^{(N)}[l+N\tau^f+j]$ for $j>0$. Thus the marking event $d^{(N)}[l+N\tau^f+j]$ is not independent from the packet sent event $\tilde{x}_i^{(N)}[l]$, which makes the above probability hard to compute. But thanks to the scaling of the marking function $h^{(N)}$, we can have "approximate" independence of these two random sequences for large N. More precisely, denote

$$\tilde{x}_{i,k}^{(N)} \triangleq \sum_{j=0}^{N\tau-1} \tilde{x}_i^{(N)}[t_i^{(N)}[k]+j]$$

which is a binomially distributed random variable and it converges weakly to an exponentially distributed random variable $\tilde{x}_{i,k}$. Consider now a sequence of marking event $\tilde{d}^{(N)}[\cdot]$ which is independent of $\tilde{x}_i[\cdot]$ and the resulting packets marking probability $\tilde{p}_i^{(N)}$. Since the length of the bottleneck queue can be varied due to the randomness of $\tilde{x}_i^{(N)}[l]$ by at most $\tilde{x}_{i,k}^{(N)}$, we have

$$
\begin{aligned}
&|p_i^{(N)}[k] - \tilde{p}_i^{(N)}[k]| \\
\leq & \left| \begin{array}{l} \prod_{j=0}^{N\tau-1} \left(1 - \Pr(\tilde{x}_i^{(N)}[t_i^{(N)}[k-2]+j]=0)\Pr(\tilde{d}^{(N)}[t_i^{(N)}[k-2]+N\tau^f+j]=0)\right) \\ -\prod_{j=0}^{N\tau-1} \left(1 - \Pr(\tilde{x}_i^{(N)}[t_i^{(N)}[k-2]+j]=0)\Pr(\zeta_{(\cdot)}{}^{(N)} \geq h((\tilde{b}^{(N)}(\cdot)+\tilde{x}_{i,k}^{(N)})/N)\right) \end{array} \right| \\
\leq & \left| \begin{array}{l} \prod_{j=0}^{N\tau-1} \left(1 - \frac{1}{N} x_i^{(N)}[t_i^{(N)}[k-2]] h(\tilde{b}^{(N)}[t_i^{(N)}[k-2]+N\tau^f+j]/N)\right) \\ -\prod_{j=0}^{N\tau-1} \left(1 - \frac{1}{N} x_i^{(N)}[t_i^{(N)}[k-2]] (h(\tilde{b}^{(N)}[t_i^{(N)}[k-2]+N\tau^f+j]/N) + h'\tilde{x}_{i,k}^{(N)}/N)\right) \end{array} \right|.
\end{aligned}
$$

From the boundedness of $p_i^{(N)}[k]$ it is easy to see that $p_i^{(N)}[k] \to \tilde{p}_i^{(N)}[k]$ almost surely. So hereafter we only have to compute the limit of $\tilde{p}_i^{(N)}[k]$.

Let us define a sequence of functionals $H_N : D[0,\tau] \to \mathbb{R}$ by

$$H_N(z) \triangleq \prod_{j=0}^{N\tau-1} \left(1 - \frac{1}{N} \int_{\frac{j\tau}{N}}^{\frac{(j+1)\tau}{N}} z(s)ds\right).$$

So $\tilde{p}_i^{(N)}[k] = 1 - H_N(x_i^{(N)}[k-2]h(\tilde{b}^{(N)}(\cdot)))$. It is known that for any bounded $z(\cdot)$, $H_N(z)$ converges uniformly to

$$H(z) \triangleq \exp\left(-\int_0^\tau z(s)ds\right).$$

Then by the Dominated Convergence Theorem, the following limit holds,

$$E\left[H_N(x_i^{(N)}[k-2]h(\tilde{b}^{(N)}(\cdot)/N)) - H(x_i^{(N)}[k-2]h(\tilde{b}^{(N)}(\cdot)/N))\right] \to 0.$$

Next suppose that the weak convergence $\tilde{b}^{(N)}(t)/N \Rightarrow \tilde{b}(t)$ and $x_i^{(N)}[t_i^{(N)}[k-2]] \Rightarrow x_i[k-2]$ hold for some random process $\tilde{b}(t)$ and random variable x_i. Then by Theorem 5.5 of [2] and the boundedness of H, $H(x_i^{(N)}[k-2]h(\tilde{b}^{(N)}/N)) \Rightarrow H(x_i[k-2]h(\tilde{b}))$, and consequently $E[H(x_i^{(N)}[k-2]h(\tilde{b}^{(N)}/N))] \to E[H(x_i[k-2]h(\tilde{b}))]$. Putting all these together, we have

$$p_i^{(N)}[k] \Rightarrow 1 - \exp\left(-x_i[k-2]\int_{(k-1)\tau-\tau^b}^{k\tau-\tau^b} h(\tilde{b}(s))ds\right), \quad N \to \infty.$$

Same as in the packet update scheme let us consider the process $x^{(N)}$ and $b^{(N)}$ which are defined as

$$\begin{aligned}\bar{x}^{(N)}[l] &= \tfrac{1}{N}\textstyle\sum_{i=1}^{N} x_i^{(N)}[l],\\ \bar{b}^{(N)}[l] &= \tfrac{1}{N}\tilde{b}^{(N)}.\end{aligned}$$

It is from the definition that

$$\begin{aligned}\bar{x}^{(N)}[t_i^{(N)}[k]] &= \bar{x}^{(N)}[t_{i-1}^{(N)}[k]]\\ &\quad +\frac{1}{N}\left(\frac{1}{\tau}1(\xi_{i,k}^{(N)} > p_i^{(N)}[k]) - \frac{1}{2}x_i^{(N)}[t_i^{(N)}[k-1]]1(\xi_{i,k}^{(N)} \le p_i^{(N)}[k])\right)\\ &\triangleq \bar{x}^{(N)}[t_{i-1}^{(N)}[k]] + \frac{1}{N}G^{(N)}[t_i^{(N)}[k]],\\ \bar{b}^{(N)}[l] &= \bar{b}^{(N)}[l-1] + \frac{1}{N}\left(\sum_{i=1}^{N} 1(\eta_{i,l}^{(N)} \le x_i^{(N)}[l - N\tau^f]/N) - C\right)^{+}_{\bar{b}^{(N)}[l-1]}\\ &\triangleq \bar{b}^{(N)}[l-1] + \frac{1}{N}F^{(N)}[l].\end{aligned}$$

With the same reasoning as in the packet update scheme we can suppose that $\{x_i^{(N)}(\cdot), \bar{b}^{(N)}(\cdot)\} \Rightarrow \{x_i(\cdot), \bar{b}(\cdot)\}$ and the situation is simpler than the case of packet update scheme. It is straightforward to show that the limit process $x_i(\cdot)$ can be represented by the random process $x(t)$, which solves the following stochastic functional difference equation:

$$x(t) = x(t-\tau) + \tfrac{1}{\tau}1(\xi(t) > p(t)) - \tfrac{1}{2}x(t-\tau)1(\xi(t) \le p(t)). \tag{2.18}$$

with the initial condition

$$x(\theta) = \frac{1}{\tau}, \quad 0 \le \theta < \tau.$$

Here $\xi(t)$ is a time independent random variable, uniformly distributed on $[0, 1]$.

The continuous time version of dropping probability $p(t)$ is defined as

$$p(t) = 1 - \exp\left(-x(t-2\tau)\int_{t-\tau-\tau^b}^{t-\tau^b} h(\bar{b}(s))\right), \tag{2.19}$$

Then by the Law of Large Numbers, we obtain

$$G(t) = \frac{1}{\tau} - E_t\left[\left(\frac{1}{\tau} + \frac{1}{2}x(t-\tau)\right)p(t)\right]. \tag{2.20}$$

Therefore it can be shown that the limit process $\bar{x}(\cdot)$ solves the following equation

$$\dot{\bar{x}}(t) = \frac{1}{\tau}G(t), \tag{2.21}$$

and the limit rate $\bar{x}(\cdot)$ is

$$\bar{x}(t) = \frac{1}{\tau}E_t\left[\int_{t-\tau}^{t} x(s)ds\right].$$

2.2.2 Flow Controlled by Proportional Fairness Controller

The many-flows asymptotics of the flows controlled by PF controller can be derived by the same techniques used in the previous subsection. Here we present the discrete models for both update schemes and associated limit fluid models directly. As in the case of TCP/Reno systems, we consider a sequence of single source/link systems, each with N flows equipped with probabilistic sending scheme. The time for the Nth system is again slotted with the duration of $1/N$ seconds. We use the same notations for the (limit) rate process, (limit) queue process, and etc. as the last subsection.

In the packet update scheme for PF controlled flows, the rate updating rule can be expressed by

$$x_i^{(N)}[l+1] = x_i^{(N)}[l] + \frac{1}{N}w - \beta\tilde{x}_i^{(N)}[l-N\tau]1(\xi_{i,l}^{(N)} < p^{(N)}[l]),$$

where the marking probability is

$$p^{(N)}[l] = h(\tilde{b}^{(N)}[l - N\tau^b]/N).$$

Then in the many-flows limit we have the same queue dynamics as in (2.13). The flow limit satisfies the following equation

$$\dot{\bar{x}}(t) = w - \beta\bar{x}(t-\tau)h(\bar{b}(t-\tau^b)). \tag{2.22}$$

For PF controlled flows with the window update scheme, the rate update rule follows the representation below,

$$x_i^{(N)}[l] = \begin{cases} \quad x_i^{(N)}[l-N\tau] + w\tau & \\ -\frac{1}{\tau}\sum_{j=1}^{N\tau}\tilde{x}_i^{(N)}[l-2N\tau+j]d^{(N)}[l-N(\tau+\tau^b)+j], & l = t_i^{(N)}[k], \\ x_i^{(N)}[l-1], & \text{otherwise,} \end{cases}$$

Again using the same methodology, one concludes that the limit process $\bar{x}(\cdot)$ solves the differential equation (2.21) with a different $G(t)$ as below,

$$G(t) = w\tau - \frac{1}{2}E_t[x(t-2\tau)]\int_{t-\tau-\tau^b}^{t-\tau^b} h(\bar{b}(s))ds. \tag{2.23}$$

The result of this section can be summarized in Table 2.1. As we can see, a general rule of thumb is that the fluid models obtained through direct differentiation with the assumption of deterministic continuous flow are in the sense of "mean field approximation" of the fluid models obtained by many-flows asymptotics, which takes the randomness of packet traffic and packet marking into account. That is, the simplifications of fluid model dynamics by replacing $E[f(x)]$ by $f(E(x))$ and treating $x(t)$ independent of $x(t-\tau)$ are made. Therefore for all the protocols and update schemes we considered in this section, only PF flow with packet update scheme has the same fluid model representation by both direct differentiation and many-flows asymptotics method, since in this case the flow dynamics is linear. Although fluid models by many-flows asymptotics are precise, studying their dynamical properties, or even calculating their equilibrium points, is a complicated task since their fluid dynamics explicitly rely on (high) moments of state variables evolved by some stochastic equations (e.g. (2.15), (2.18), etc.). Hence in the subsequent discussions, we follow a common practice that focuses entirely on the fluid models based on the direct differentiation method. This approach has been used successfully on study of the equilibrium and dynamical properties of congestion controlled networks and has been shown to describe the system very accurately [9, 10, 11, 12, 13]. We run simulations for TCP/Reno flows with the window update scheme for 200 flows and 2000 flows in a single source/link network with proportional marking function $h(b) = 0.02q$. The paths of their corresponding mean flow rates with comparison to what is predicted by the direct differentiation method is presented in Figure 2.1. As we can see, in this case the fluid model by the direct differentiation method is a close match to the real model.

2.3 Network of Congestion Controlled Flows and Flow Optimization

The fluid model of the previous section can be readily extended to the network case. We consider a N-user L-bottleneck-link network (Fig. 2.2). By bottleneck links we mean those links whose aggregate traffic rates are equal to their bandwidths

Table 2.1: Fluid Models for TCP/Reno and Proportional Fairness Controller

	TCP/Reno		PF Controller	
	Packet Update	Window Update	Packet Update	Window Update
Direct Differentiation	(2.2), (2.3), (2.4)	(2.3), (2.6), (2.7), (2.9)	(2.3), (2.4), (2.10)	(2.3), (2.11), (2.12)
Many-Flows Asymptotics	(2.13), (2.15), (2.14), (2.16), (2.17)	(2.13), (2.18), (2.19), (2.20), (2.21)	(2.13), (2.22)	(2.13), (2.21), (2.23)

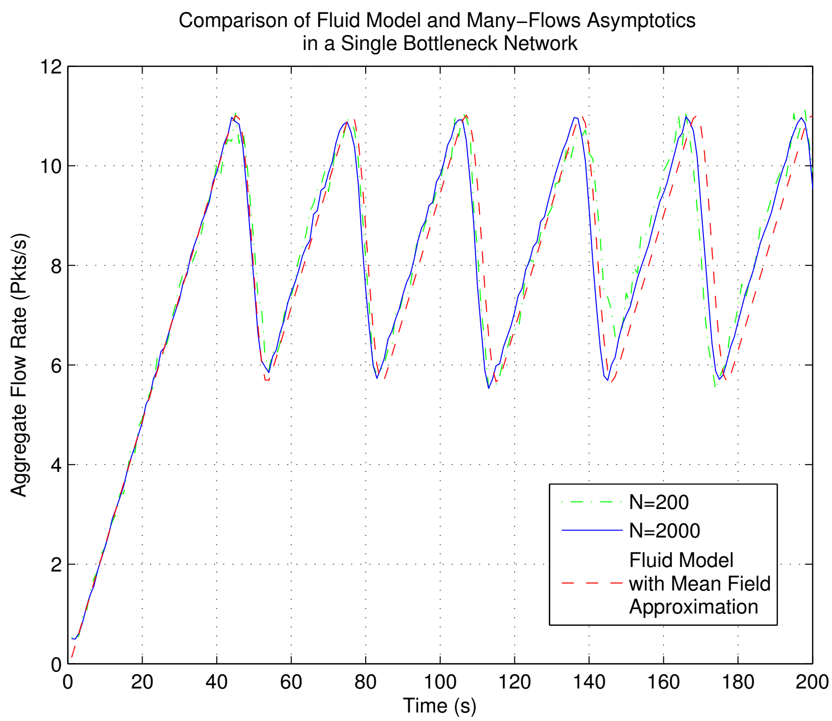

Figure 2.1: TCP/Reno flow rates in a single source/link network with proportional marking. The update scheme is window based. The round-trip delay is 2 seconds and the link bandwidth is 10 packets/second. Plotted are simulated paths with 200 flows, 2000 flows, and the fluid model derived from direct differential method (see equations (2.3), (2.6), (2.7), (2.8), (2.9))

in equilibrium. We suppose that the bottleneck links are known *a priori* and all other non-bottleneck links are "transparent" to user traffics. So from now on we refer to bottleneck links simply as "links". Each user has a single fixed path, consisting of one or more links, to send traffic. If a user has multiple sessions of traffic simultaneously then we "split" the user into multiple users. So we hereafter use user, flow and traffic on this single path interchangeably (notice flow used in this paper is different from that in multicommodity-flow problem). Each flow is indexed by a number in $[N] \triangleq \{1, \ldots, N\}$ and each link is indexed by a number in $[L] \triangleq \{1, \ldots, L\}$. We use a $L \times N$ 0-1 routing matrix R to describe this flow/link relationship. That is $R_{ji} = 1$ if the flow i passes the link j and $R_{ji} = 0$ otherwise. We denote the bandwidth of link j by $c_j > 0$.

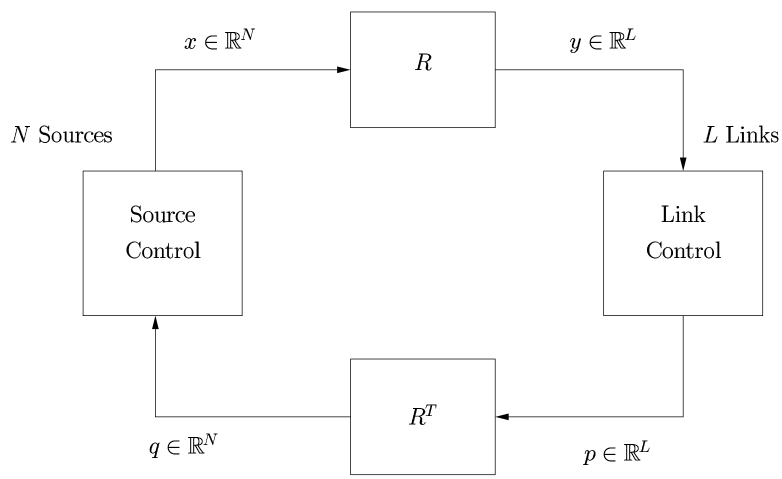

Figure 2.2: Network of elastic traffic

Assuming all the flows follow TCP/Reno protocol with the packet update scheme, we can write the flow dynamics according to (2.2), (2.3), and (2.4) as below,

$$\begin{cases} \dot{x}_i(t) = \frac{x_i(t-\tau_i)}{x_i(t)} \left(\frac{1}{\tau_i^2} - \left(\frac{1}{\tau_i^2} + \frac{1}{2} x_i(t)^2 \right) q_i(t) \right) \\ \dot{b}_j(t) = (y_j(t) - c_j)^+_{b_j}. \end{cases}$$

Here the aggregated rate $y_j(\cdot)$ and the aggregate marking probability $q_i(\cdot)$ can be expressed as

$$y_j(t) = \sum_{i=1}^{N} R_{ji} x_i(t - \tau_{ij}^f),$$
$$q_i(t) = \sum_{j=1}^{N} R_{ji} h_j(b_j(t - \tau_{ij}^b)).$$

Similarly, for the TCP/Reno flows with the window update scheme, their dynamics (see (2.3), (2.6), (2.7), and (2.9)) can be expressed as,

$$\begin{cases} x_i(t) = x_i(t-\tau_i) + \frac{1}{\tau_i} - \left(\frac{1}{\tau_i} + \frac{1}{2}x_i(t-\tau_i)\right) q_i(t), \\ \dot{b}_j(t) = (y_j(t) - c_j)_{b_j}^+. \end{cases}$$

Different than those in the packet update scheme, the aggregate rate and marking probability are of the following form,

$$y_j(t) = \sum_{i=1}^{N} R_{ji} \frac{1}{\tau_i} \int_{t-\tau_i-\tau_{ij}^f}^{t-\tau_{ij}^f} x_i(s) ds,$$
$$q_i(t) = 1 - \exp\left(-x_i(t-2\tau_i) \sum_{j=1}^{N} R_{ji} \int_{t-\tau_i-\tau_{ij}^b}^{t-\tau_{ij}^b} h_j(b_j(s)) ds\right).$$

It is worth to remark that the aggregate marking probability for an individual packet is a multiplication of the marking probabilities exercised by the intermediate bottleneck links which the packet passes, assuming the marking events are independent from each other. But in the fluid model we only concern events occurred in infinitesimal durations. Therefore the aggregate marking probability appears in an additive form. This is precise to the extent of the mean field approximation and no assumption of “small” marking probability is needed for this to be valid.

We are interested in the equilibrium rate of the network flow system. From the network dynamic equations, the equilibrium rates x_i^* and queue lengths b_j^* satisfies the following equations in the packet update scheme:

$$\begin{cases} \left(\frac{1}{\tau_i^2}+\frac{1}{2}x_i^{*2}\right)\sum_{j=1}^N R_{ji}h_j(b_j^*)=\frac{1}{\tau_i^2}, \\ \sum_{i=1}^N R_{ji}x_i^* = c_j, \forall j \text{ s.t. } b_j^*>0. \end{cases}$$

and in the window update scheme:

$$\begin{cases} \left(\frac{1}{\tau_i}+\frac{1}{2}x_i^{*}\right)\left(1-\exp\left(-\tau_i\bar{x}_i\sum_{j=1}^N R_{ji}h_j(b_j^*)\right)\right)=\frac{1}{\tau_i}, \\ \sum_{i=1}^N R_{ji}x_i^* = c_j, \forall j \text{ s.t. } b_j^*>0. \end{cases}$$

Then the equilibrium rates for both update schemes depend on the aggregated feedback value of marking probabilities in the follow way,

$$\text{Packet Update Scheme:} \quad \frac{1}{1+\frac{1}{2}\tau_i^2 x_i^{*2}} \triangleq F_1(\bar{x}_i)=\sum_{j=1}^N R_{ji}h_j(b_j^*)$$

$$\text{Window Update Scheme:} \quad \frac{1}{\tau_i x_i^*}\ln\left(\frac{2}{\tau\bar{x}_i}+1\right) \triangleq F_2(\bar{x}_i)=\sum_{j=1}^N R_{ji}h_j(b_j^*).$$

In order to characterize the above equilibrium rates, consider the utility functions defined as follows,

$$\text{Packet Update Scheme:} \quad U_i(x) \triangleq \int F_1(x)dx = \frac{\sqrt{2}}{\tau_i}\arctan\frac{\tau_i x}{\sqrt{2}}$$

$$\text{Window Update Scheme:} \quad U_i(x) \triangleq \int F_2(x)dx = \frac{1}{\tau_i}\mathrm{Li}_2\left(-\frac{2}{\tau_i x}\right)$$

where the function $\mathrm{Li}_2(x) : (\infty, 1] \to \mathbb{R}$ is dilogarithm function which is defined as $\mathrm{Li}_2(x) \triangleq \sum_{k=1}^{\infty} x^k/k^2$. It can be verified immediately that the utility functions $U_i(\cdot)$ in both update schemes are strictly monotonically increasing concave functions. Kelly formulated the network flow optimization problem as below [1]

$$\begin{aligned} P = \max_{x_i\geq 0}\sum_{i=1}^N U_i(x_i) \\ \text{s.t. } \sum_{i=1}^N R_{ji}x_i \leq c_j, \forall j \in [L]. \end{aligned} \tag{2.24}$$

for appropriately chosen utility functions $U_i(\cdot)$. The corresponding dual optimization problem is

$$D = \min_{p_j \geq 0} \sum_{j=1}^{L} \max_{x_i \geq 0} \left(U_i(x_i) - x_i \sum_{j=1}^{L} R_{ji} p_j \right) + \sum_{j=1}^{L} p_j c_j, \tag{2.25}$$

where $\{p_j\}$ is a vector of Lagrange dual variable.

Since we have strictly monotonic objective function defined on an nonempty convex compact set in the primal optimization problem (2.24), there exists a unique solution to this problem, which we denote as $\{x_i^{*'}\}$. By the concavity of the primal objective function and Slater's condition, strong duality holds, that is, $P = D$. The solution to the dual problem $\{p_j^{*'}\}$ satisfies the first order condition

$$\begin{aligned} U_i'(x_i^{*'}) &= \textstyle\sum_{j=1}^{L} R_{ji} p_j^{*'}, \\ p_j^{*'} (\textstyle\sum_{i=1}^{N} R_{ji} x_i^{*'} - c_j) &= 0. \end{aligned} \tag{2.26}$$

Back to the TCP/Reno system, from the definition and the assumption that the feasible rates are all above $1/\tau_i$, it is true that $U_i'(x_i^{*'}) \in [0,1]$. Therefore for all j, $p_j^{*'} \in (0,1)$ holds. If we define $\bar{x}_i = \bar{x}_i', \forall i$ and $h_j(b_j^*) = \bar{p}_j, \forall j'$, then the equilibrium equations hold and we have just shown the existence of the equilibrium rates and queue sizes. Conversely, from the strong duality any $\{\bar{x}_i\}$ and $\{\bar{p}_j\}$ that satisfies the first order condition (2.26) constitutes the optimal solution to the primal (2.24) and dual (2.25) problem, which is unique. Therefore we have unique equilibrium rates and queue sizes for both packet and window update schemes with TCP/Reno controller. Similar arguments can be shown to hold for the case of PF controller, in which the utility function is $U_i(x) = \frac{w}{\beta} \log x$.

Since the primal optimization problem (2.24) possesses a separable structure in both the objective function and the constraints, the dual problem (2.25) has a nice interpretation of decentralized optimization. The value of $U_i(x_i)$ indicates the benefit the source receives by sending its flow at rate x_i and the dual variable p_j can be interpreted as the unit price charged by accessing link j. Then the dual optimization (2.25) can be viewed that each source selects its transmission rate x_i such that its profit $U_i(x_i) - x_i \sum_{j=1}^{L} R_{ji} p_j$ is maximized.

Remark 2.1 *It is worth to note that this model of network flow optimization is a special case of the classical pure exchange economy* $(I, \mathcal{E})$*, in which* I *is a finite set of consumers and a map* $\mathcal{E} : I \to \mathbb{R}_+^{L+1} \times \mathcal{P}$ *assigns to each consumer* $i \in I$ *an initial endowment* $e(i) \in \mathbb{R}_+^{L+1}$ *and a preference relation* $\succeq_i \in \mathcal{P}$*. In our context each network user corresponds to a consumer and each link corresponds to a commodity, and the additional* $(L+1)$*th commodity is the "numeraire", upon which all the other commodities are evaluated. The initial endowment is such that* $\sum_{i \in I} e_j(i) = c_j$ *for all* $j = 1, \cdots, L$*. The preference relation* $\succeq_i$ *on* $\mathbb{R}_+^{L+1}$ *is a subset of* $\mathbb{R}_+^{L+1} \times \mathbb{R}_+^{L+1}$*. In the network flow optimization, for any two consumption vectors* (x^1, m^1) *and* (x^2, m^2)*, where* $x^k \in \mathbb{R}_+^L$ *is the bandwidth allocation and* m^k *is the numeraire, we have* $(x^1, m^1) \succeq_i (x^2, m^2)$ *if and only if* $U_i(\min_{j:R_{ji}=1} x_j^1) + m^1 \geq U_i(\min_{j:R_{ji}=1} x_j^2) + m^2$*. The Pareto optimal of the economy is exact the solution (in terms of effective bandwidth allocation) of the network optimization problem (2.24), and the dual solution with the price concept is exactly the competitive equilibrium (or Walrasian equilibrium) of the economy. A direct consequence of the competitive equilibrium is that the equilibrium is fair in the sense of envy-freeness of the "net trade" [14] (another popular interpretation of fairness only considers a special type* *of utility functions [15]). That is, given the Pareto optimal solution, the allocations of user* i_1 *and* i_2 *are such that* $(x(i_1), m(i_1)) \succeq_{i_1} (x(i_2), m(i_2)) - e(i_2) + e(i_1)$ *and* $(x(i_2), m(i_2)) \succeq_{i_2} (x(i_1), m(i_1)) - e(i_1) + e(i_2)$*. In other words, no one desires to replace his own change of allocation with somebody else's. Special to our network flow optimization problem, an important feature of the preference relation is that it is quasi-linear. Consequently, the bandwidth allocation part does not depend on the initial individual endowment of numeraire and bandwidths, as long as the aggregate endowment is fixed. Therefore the bandwidth part of any core allocations is the same as the competitive equilibrium, since we have unique Pareto optimum in terms of effective bandwidths and any core allocation is a Pareto optimum. That is to say, no coalition of a subset of all network users can achieve better utilities for all its members than those of competitive equilibrium. This is a rather strong property since this kind of core equivalence is usually only achieved in large economies [16].*

In the network algorithms, different utility functions represents different congestion control protocols, as we have already seen in the case of TCP/Reno and proportional fairness controller. It is shown in [17] that any link marking algorithms which satisfy

$$\sum_{i=1}^{N} R_{ji} x_i^* \leq c_j \text{ with equality if } p_j > 0$$

solve the network optimization problem in equilibrium. AQMs such as RED, PI, RED, droptail, and etc. all satisfy this condition. Therefore TCP/AQM protocols can be seen as decentralized primal-dual algorithms to solve the global network optimization problem. In the subsequent chapters we will base all of our protocol studies and designs on this optimization framework.

Chapter 3

Stability Results for Networks with Time Varying Delays - Classical Source/Link Controllers

3.1 Introduction

Recently there have been extensive studies [18, 19, 20, 21, 22] in the stability problem of the distributed algorithms for network with elastic traffic. This problem stems from Internet congestion control [23] and router design. The objective of Internet congestion control [23] is to allocate network bandwidth among Internet users in a fair and efficient way. It is known that the equilibrium value of bandwidth allocation is the solution to a centralized utility maximization problem of a whole network [1, 17]. To solve the network optimization problem, decentralized algorithms can be designed from the dualization of the original problem: source rates ("primal variables") are decoupled from each other at each link constraint by introducing congestion signals or prices ("dual variables") such as packet dropping probabilities. Each user updates its rate according to the aggregate congestion prices along the links that its traffic traverses and in the meantime each link adjusts its price by its aggregate rate of arriving traffic. Together the user and link dynamics drive the system to the rate and price equilibrium without knowing their actual values in prior if the dynamics converge. This summarizes the common structure of some distributed algorithms for network with elastic traffic, including widely implemented TCP/AQM (Transmission Control Protocol with Active Queue Management) system. Therefore the ability of these algorithms to converge from any initial values is required for implementation. Without stability user traffic rates will never reach the optimal equilibrium and will oscillate which may cause link under-utilization and frequent packet drops.

The system of network with elastic traffic is a nonlinear system with heterogeneous delays. Early studies on the stability of network with elastic traffic focus on either linearized version of the algorithm [18, 10] by frequency domain approach or delay-free case [21] by Lyapunov-based methods. Since the system does not know its equilibrium values in prior it usually can not be guaranteed that the system operates in a region near equilibrium. Therefore the stability condition from the analysis of the linearized version can only give us limited information about the dynamic behavior of the network. The analysis of the effect of delays in the network is important since their existence may bring instability of the network (so we are only interested in network with delay-dependent stability, which is different from delay-independent stability studied in [20]). In fact all three types of distributed algorithms studied in this paper are globally stable in a delay-free situation. In addition, delays in the network are usually not known exactly and they are time-varying in nature because part of delays are caused by queuing latencies at routers which change frequently according to their congestion levels. Another desirable feature of the distributed algorithm is that the stability condition can be satisfied with network information that can be accessed by users and links. This is important for the implementation of the algorithm since to avoid extra communication costs we wish each user or link collects relevant information only from their "local" measurements and different users or links do not cooperate with each other. Therefore in this paper we intend to analyze the global stability conditions of three types of popular distributed algorithm for network with elastic traffic and time-varying delays and also show that these conditions can be satisfied by each user and link individually so that the system-wide stability is ensured.

Our approach to the stability problem of network system with heterogeneous delays is to use Lyapunov-Krasovskii method, which relies on Lyapunov-Krasovskii functional instead of Lyapunov function in the analysis of delay free systems. It is one of the most general methods in analyzing delay system and it can be shown that the results from Small-Gain Theorem [19] (which is equivalent to Lyapunov-Razumikhin method) can be obtained by Lyapunov-Krasovskii method and improved by a better choice of Lyapunov-Krasovskii functional. Generally, for systems with time-varying delays, stability conditions obtained by Lyapunov-Razumikhin method have advantage over those by Lyapunov-Krasovskii method where the former does not depend on the time-derivative of delays but the latter does. However, due to the special structure of the system of network with elastic traffic, our results based on Lyapunov-Krasovskii method are independent of the time-derivative of delays, which is desirable for the implementation of the algorithms since the delays in the network can be "jittery".

In Section 3.2 we present our network model in detail and introduce three popular distributed algorithms for network with elastic traffic to study in detail. We discuss previous results and methods on the stability problem and make comparisons in Section 3.3. In Section 3.4 we present our main results on stability conditions by Lyapunov-Krasovskii method for each algorithm in subsections. Because of the "distributed" structure of the stability conditions we discuss the implementation issues in Section 3.5 and finally conclude in Section 3.6.

3.2 Network Model

It is known that fair bandwidth allocation can be understood as a solution to network optimization problem of some user utility function. The network optimization problem can be written as follows (recall (2.24),

$$\max_{x>0} \textstyle\sum_{i=1}^{N} U_i(x_i) \tag{3.1}$$

$$\text{s.t. } Rx \leq c$$

where x_i is the allocated rate of flow i and $U_i(\cdot)$ is the utility function of user i, which is a strictly concave, non-decreasing function. Here the designated utility function of an individual user reflects the system-wise fairness requirement [15], not individual user preference per se as in the game theoretic framework [24]. In some situations we need the conditions on maximally allowed transmission rate for each user to ensure global stability. This is done by constraining the maximum second derivative of individual utility by constant η_i for each user i as $U_i''(x) \leq -\eta_i < 0, \forall x \geq 0$, since $U_i(\cdot)$ is a strictly concave and non-decreasing, $U_i''(x) \nearrow 0, x \to \infty$. The constraints of this optimization problem are simply bandwidth limitations. In the case of TCP, the approximate utility function is $-1/(\tau^2 x)$ where τ is the round-trip time (compared to the result in Section 2.3). Network optimization of this particular utility function can be interpreted as a weighted minimum transmission time problem. A salient feature of this optimization problem is that by dualization the primal variable x_is in the link constraints can be decoupled by the equivalent optimization

$$\min_{p>0} \sum_{i=1}^{N} \max_{x>0} \left(U_i(x_i) - x_i \sum_{j=1}^{L} R_{ji} p_j \right) + \sum_{j=1}^{L} c_j p_j \tag{3.2}$$

where p_j is the dual variable and is interpreted as "price" at link j. Then by gradient method there exist algorithms, which will be introduced below, to solve the optimization problem completely in a distributed way where the exchange of state variables between users and links is done by aggregate price q_i at each user i and by aggregate rate y_j at each link j. Hence the algorithms do not require extra communications to obtain information about network congestion conditions and user decisions.

Before presenting the distributed algorithms we must specify the network delays. We denote at time t at link j the forward delay from user i to link j as $\tau_{ij}^{f}(t)$. In other words if a packet of flow i leaves user i at time $t - \tau_{ij}^{f}(t)$ then it arrives at link j at time t. Similarly we denote at time t at user i the backward delay from link j to user i as $\tau_{ij}^{b}(t)$. Therefore the round-trip delay at time t seen by user i is $\tau_{ij}^{b}(t) + \tau_{ij}^{f}(t - \tau_{ij}^{b}(t))$ for any link j on the path of flow i. In some cases we omit forward delays τ_{ij}^{f} since bottleneck links often appear at the network entrance point and thus forward delays only constitute a small portion of the total round-trip delay. In general we have the network model as follows, Network model:

$$\begin{cases} q_i(t) = \sum_{j=1}^{L} R_{ji} p_j(t - \tau_{ij}^b), & \forall i \in [N] \\ y_j(t) = \sum_{i=1}^{N} R_{ji} x_i(t - \tau_{ij}^f), & \forall j \in [L]. \end{cases} \tag{3.3}$$

There are different approaches to apply the gradient method to solve the network optimization problem. In [1] a distributed algorithm based on primal optimization problem (3.1) is proposed. Each source uses first-order dynamics to update his flow rate and each link uses static penalty function to prevent aggregate flow to exceed his bandwidth capacity. We call it "primal law" algorithm and it has the follow form:

$$\begin{cases} \dot{x}_i = K_i(U_i'(x_i) - q_i), & \forall i \in [N] \\ p_j = h_j(y_j), & \forall j \in [L]. \end{cases} \tag{3.4}$$

where $K_i > 0$ for each user i and $h_j : \mathbb{R}_+ \to \mathbb{R}_+$ is an increasing function for each link j with derivative bounded by constant ξ_j as $0 < h_j'(x) \leq \xi_j, \forall x \geq 0$. The equilibrium can be close to the optimal solution of (3.1) by appropriately chosen penalty functions $h_j(\cdot)$s. But this algorithm does not reflect the queuing dynamics taking place in the actual network. Also in [1] a gradient method on the dual variables p_js result in a different distributed algorithm. Now each link uses first-order dynamics to update its price, just like the dynamics of queue accumulation and depletion, and each user uses static function to solve individual utility maximization as in the dual problem (3.2). This is in turn called "dual law" algorithm and it can be written as

$$\begin{cases} x_i = U_i'^{-1}(q_i), & \forall i \in [N] \\ \dot{p}_j = \Gamma_j(y_j - c_j)^+_{p_j}, & \forall j \in [L]. \end{cases} \tag{3.5}$$

where $\Gamma_j > 0$ for each link j. Recall the form $[u]^+_v$ takes the value u if $v > 0$ and $\max\{u, 0\}$ if $v = 0$. In addition, better convergence speed can be achieved by using first-order dynamics in both user and link update laws. This is the same as Lagrangian method for convex programming in [25] and TCP/AQM systems can be modeled in this framework. We call it "primal/dual law" algorithm and it has the form below

$$\begin{cases} \dot{x}_i = K_i(U_i'(x_i) - q_i), & \forall i \in [N] \\ \dot{p}_j = \Gamma_j(y_j - c_j)^+_{p_j}, & \forall j \in [L]. \end{cases} \tag{3.6}$$

3.3 Literature Review

The major difficulty of stability analysis of these distributed flow control algorithms lies in system nonlinearity and existence of delay. Initial studies often focuses on linearized systems [18, 17, 12, 26, 10] and their results lead to the understanding and the design of control algorithms such as RED [12], PI [26], scalable controller [9], etc. But as we know that the analysis of a linearized system only guarantees the local behavior of the system, global behavior can be qualitatively different from local behavior even though the delay-free system is globally asymptotically stable as manifested by a result implied in [27]:

Proposition 3.1 *Consider a scalar delay differential equation*

$$\dot{x}(t) = -\delta x(t) + w(x(t-h)), \delta > 0.$$

Then for every $\alpha \geq 0$ there exists a smooth strictly decreasing, bounded below function w with $-w'(0) = \alpha$ and $w(0) = 0$, and such that the above equation has a nontrivial periodic solution which is hyperbolic, stable, and exponential attracting with asymptotic phase (so the trivial solution of the equation may not be globally stable although it is locally stable).

Therefore existing results on distributed algorithms in which both the local stability with time delays and global stability without time delays [1, 21] hold do not guarantee the stability of global behavior of network system with time delays. Desirable properties from local analysis, such as in scalable controller where stability can be maintained independent of network, need further investigation in the global region. Indeed, recent studies [22, 28] only verify the network-dependent global stability condition for scalable controller. We will briefly discuss two main approaches used in the literature for the study of global stability of distributed algorithms on network flows.

Consider the delay differential equation of the following form,

$$\dot{x}(t) = f(t, x_t), \tag{3.7}$$

where we define the function $x_t \in C(\triangleq C([-\tau, 0], \mathbb{R}^n))$ by $x_t(\theta) = x(t+\theta), \theta \in [-\tau, 0]$ and for any $\phi \in C$ define its norm by $\|\phi\| = \sup_{\theta \in [-\tau,0]} |\phi(\theta)|$. Extended from Lyapunov-based method in stability analysis of delay-free systems, two main approaches for dealing with the stability of time delay systems have been widely used in the past. The first is based on Lyapunov-Krasovskii functionals and the second is based on Razumikhin Theorem. The first method requires the construction of a nonnegative functional with decreasing values along the solution trajectory but the second method only asks for a nonnegative function whose value decreases at some special moments. The theorems related to these two methods can be stated as follows.

Theorem 3.1 (Lyapunov-Krasovksii, Theorem 5.2.1 of [29]) *Suppose that the functional $f : \mathbb{R} \times C \to \mathbb{R}^n$ takes bounded sets of $\mathbb{R} \times C$ into bounded sets of $\mathbb{R}^n$, and $u, v, w : \mathbb{R}_+ \to \mathbb{R}_+$ are continuous nondecreasing functions with $u(s), v(s), w(s) > 0$ for $s \neq 0$ and $u(0) = v(0) = 0$. If there exists a continuous functional $V : \mathbb{R} \times C \to \mathbb{R}$ such that*

(i) $u(|\phi(0)|) \leq V(t, \phi) \leq v(\|\phi\|)$,

(ii) $\dot{V}(t, \phi) \triangleq \limsup_{h \to 0^+} h^{-1}[V(t+h, x_{t+h}(t, \phi)) - V(t, \phi)] \leq -w(|\phi(0)|)$,

then the solution $x = 0$ of the equation (3.7) is uniformly asymptotically stable.

Theorem 3.2 (Lyapunov-Razumikhin, Theorem 5.4.2 of [29]) *Suppose functions f, u, v, w satisfy the same conditions as in the statement of Theorem 3.1. Assume that there exists a continuous function $V : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ such that*

$$u(|x|) \leq V(t, x) \leq v(|x|), \quad t \in \mathbb{R}, \quad x \in \mathbb{R}^n,$$

and there is a continuous nondecreasing function $r : \mathbb{R}_+ \to \mathbb{R}_+$, $r(s) > s$, such that

$$\dot{V}(t, \phi(0)) \triangleq \limsup_{h \to 0^+} h^{-1}[V(t+h, x(t, \phi)(t+h)) - V(t, \phi(0))] \leq -w(|\phi(0)|),$$
$$\text{if} \quad V(t+\theta, \phi(\theta)) < r(V(t, \phi(0))), \quad \forall \theta \in [-\tau, 0],$$

then the solution $x = 0$ of the equation (3.7) is uniformly asymptotically stable.

Due to relative simplicity of constructing functions over functionals, Lyapunov-Razumikhin method or its equivalence has been applied to the stability analysis of network flow algorithms. Deb and Srikant [30] used this method to study single source/link network with Kelly's primal algorithm. Fan et al [19] used ISS Small-Gain Theorem, which is equivalent to Razumikhin method [31], to study both primal and dual algorithm in general networks. There are several kinds of conservativeness when we adopt Razumikhin method to obtain stability conditions. First is due to the crude estimate of the worst case of delayed dynamics to ensure the decrease of Lyapunov-Razumikhin function over some critical moments. This deficiency is actually shared by some most common selections of Lyapunov-Krasovskii functional and it is difficult to improve. For example if the single source update law is the following

$$\dot{x}(t) = k\left(w - \frac{1}{xU'(x)}x(t-\tau)p(x(t-\tau))\right)$$

with $U(x) = -\frac{1}{ax^a}$ and $p = \left(\frac{x}{C}\right)^b$ for some positive constants a, b, and C. Then Deb and Srikant's result shows that the system is globally asymptotically stable when $\tau \leq ck^{-1}M^{a-b}l^{a+1}$ where M and l are upper and lower bound of rate x and c is some constant independent of M and l when M is large and l is small. However by contracting mapping method [20] the system is globally asymptotically stable for arbitrary large τ if $a < b+1$. Second cause of conservativeness brought by the application of Razumikhin method for the analysis of general networks is due to the fact that the calculation of the critical moments $\{t \in \mathbb{R}_+ : V(t+\theta, \phi(\theta)) < r(V(t, \phi(0)), \forall \theta \in [-\tau, 0]\}$ when the value of Razumikhin function decreases requires collective knowledge of states. Therefore 1) it is hard for a distributed algorithm to meet the stability conditions depending on global information; 2) the stability conditions themselves become overly restrictive since any users and links, even with "good" delay parameters, have to comply with the global stability requirement. In contrast a carefully constructed Lyapunov-Krasovskii functional decouples the system dynamics during the analysis so that the stability of the dynamics of each user or link can be satisfied individually with the improvement of stability region. We can even design Lyapunov-Krasovskii functional (see Appendix A) to proved the global stability of rate controlled networks described by Ranjan and La [20, 32], in which the stability results are obtained by the analysis of solution trajectory. This is one of the major themes we explore in Section 3.4. There we also prove that our estimate of stability region is better than those obtained in [19] by the equivalence of Razumikhin method.

Ranjan and La [20, 32] have shown a family of distributed algorithms which possess a remarkable feature of delay-independent stability. In their algorithms, the network is asymptotically stable when the delays can be arbitrarily large and time varying. The idea is that the dynamics of the algorithms can be considered as contraction mappings so that the future trajectories are confined within initial invariant region. However their algorithms do not really solve the optimization problem (3.1). That is, the final equilibrium point is not the solution to the optimization problem unless the network knows exact information of users, such as user utilities and number of users, etc., which is not a desirable requirement for distributed algorithms. Let us illustrate this point briefly. Suppose we have a bottleneck link dynamics G

with bandwidth $c > 0$ such that the system with one or more users accessing the bottleneck link converges to a unique solution of some optimization problem if the users' utility functions belong to a class of utility functions. Specifically for a single user/link system consider a primal/dual algorithm with rate x and price q as below

$$\begin{aligned} \dot{x}(t) &= \alpha^{-1} F(\alpha x(t), \alpha x(t-T), q(t-T)) \\ \dot{q}(t) &= G(x(t), q(t)), \end{aligned} \tag{3.8}$$

where α is a weight parameter in user's utility function. For n identical users accessing the bottleneck link the system is as follows,

$$\begin{aligned} \dot{x}_i(t) &= F(x_i(t), x_i(t-T), q(t-T)), \quad i \in [n] \\ \dot{q}(t) &= G\left(\sum_i x_i(t), q(t)\right). \end{aligned} \tag{3.9}$$

We assume that the both of above systems are globally asymptotically stable for all $0 \leq \alpha \leq 1$ and $n \geq 1$ with arbitrarily large T in some invariant regions from the argument of contraction mapping. Suppose that the equilibrium points of both systems achieve the unique optimal solution of corresponding optimization problems. We try to show this is false. First let us consider the single user system (3.8). Denote the invariant region of the system with parameter α by $I_x(\alpha) \times I_q(\alpha)$. By common techniques of contraction mapping, there exist a function $\mathcal{F}_\alpha : I_x(\alpha) \times I_q(\alpha) \to I_x(\alpha)$ and $\mathcal{G} : I_x(\alpha) \to I_q(\alpha)$ such that $F(\alpha \mathcal{F}_\alpha(x, q), \alpha x, q) \equiv 0$ and $G(x, \mathcal{G}(x)) \equiv 0$. Additionally it requires that for any $I_1 \times I_2 \subseteq I_x(\alpha) \times I_q(\alpha)$, $\mathcal{F}_\alpha(I_1, I_2) \times \mathcal{G}(I_1) \subset I_1 \times I_2$. From the assumption the equilibrium x is the bottleneck bandwidth c for all feasible α. By definition $\mathcal{F}_1(x, q) = \alpha^{-1} \mathcal{F}_\alpha(\alpha x, q)$. So we have $\mathcal{F}(\alpha c, \mathcal{G}(c)) =$

αc. Since it holds for all $0 \le \alpha \le 1$ one concludes that $\mathcal{F}(x, \mathcal{G}(c)) = x$ for all $0 \le x \le \bar{x}$ for some constant $\bar{x}$. Returning back to the multiple users system (3.9), one immediately sees that any point with $\sum_i x_i = c$ is an equilibrium point of the equations. Therefore the solution violates the uniqueness condition. So one must seek different constructions of distributed algorithms to solve network optimization problems, especially of types (3.5,3.6) which achieve the exact intended solution when in equilibrium.

Our construction of Lyapunov-Krasovskii functional is similar to some canonical form for linear delay systems. It is known that for delay-dependent stability of linear delay system of the following

$$\dot{x}(t) = Ax(t) + A_d x(t-\tau), \tag{3.10}$$

we can transform it into the form below

$$\dot{x}(t) = (A + A_d)x(t) - A_d \int_{t-\tau}^{t} \dot{x}(s)ds$$

and then use the Lyapunov-Krasovskii functional for some positive definite matrices P, R,

$$V(x_t) = x^T(t)Px(t) + \int_{-\tau}^{0}\int_{t+\theta}^{t} \dot{x}^T(s)A_d^T R A_d \dot{x}(s)dsd\theta.$$

The upperbound of the derivative of the functional along the trajectory is estimated by completing the squares. This type of transformation and Lyapunov-Krasovskii functional can be adapted to the case of time-varying delays as well. But usually the stability requires the delays to have uniformly bounded derivatives $d < 1$ [33]. It turns out that the stability of our system with time-varying delays does not depend on the time derivatives of delays. Let us intuitively illustrate this point by the following observation,

Proposition 3.2 *Suppose the delay τ in (3.10) is uniformly bounded function of time $0 < \tau(t) < \bar{\tau}, \forall t \geq 0$ and the matrices A and A_d in (3.10) are symmetric matrices. Then the solution $x = 0$ of the equation (3.10) is asymptotically stable if $A + A_d \leq 0$ and*

$$\begin{bmatrix} -\frac{2}{\bar{\tau}} + Q & A_d \\ A_d & -Q \end{bmatrix} < 0$$

for some positive definite matrix Q.

Proof: Define the Lyapunov-Krasovskii functional $V(\phi)$ as

$$V(\phi) = -\phi(0)^T(A + A_d)\phi(0) + \int_{-\bar{\tau}}^{0}\int_{\theta}^{0} \dot{\phi}(s)^T Q \dot{\phi}(s) ds d\theta.$$

Then the derivative of V along the solution trajectory of (3.10) is

$$\begin{aligned}
\dot{V}(x_t) &= -2\dot{x}^T(t)\dot{x}(t) - 2\dot{x}^T(t)A_d\int_{-\tau(t)}^{0}\dot{x}(s)ds \\
&\quad +\bar{\tau}\dot{x}^T(t)Q\dot{x}(t) - \int_{-\bar{\tau}}^{0}\dot{x}^T(s)Q\dot{x}(s)ds \\
&\leq -2\dot{x}^T(t)\dot{x}(t) + \bar{\tau}\dot{x}^T(t)A_dQ^{-1}A_d\dot{x}(t) + \int_{-\bar{\tau}}^{0}\dot{x}^T(s)Q\dot{x}(s)ds \\
&\quad +\bar{\tau}\dot{x}^T(t)Q\dot{x}(t) - \int_{-\bar{\tau}}^{0}\dot{x}^T(s)Q\dot{x}(s)ds \\
&\leq -\dot{x}^T(t)(2 - \bar{\tau}(A_dQ^{-1}A_d + Q))\dot{x}(t).
\end{aligned}$$

Then by Shur's complement we have the conclusion. ■

As we see in the proof it is the symmetry of matrix A and A_d that leads to the stability condition independent of time derivative of delay. Such structural property exists in all our algorithms. For example the linearized version of the primal law algorithm (3.4) can be written as below

$$\dot{x}(t) = \Upsilon x(t) - R^T \Xi R x(t - \tau(t))$$

for some diagonal matrix Υ and Ξ if all delays are the same. Therefore we expect that our stability conditions do not depend on the time derivative of delays as well and we will show that it is indeed correct in the next section.

3.4 Stability Results

In this section we give our main results on three distributed algorithms. In all situations we denote equilibrium rate of flow i, equilibrium rate of aggregate flows at link j, equilibrium congestion at link j, equilibrium aggregate congestion at user i by x_i^*, y_j^*, p_j^*, and q_i^*, respectively. Also we assume delay functions $\tau_{ij}^f(t)$ and $\tau_{ij}^b(t)$ are all bounded above by $\bar{\tau}_{ij}^f$ and $\bar{\tau}_{ij}^b$.

3.4.1 Primal Control Law

A sufficient condition can be derived by Lyapunov-Krasovskii functional W_p of the following form,

$$\begin{aligned} W_p(\phi) &= -\sum_{i=1}^{N} \int_{x_i^*}^{\phi_i(0)} (U_i'(u) - q_i^*) du + \sum_{j=1}^{L} \int_{y_j^*}^{\psi_j} (h_j(u) - p_j^*) du \\ &+ \sum_{i=1}^{N} \sum_{j=1}^{L} \sum_{n=1}^{N} \frac{\xi_j}{2} R_{ji} R_{jn} \int_{-\bar{\tau}_{ij}^f - \bar{\tau}_{nj}^b}^{0} \int_{\theta}^{0} \dot{\phi}_n(\zeta)^2 d\zeta d\theta \end{aligned}$$

for any $\phi \in C([-\bar{\tau}, 0], \mathbb{R}^N)$ and $\psi_j = \sum_{i=1}^{N} R_{ji}\phi_i(0), \forall j \in [L]$. From the monotonicity of $U_i'(\cdot)$ and $h_j(\cdot)$ it is straightforward that the above functional is nonnegative.

Define $\hat{y}_j(t) = \sum_{i=1}^{N} R_{ji} x_i(t)$. The derivative of the first two parts of $W_p(x_t)$ along the solution trajectory is

$$
\begin{aligned}
& -\sum_{i=1}^{N} K_i(U_i'(x_i) - q_i)(U_i'(x_i) - q_i^*) + \sum_{j=1}^{L} (h_j(\hat{y}_j) - p_j^*) \sum_{n=1}^{N} R_{jn} K_n (U_n'(x_n) - q_n) \\
= \; & -\sum_{i=1}^{N} K_i(U_i'(x_i) - q_i)(U_i'(x_i) - q_i^*) + \sum_{n=1}^{N} K_n (U_n'(x_n) - q_n) (\sum_{j=1}^{L} R_{jn} h_j(\hat{y}_j) - q_n^*) \\
= \; & -\sum_{i=1}^{N} K_i(U_i'(x_i) - q_i)^2 - \sum_{i=1}^{N} K_i(U_i'(x_i) - q_i) \left(\sum_{j=1}^{L} R_{ji} h_j(\hat{y}_j) - q_i \right) \\
\leq \; & -\sum_{i=1}^{N} K_i(U_i'(x_i) - q_i)^2 + \sum_{i=1}^{N} |\dot{x}_i| \sum_{j=1}^{L} R_{ji} \sum_{n=1}^{N} R_{jn} \xi_j \int_{-\tau_{ij}^b(t) - \tau_{nj}^f(t - \tau_{ij}^b(t))}^{0} |\dot{x}_n(t+\theta)| d\theta \\
\leq \; & -\sum_{i=1}^{N} K_i(U_i'(x_i) - q_i)^2 + \sum_{i=1}^{N} \sum_{j=1}^{L} \sum_{n=1}^{N} \frac{\xi_j}{2} R_{ji} R_{jn} \left((\bar{\tau}_{ij}^b + \bar{\tau}_{nj}^f) \dot{x}_i^2 + \int_{-\bar{\tau}_{ij}^b - \bar{\tau}_{nj}^f}^{0} \dot{x}_n(t+\theta)^2 d\theta \right).
\end{aligned}
$$

Here we use the assumption that $0 \leq h_j'(y) \leq \xi$ holds uniformly. The derivative of the third part is as following,

$$
\sum_{i=1}^{N} \sum_{j=1}^{L} \sum_{n=1}^{N} \frac{\xi_j}{2} R_{ji} R_{jn} \left((\bar{\tau}_{ij}^b + \bar{\tau}_{nj}^f) \dot{x}_n(t)^2 + \int_{-\bar{\tau}_{ij}^b - \bar{\tau}_{nj}^f}^{0} \dot{x}_n(t+\theta)^2 d\theta \right).
$$

By adding them up we arrive at the following inequality,

$$
\dot{W}_p(x_t) \leq -\sum_{i=1}^{N} \left(K_i^{-1} - \sum_{j=1}^{L} \sum_{n=1}^{N} \xi_j R_{ji} R_{jn} (\bar{\tau}_{ij}^b + \bar{\tau}_{nj}^f) \right) \dot{x}_i(t)^2.
$$

Therefore we have the following theorem,

Theorem 3.3 *If the network optimization problem uses primal control law (3.4), then the optimal solution is globally asymptotically stable if the following inequality is satisfied for every $i \in [N]$*

$$
\sum_{j=1}^{L} \sum_{n=1}^{N} \xi_j R_{ji} R_{jn} (\bar{\tau}_{ij}^b + \bar{\tau}_{nj}^f) < K_i^{-1}.
$$

It is instructive to compare the results obtained by this Lyapunov-Krasovskii functional approach to the ones obtained by ISS small-gain theorem in [19]. Since it is well known that the system is ISS is equivalent to the existence of ISS-Lyapunov functions [34] and ISS small-gain theorems can also be proved by judicious construction of Lyapunov functions [35, 36], it follows then that the Lyapunov-Krasovskii approach gives better stability regions as those obtained in [19]. In the case of primal control law, [19] gives the global stability condition

$$\sqrt{2LN}\bar{\tau}\|R\|_2\bar{K}(\eta_1\eta_2^{-1}\bar{K}\underline{K}^{-1}+1)\xi(\bar{K}\underline{K}^{-1}\eta_2^{-1}\|R\|_2^2\xi+1)<1$$

where η_1 and η_2 are constants such that $-\eta_1 \leq U_i''(x) \leq -\eta_2 < 0, \forall i \in [N]$, and we define $\bar{\tau} = \max_i \tau_i$, $\bar{K} = \max_i K_i$, and $\underline{K} = \min_i K_i$. It is easy to see that the stability region of $\bar{\tau}$ is contained in

$$2\sqrt{2LN}\|R\|_2\bar{K}\xi\bar{\tau} < 1.$$

While the stability region obtained from Theorem 3.3 does not depend on η_1 and η_2, it contains the region

$$\bar{K}\xi\|R^TR\|_\infty\bar{\tau} < 1.$$

So in order to see whether the Lyapunov-Krasovskii approach gives better estimate of the stability region, it is sufficient to show

$$2\sqrt{2LN}\|R\|_2 \geq \|R^TR\|_\infty.$$

In fact we have the following Proposition,

Proposition 3.3 *For any $L \times N$ 0-1 matrix R, the following inequality holds*

$$\sqrt{LN}\|R\|_2 \geq \|R^TR\|_\infty.$$

Proof: Define $N \times L$ all 1 matrix Φ. Since $\|\Phi\|_2 = \sqrt{LN}$ we have

$$\sqrt{LN}\|R\|_2 = \|\Phi\|_2\|R\|_2 \geq \|\Phi R\|_2.$$

Note ΦR is a $N \times N$ matrix with all row vectors the same. Let us write its row vector as $[a_1, \ldots, a_N]$ and it is easy to see that $\sum_{i=1}^{N} a_i =$ number of 1's in R. Since we have $\|\Phi R\|_2 = \sqrt{\rho(\Phi R R^T \Phi^T)}$ where $\rho(A)$ denotes the spectral radius of matrix A, and the components of the matrix $\Phi R R^T \Phi^T$ are all $\sum_{i=1}^{N} a_i^2$, it follows

$$\|\Phi R\|_2 = \sqrt{N \sum_{i=1}^{N} a_i^2} \geq \sum_{i=1}^{N} a_i \geq \|R^T R\|_\infty.$$

■

Therefore we see that the Lyapunov-Krasovskii functional based method indeed gives better estimate than the ISS small-gain theorem in the primal control law.

3.4.2 Dual Control Law

We will also use a sufficient global stability condition by Lyapunov-Krasovskii functional approach. Define functional W_d as,

$$W_d(\phi) = -\sum_{i=1}^{N} \int_{q_i^*}^{\psi_i} (U_i'^{-1}(u) - x_i^*) du + \sum_{j=1}^{L} \sum_{i=1}^{N} \sum_{l=1}^{L} \frac{1}{2\eta_i} R_{ji} R_{li} \int_{-\bar{\tau}_{ij}^f - \bar{\tau}_{il}^b}^{0} \int_{\theta}^{0} \dot{\phi}_j(\zeta)^2 d\zeta d\theta$$

for any $\phi \in C([-\bar{\tau}, 0], \mathbb{R}^L)$ and $\psi_i = \sum_{j=1}^{L} R_{ji}\phi_j(0), \forall i \in [N]$. Again it is easy to see that W_d is nonnegative from the monotonicity of $U_i'(\cdot)$. For notational simplicity we define $\hat{q}_i(t) = \sum_{j=1}^{L} R_{ji} p_j(t)$. Then similar to the case of primal control law, we take the derivative of $W_d(p_t)$ along the solution trajectory and the first part becomes

$$
\begin{aligned}
& -\sum_{i=1}^{N}(U_i'^{-1}(\hat{q}_i(t)) - x_i^*)\sum_{j=1}^{L}R_{ji}\Gamma_j\sum_{n=1}^{N}R_{jn}(U_l'^{-1}(q_l(t-\tau_{nj}^f)) - x_n^*) \\
= \; & -\sum_{j=1}^{L}\Gamma_j\sum_{i=1}^{N}R_{ji}(U_i'^{-1}(\hat{q}_i(t)) - x_i^*)\sum_{n=1}^{N}R_{jn}(U_n'^{-1}(q_l(t-\tau_{nj}^f)) - x_n^*) \\
= \; & -\sum_{j=1}^{L}\Gamma_j\left(\sum_{i=1}^{N}R_{ji}U_i'^{-1}(q_i(t-\tau_{ij}^f)) - c_i\right)^2 \\
& -\sum_{j=1}^{L}\Gamma_j\sum_{i=1}^{N}R_{ji}(U_i'^{-1}(q_i(t-\tau_{ij}^f)) - U_i'^{-1}(\hat{q}_i(t))) \\
& \times\sum_{n=1}^{N}R_{jn}(U_n'^{-1}(q_n(t-\tau_{nj}^f)) - x_n^*) \\
\leq \; & -\sum_{j=1}^{L}\Gamma_j^{-1}\dot{p}_j(t)^2 + \sum_{j=1}^{L}|\dot{p}_j(t)|\sum_{i=1}^{N}\eta_i^{-1}R_{ji}\sum_{l=1}^{L}R_{li}\int_{-\tau_{ij}^f(t)-\tau_{il}^b(t-\tau_{ij}^f(t))}^{0}|\dot{p}_l(t+\theta)|d\theta \\
\leq \; & -\sum_{j=1}^{L}\Gamma_j^{-1}\dot{p}_j(t)^2 + \sum_{j=1}^{L}\sum_{i=1}^{N}\sum_{l=1}^{L}(2\eta_i)^{-1}R_{ji}R_{li} \\
& \times\left((\bar{\tau}_{ij}^f + \bar{\tau}_{il}^b)\dot{p}_j(t)^2 + \int_{-\bar{\tau}_{ij}^f-\bar{\tau}_{il}^b}^{0}\dot{p}_l(t+\theta)^2 d\theta\right). \qquad (3.11)
\end{aligned}
$$

The derivative of the second part becomes

$$
\sum_{j=1}^{L}\sum_{i=1}^{N}\sum_{l=1}^{L}(2\eta_i)^{-1}R_{ji}R_{li}\left((\bar{\tau}_{ij}^f + \bar{\tau}_{il}^b)\dot{p}_j(t)^2 - \int_{-\bar{\tau}_{ij}^f-\bar{\tau}_{il}^b}^{0}\dot{p}_j(t+\theta)^2 d\theta\right).
$$

Add these two together and we obtain the following inequality

$$
\dot{W}_d(p_t) \leq -\sum_{j=1}^{L}\left(\Gamma_j^{-1} - \sum_{i=1}^{N}\sum_{l=1}^{L}\eta_i^{-1}R_{ji}R_{li}(\bar{\tau}_{ij}^f + \bar{\tau}_{il}^b)\right)\dot{p}_l(t)^2.
$$

Therefore we conclude with the following theorem with regard to the global stability of dual control law,

Theorem 3.4 *If the dual control law (3.5) is used for solving the network optimization problem, then the optimal solution is globally asymptotically stable if the* *following inequality is satisfied for every* $j \in [L]$

$$\sum_{i=1}^{N}\sum_{l=1}^{L}\eta_i^{-1}R_{ji}R_{li}(\bar{\tau}_{ij}^f+\bar{\tau}_{il}^b)<\Gamma_j^{-1}$$

Again the result by ISS Small-Gain Theorem in [19] shows that the network with dual law algorithm and fixed uncertain delays is globally asymptotically stable if the following condition holds

$$\sqrt{2LN}\bar{\tau}\bar{\Gamma}\|R\|_2\left(\eta_2^{-1}+\frac{\eta_1\bar{\Gamma}\|R\|_2^2}{\eta_2^2\underline{\Gamma}\underline{\sigma}(R)^2}\right)<1 \tag{3.12}$$

where $\bar{\Gamma}=\max_j\Gamma_j$, $\underline{\Gamma}=\min_j\Gamma_j$, and $\underline{\sigma}(R)$ is the smallest singular value of routing matrix R. As in the case of primal law algorithm, Theorem 3.4 gives the maximum delay bound which is equivalent to

$$\eta_2^{-1}\bar{\Gamma}\|R^TR\|_\infty\bar{\tau}<1.$$

But from Proposition 3.3 we know that the stability region of $\bar{\tau}$ from the above inequality strictly contains that by the following inequality

$$\sqrt{2LN}\eta_2^{-1}\bar{\Gamma}\bar{\tau}<1$$

which in turn strictly contains the stability region from the condition (3.12). Hence our result from Lyapunov-Krasovskii method yields better estimate of stability region.

3.4.3 Primal/Dual Control Law

Here we assume the forward delay τ_{ij}^f are constants throughout the time. Since bottleneck links are usually concentrated at the network entrance points so the forward delays are mostly due to propagation latencies. We define our Lyapunov-Krasovskii functional as

$$\begin{aligned} W_{pd}(\phi,\psi) &= \sum_{i=1}^{N}\int_{x_i^*}^{\phi_i}\frac{u-x_i^*}{K_i}du+\sum_{j=1}^{L}\int_{p_j^*}^{\psi_j}\frac{v-p_j^*}{\Gamma_j}dv \\ &+ \sum_{j=1}^{L}\frac{\Gamma_j}{2}\sum_{i=1}^{N}\sum_{n=1}^{N}R_{ji}R_{jn}\sqrt{\frac{\bar{\tau}_{nj}^b+\bar{\tau}_{ij}^f}{\bar{\tau}_{ij}^b+\bar{\tau}_{nj}^f}}\int_{-\bar{\tau}_{ij}^b-\bar{\tau}_{nj}^f}^{0}\int_{\theta}^{0}(\phi_n(\zeta)-x_n^*)^2 d\zeta d\theta, \end{aligned}$$

where $\phi\in C([-\bar{\tau},0],\mathbb{R}^N)$ and $\psi\in\mathbb{R}^L$. For a particular solution of the primal/dual system, we define auxiliary state variables $z_i(t)=\int_{-\infty}^{t}(x_i(s)-x_i^*)ds$ for $i\in[N]$ and define $\hat{p}_j(t)=\Gamma_j\sum_{i=1}^{N}R_{ji}z_i(t)$. It is then easy to see that $\dot{z}_i(t)=x_i(t)-x_i^*$. Take the derivative of $W_{pd}(x_t,\hat{p}(t))$ along the trajectory and the first two parts become

$$\begin{aligned} & \sum_{i=1}^{N}(x_i-x_i^*)(U_i'(x_i)-q_i)+\sum_{j=1}^{L}(\hat{p}_j-p_j^*)\sum_{i=1}^{N}R_{ji}(x_i-x_i^*) \\ = & \sum_{i=1}^{N}(x_i-x_i^*)(U_i'(x_i)-q_i^*)+\sum_{i=1}^{N}(x_i-x_i^*)\sum_{j=1}^{L}R_{ji}(\hat{p}_j(t)-p_j(t-\tau_{ij}^b(t))) \\ = & \sum_{i=1}^{N}(x_i-x_i^*)(U_i'(x_i)-q_i^*)+\sum_{i=1}^{N}(x_i-x_i^*)\sum_{j=1}^{L}R_{ji}\Gamma_j\sum_{n=1}^{N}R_{jn}\int_{-\tau_{ij}^b(t)-\tau_{nj}^f}^{0}(x_n(t+\theta)-x_n^*)d\theta \\ \le & \sum_{i=1}^{N}(x_i-x_i^*)(U_i'(x_i)-q_i^*)+\sum_{j=1}^{L}\Gamma_j\sum_{i=1}^{N}\sum_{n=1}^{N}R_{ji}R_{jn} \\ & \times\int_{-\tau_{ij}^b(t)-\tau_{nj}^f}^{0}|x_i(t)-x_i^*(t)||x_n(t+\theta)-x_n^*|d\theta \\ \le & \sum_{i=1}^{N}(x_i-x_i^*)(U_i'(x_i)-q_i^*)+\sum_{j=1}^{L}\frac{\Gamma_j}{2}\sum_{i=1}^{N}\sum_{n=1}^{N}R_{ji}R_{jn}\sqrt{(\bar{\tau}_{ij}^b+\bar{\tau}_{nj}^f)(\bar{\tau}_{nj}^b+\bar{\tau}_{ij}^f)}(x_i(t)-x_i^*)^2 \\ & +\sum_{j=1}^{L}\frac{\Gamma_j}{2}\sum_{i=1}^{N}\sum_{n=1}^{N}R_{ji}R_{jn}\sqrt{\frac{\bar{\tau}_{nj}^b+\bar{\tau}_{ij}^f}{\bar{\tau}_{ij}^b+\bar{\tau}_{nj}^f}}\int_{-\bar{\tau}_{ij}^b-\bar{\tau}_{nj}^f}^{0}(x_n(t+\theta)-x_n^*)^2 d\theta. \end{aligned}$$

After adding the derivative of the third part we obtain

$$\begin{aligned}\dot{W}_{pd}(x_t,\hat{p}(t)) &\le \sum_{i=1}^{N}(x_i-x_i^*)(U_i'(x_i)-q_i^*)\\ &\quad+\sum_{j=1}^{L}\Gamma_j\sum_{i=1}^{N}\sum_{n=1}^{N}R_{ji}R_{jn}\sqrt{(\bar{\tau}_{ij}^b+\bar{\tau}_{nj}^f)(\bar{\tau}_{nj}^b+\bar{\tau}_{ij}^f)}(x_i-x_i^*)^2\\ &\le -\sum_{i=1}^{N}\eta_i(x_i-x_i^*)^2\\ &\quad+\sum_{i=1}^{N}\sum_{j=1}^{L}\sum_{n=1}^{N}\Gamma_jR_{ji}R_{jn}\sqrt{(\bar{\tau}_{ij}^b+\bar{\tau}_{nj}^f)(\bar{\tau}_{nj}^b+\bar{\tau}_{ij}^f)}(x_i-x_i^*)^2.\end{aligned}$$

Therefore we conclude with the following global stability criterion

Theorem 3.5 *If the primal/dual control law (3.6) is used for solving the network optimization problem, then the optimal solution is globally asymptotically stable if the following inequality is satisfied for every $i \in [N]$*

$$\sum_{j=1}^{L}\sum_{n=1}^{N}\Gamma_jR_{ji}R_{jn}\sqrt{(\bar{\tau}_{ij}^b+\bar{\tau}_{nj}^f)(\bar{\tau}_{nj}^b+\bar{\tau}_{ij}^f)} < \eta_i.$$

3.5 Implementation Issues

The implementation of a distributed algorithm over the network requires each user and link to set their parameters by information easily available to them so that the resulting algorithm leads the system to the optimal equilibrium state eventually. Although our stability conditions are distributed in nature, for example in the primal law algorithm the stability is achieved when all of individual user's stability conditions are satisfied, there are still some measurement issues to consider. A major obstacle is how to measure the packet forward delays, since it is impossible that every network users and link routers have their clocks in synchronization and even their local times may be different due to geographic diversity. Therefore we make an assumption that the main component of packet forward delays is between bottleneck routers and we neglect the delays between user hosts and their edge routers. The delays between bottleneck routers can be measured in a much slow time scale by messaging between nearby routers and accumulated in a routing table as a metric. Therefore we obtain an estimate of forward delays for each bottleneck routers. Now we discuss the implementation of each of distributed algorithms:

1. Primal Control Law: The stability condition in Theorem 3.3 can be written as

$$\sum_{j=1}^{L} \xi_j R_{ji} \sum_{n=1}^{N} R_{jn} \bar{\tau}_{ij}^{b} + \sum_{j=1}^{L} \xi_j R_{ji} \sum_{n=1}^{N} R_{jn} \bar{\tau}_{nj}^{f} < K_i^{-1}$$

   for every user i. For each link let us set $\xi_j = (\sum_{n=1}^{N} R_{jn})^{-1} \xi$ for some globally known constant ξ. In other words we set the penalty function $h_j(x)$ of each link j to be of the form $h(x)/\sum_{n=1}^{N} R_{jn}$ for some positive function $h(x)$ with $0 < h'(x) < \xi$. Here $\sum_{n=1}^{N} R_{jn}$ is exactly the number of flows entering link j. Also from our assumption each router j has the estimate of $\sum_{n=1}^{N} R_{jn} \bar{\tau}_{nj}^{f} / \sum_{n=1}^{N} R_{jn}$ which is the average forward delay. Here is how the algorithm is implemented. Each packet of flow i has two field, one records the number of bottleneck links it passes and the other records the cumulative average forward delay. When entering a bottleneck link each field is updated accordingly. Upon receipt of the packet the receiver sends back the acknowledgment packet with these two fields. The user i then adjust K_i so that

$$K_i < \frac{1}{\bar{\tau}_i \xi \times \text{number of links passed by } i + \xi \times \text{cumulative average forward delay}}$$

   and global stability can then be guaranteed.

2. Dual Control Law: In this situation the routers do not have to estimate the forward delays since the stability condition in Theorem 3.4 only involves the delay term $\bar{\tau}_{ij}^f + \bar{\tau}_{il}^b$, which is bounded by $\frac{3}{2}\bar{\tau}_i$ and can be estimated by user i. Therefore each user i adapts his η_i to be $\frac{3}{2}\bar{\tau}\sum_{l=1}^{L} R_{li}\eta$ for some globally known constant η. This can be achieved by either changing his maximum transmission rate or changing his utility function as in the case of scalable controller [9]. The parameter $\sum_{l=1}^{L} R_{li}$ is the number of links flow i passes and can be estimated in the same way as the case of primal law algorithms. Then each link j only needs to set his scaling parameter Γ_j so that

$$\Gamma_j < \frac{\eta}{\text{number of flows entering link } j}$$

to ensure global stability.

3. Primal/Dual Control Law: The stability condition stated in Theorem 3.5 can be relaxed into the form below for easy implementation,

$$\sum_{j=1}^{L}\sum_{n=1}^{N}\Gamma_j R_{ji} R_{jn}(\bar{\tau}_i + \bar{\tau}_n) < 2\eta_i,$$

since

$$\sqrt{(\bar{\tau}_{ij}^b + \bar{\tau}_{nj}^f)(\bar{\tau}_{nj}^b + \bar{\tau}_{ij}^f)} <= \frac{1}{2}(\bar{\tau}_{ij}^b + \bar{\tau}_{nj}^f + \bar{\tau}_{nj}^b + \bar{\tau}_{ij}^f) = \frac{1}{2}(\bar{\tau}_i + \bar{\tau}_n).$$

To implement the primal/dual algorithm over the network, each user i must report to each router its flow passes his maximum round-trip time $\bar{\tau}_i$. This can be done by either piggybacking the information along the data packet or by special control messages which can be sent in a much slow time scale. As in the case of primal law algorithm, the scaling parameter Γ_j for each link j is set to be $\Gamma / \sum_{n=1}^{N} R_{jn}$ for some globally known constant Γ. Again each packet of flow i uses two special field to record the number of links the flow traverses $\sum_{j=1}^{L} R_{ji}$ and the cumulative average round-trip time $\sum_{j=1}^{L} R_{ji} \sum_{n=1}^{N} R_{jn}\bar{\tau}_n / \sum_{n=1}^{N} R_{jn}$. Then user i has to ensure the minimum derivative of his utility function η_i to satisfy

$$\eta_i > \frac{1}{2}\Gamma(\bar{\tau}_i + \text{ cumulative average round-trip time })$$

by either restricting his maximally allowed transmission rate or changing his utility function. Thus the global stability can be achieved.

3.6 Discussion and Conclusion

We have systematically presented analytic methodology and results on the stability of distributed algorithms of network with elastic traffic with time-varying delays. Different methods of stability analysis have been compared and we show that our constructions of Lyapunov-Krasovskii functional yields better results in terms of stability region. Additionally as our results show that the global stability of the whole network system is achieved by separate conditions for individual users or links, we can design the implementations of these algorithms with parameters which can be set adaptively in a distributed way for changing network conditions.

We have mentioned the conservativeness caused by over-estimate of worst case dynamics in Lyapunov-based method is difficult to overcome. In some situations, the necessary stability condition of a single source/link network does not restrict the maximum transmission rate, while our results do. A simple example is Low's scalable controller [9] where the single source/link network uses dual control law and the single user has utility function $U(x) = \frac{\bar{\tau}x}{\alpha}(\log\frac{\hat{x}}{x}+1)$ with the maximum rate $\hat{x}$, the maximum delay $\bar{\tau}$ and some parameter α. The system dynamics obey the following equation

$$\begin{aligned} x(t) &= \hat{x}e^{-\frac{\alpha p(t-\tau(t))}{\bar{\tau}}}, \\ \dot{p}(t) &= \frac{x(t)-c}{c}, \end{aligned}$$

with usual notations p as link price and c as link capacity. By simple manipulation the above equations are equivalent to

$$\dot{\tilde{p}}(t) = e^{-\frac{\alpha\tilde{p}(t-\tau(t))}{\bar{\tau}}} - 1, \tag{3.13}$$

where we define $\tilde{p} = p - p^*$. It is worth mentioning that the global stability conditions of this system obtained in [37, 28] actually depend on the maximum rate $\hat{x}$ and the network capacity c since their analysis are based on sector nonlinearity. However a simple usage of contraction mapping leads to a better sufficient condition of $\alpha < 1$ for the global stability as follows. Suppose initially $\tilde{p}(\theta) \in [-m_0, M_0]$, for $-\bar{\tau} \le \theta \le 0$ and $-m_0 \le -1 < 0 \le \bar{\tau}(\exp(\alpha m_0/\bar{\tau}) - 1) \le M_0$. Suppose that $\tilde{p}$ achieves the next maximum M_1 at time t_1. Then there exists a time instance $s_1 < t_1$ closest to t_1 such that $\tilde{p}(s_1) = 0$ and $t_1 - s_1 = \tau(t_1) \le \bar{\tau}$. Therefore $M_1 = \int_{s_1}^{t_1} \dot{\tilde{p}}(\xi - \tau(\xi))d\xi \le \bar{\tau}(\exp(\alpha m_0/\bar{\tau})-1)$. Similarly the next minimum satisfies $m_1 \le \bar{\tau}(\exp(-\alpha M_0/\bar{\tau})-1)$. By defining function $f_\alpha(x) \triangleq \bar{\tau}(\exp(-\alpha x/\bar{\tau}) - 1)$ we can get recursive relations $M_{n+1} \le f_\alpha(m_n)$ and $m_{n+1} \le f_\alpha(M_n)$. Since 0 is the only fixed point of $f_\alpha \circ f_\alpha$ and it is locally stable when $\alpha < 1$, by Sharkovsky's Theorem the fixed point of f_α is globally stable if $\alpha < 1$. So the set $[-m_n, M_n]$ converges to zero and we reach the global stability of (3.13). Although this condition is not the best we can get since we bounded the dynamics in a very coarse way, but the result is still appealing because of its independence of network parameters. It means that the efficiency of a stable network is not hampered by its scale.

A recent more careful study [38] of the solution trajectory of scalar FDE renders the following result:

Theorem 3.6 (Liz, Tkachenko, and Trofimchuk) *Consider the following scalar functional differential equation*

$$\dot{x}(t) = f(t, x_t) \tag{3.14}$$

where $f : \mathbb{R} \times \mathcal{C} \to \mathbb{R}$ *is a continuous functional, where* $\mathcal{C} \triangleq C([-1, 0], \mathbb{R})$*. Suppose* $x(t) \equiv 0$ *is the unique equilibrium of the equation. Also the following generalized Yorke condition holds for either* $-3/2 \leq a < 0$ *and* $b > 0$ *or* $-3/2 < a < 0$ *and* $b = 0$*:*

$$\frac{aM(\phi)}{1 + bM(\phi)} \leq f(t, \phi) \leq \frac{-aM(-\phi)}{1 - bM(-\phi)}$$

where the first inequality holds for all $\phi \in \mathcal{C}$ *and the second inequality holds for all* $\phi \in \mathcal{C}$ *such that* $\min_{s \in [-1,0]} \phi(s) > -1$*. The functional* $M : \mathcal{C} \to \mathbb{R}$ *is defined as* $M(\phi) \triangleq \max\{0, \max_{s \in [-1,0]} \phi(s)\}$*. Then all the solutions of (3.14) converges to 0 as* $t \to \infty$*.*

It should be noted that the $3/2$ bound is the best attainable for general $f(t, x_t)$, since for linear functional $f(t, x_t) = -ax(t - \tau(t))$ it is demonstrated in [39] that for

$a = 3/2$ there exists $\tau(t)$ with $0 \leq \tau(t) \leq 1$ such that the solution is nonconverging and periodically varying.

A simple application of Theorem 3.14 to the scalable controller network (3.13) gives an improved bound $\alpha \leq 3/2$ for global stability. In addition, one can solve the global stability problem for scalable controller with the window update scheme, in which the rate update takes place every round-trip time and the fluid model is viewed as an aggregate of flows with different updating time instances. This can be understood as the source maintains a window of received packets for estimating the marking probability. See Section 2.1.1 for details. By exactly the same reasoning as the case of TCP/Reno, the fluid model can be written as an FDE with distributed delays as follows,

$$\dot{\tilde{p}}(t) = \frac{1}{\tau(t)} \int_{t-\tau(t)}^{t} \exp\left(-\frac{\alpha}{\bar{\tau}\tau(s)} \int_{s-\tau(s)}^{s} \tilde{p}(\theta)d\theta\right) ds - 1 \triangleq f_p(t, \tilde{p}_t),$$

where we suppose that $0 < \underline{\tau} \leq \tau(t) \leq \bar{\tau}$ for all t and $\tilde{p}_t \in C([-2\bar{\tau}, 0], \mathbb{R})$. Let us only consider the case where $-p_m = -\min_{s\in[-2\tau,0]} \tilde{p}_t(s) < 0 < p_M = \max_{s\in[-2\tau,0]} \tilde{p}_t(s)$, then we have

$$\exp\left(-\frac{\alpha}{\bar{\tau}} p_M\right) - 1 \leq f_p(t, \tilde{p}_t) \leq \exp\left(\frac{\alpha}{\bar{\tau}} p_m\right) - 1.$$

The upper and lower bound of $f_p(\cdot, \cdot)$ is the same as in the packet update model (3.13). Then the result of Theorem 3.14 can be directly applied and the stability bound of the system is $\alpha \in [0, 3/4]$.

Although we have quite strong result on the global stability for the single source/link network with time-varying delay, unlike delay-independent stability in [20], the method of contraction mapping based on solution trajectory for the study of delay-dependent stability quickly becomes unmanageable when the system dimension, i.e. the number of bottleneck links in the case of source controller, becomes greater than 1. In the next chapter, we will introduce new suite of protocols whose parameters only depend on local measurements and which achieve global stability for arbitrary networks.

Chapter 4

Design of Scalable and Distributed Control Laws

4.1 Introduction

The design objective of congestion control algorithms of TCP type over the Internet [23] is to achieve efficient and fair usage of bandwidth for each user with limited information on the user's network environment. The necessity of the requirement for limited (or local) information is a natural consequence of the size of the system we deal with. By limited information we mean only information which can be measured or obtained by each user directly through her interaction only with the part of the network relevant to her flow. For example TCP operates explicitly on the knowledge of the losses of user's packets, which can be seen as a congestion message sent by the intermediate routers, and implicitly on the round-trip delay of each user's flow through a self-clock mechanism, which is a direct measurement by the end user. On the other hand efficiency and fairness are design goals which depend on various combinations of different flows, which are definitely non-local to each user per se. However, by adopting an optimization framework to interpret the efficient and fair bandwidth allocation [1, 17, 15], one can immediately reformulate the original large coupled problem into smaller decoupled problems via duality. In essence, each user tries to maximize her own utility function (induced by the fairness requirement) which is a function of her flow rate (primal variables of the optimization problem). At the same time congestion messages generated at each router by Active Queue Management (AQM) can be seen as dual variables (or Lagrange multipliers for the bandwidth constraints). Then the distributed algorithm is executed between all users and routers in the network through the exchange of primal and dual variables. The original design goal is translated into the ability of the distributed algorithm to reach the global optimal point eventually.

Given the separable nature of the network optimization problem in our context, an immediate candidate for distributed algorithms comes from dual gradient methods [25, 17], in which the dual variables are updated based on a gradient approach and the primal variables are obtained directly by solving the first-order optimality condition. This is generally termed "dual law" since only the calculation of the dual variables has dynamics. A variant of the dual law algorithm, in which primal variables are also updated according to a certain kind of dynamics, is called "primal/dual law" and this actually corresponds to the Lagrangian method in the theory of optimization [25]. The only equilibrium point of these two algorithms is the solution of the global optimization problem. In reality TCP with AQM, which has a pure integrator term, can be modeled as a primal/dual law algorithm. Furthermore, there is a class of "primal law" algorithms, which can model AQM with arbitrary random dropping functions, but in a strict sense those algorithms do not solve the network optimization problem, since their equilibrium points are not guaranteed to be the optimal solution, although they can be arbitrarily close to the optima [1].

A major cause of problems in the aforementioned distributed algorithms is the existence of delays in the network. Information obtained from the network in order to update primal or dual variables is usually subject to delays due to the time spent on computation, propagation, and queues. This information staleness is one of the major destabilizing factors for the algorithm dynamics and it is well known that TCP/AQM algorithms do not scale with large delay and bandwidth: they either result in low utilization of the network resources or display perpetual fluctuations of flow rates. Many research efforts have been devoted to this issue. First results on a scalable control law were proposed by Low and Paganini et al for a particular utility function [9], and subsequently they extended their result for general utility functions [40]. But both protocols are only verified (validated) for a linearized situation by Vinnicombe's results on TCP/AQM network control with heterogeneous delays [10]; however the global behavior results of their scalable control law are restricted to a single source/link network [37]. General approaches of global stability analysis include Lyapunov-Krasovskii methods, Lyapunov-Razumikhin methods [29], and contraction mapping methods. Stability conditions for the primal law and dual law algorithms are obtained by Fan et al [19] by employing a Razumikhin equivalent method, but their condition requires global information about the network. A contraction mapping method is used by Ranjan et al [20] to analyze a class of congestion control algorithms which enjoys stability with arbitrary large delays.

Our work intends to design a scalable and distributed control algorithm for the network flow optimization problem such that the algorithm has global stability and only requires local information for both users and routers. Specifically each user only needs to know the number of bottlenecks his flow traverses, the round-trip delay, and the aggregate congestion of his flow, and each router only needs to know the number of flows and the aggregate flow it has. In this way such a controller has a nice plug-and-play property which is desirable for actual implementation. The paper is organized as follows. In Section 4.2 we will present the network model and problem formulation. We will put forward our design principles there. In Section 4.3 the general properties that valid controllers must have are discussed based on some of our design principles. The scalable controller is then designed in Section 4.4 and its global stability is proved. Final discussion and conclusions are given in Section 4.5.

4.2 Problem Formulation

The network considered in this paper is similar to the one in [40] and consists of N users and L bottleneck links (all those links whose bandwidths are fully utilized at equilibrium). We use the notation $[n]$ for the set $\{1, \cdots, n\}$ and the operator $|\cdot|$ for set cardinality. Therefore we have for the user set $[N]$ and for the bottleneck link set $[L]$. In reality network links other than bottleneck links may have effects on the dynamics of network flows. But for simplicity we only consider those bottleneck links, which we abbreviate as "links" hereafter. Each user i has a fixed flow path $r_i \subset [L]$ to send a file with infinite length. In other words we only consider persistent flows. Also each router at link j has a set $f_j \subset [N]$ of accessing flows. The routing matrix $R \in \{0,1\}^{L\times N}$ is defined as

$$R_{ji} \triangleq \begin{cases} 1, & j \in r_i, \\ 0, & j \notin r_i. \end{cases}$$

We denote by x_i the flow rate of user i and by p_j the congestion information on link j. Due to the packet forward delay incurred during the transmission of flow packets, the aggregate flow rate seen by the router at link j at time t is

$$y_j(t) = \sum_{i \in f_j} x_i(t - \tau_{ij}^f) = R_{j \cdot} \begin{bmatrix} x_1(t - \tau_{1j}^f) \\ \vdots \\ x_N(t - \tau_{Nj}^f) \end{bmatrix} \tag{4.1}$$

where τ_{ij}^f is the forward delay from user i on link j. Accordingly, the aggregate congestion information received by each user i at time t is

$$q_i(t) = \sum_{j \in r_j} p_i(t - \tau_{ij}^b) = [p_1(t - \tau_{i1}^b), \cdots, p_L(t - \tau_{iL}^b)] R_{\cdot i} \tag{4.2}$$

where τ_{ij}^b is the backward delay from link j to user i. For reasons of fast computation and small communication cost, routers cannot differentiate individual flows and users cannot differentiate congestion levels of individual links. All they have access to are aggregate information and we will show that these are actually sufficient for our purposes.

An important assumption made now is that both forward delays and backward delays are time invariant, which is a valid approximation when routers have small buffers compared to the product of bandwidth and propagation delays. Then the observation that $\tau_{ij}^b \geq \tau_{ij}^f$ usually holds if the reverse route is symmetric with respect to the forward route. We also use the following definition

$$\tau_i \triangleq \tau_{ij}^f + \tau_{ij}^b, \forall j \in r_i, \tag{4.3}$$

which is the round-trip delay of flow i. Again it consumes extra communication bits to accumulate information about forward delays and backward delays separately, and in contrast it is straightforward for each user to measure the round-trip delay. Therefore it is much more desirable to design algorithms whose parameters depend not on the forward/backward delays separately but only on the round-trip delays.

As mentioned in Section 4.1 the problem of efficient and fair allocation of network bandwidths can be cast into the problem of network optimization over flows. This optimization problem is a classical convex programming problem with linear constraints:

$$\begin{aligned} &\max_{x_i \geq 0} \textstyle\sum_{i\in[N]} U_i(x_i) \\ &\text{s.t. } Rx \leq c. \end{aligned} \tag{4.4}$$

where each function $U_i : \mathbb{R}_+ \to \mathbb{R}$, which is understood as the utility function associated with user i, is a strictly concave, continuously differentiable nondecreasing function and c is a L-dimensional vector whose jth component represents the bandwidth of link j. As usual we assume that $U_i'(x) \to \infty$ as $x \to 0$. The relation between the role of utility functions and fairness criteria has been clarified by [1, 15]: it turns out that many practical concepts of fairness are equivalent to the right selection of utility functions. As a consequence of our assumptions the network optimization problem (4.4) has a unique solution at which all the constraints are satisfied with equality, i.e. we attain efficient usage of network resources. We use the notation $\cdot^*$ to denote the equilibrium value from (or induced by) the network optimization problem, for example p_j^* is the equilibrium congestion information on link j.

The standard approach to solve this global optimization problem (4.4) in a distributed manner is to solve the dual problem instead:

$$\min_{p_j \geq 0} \sum_{i \in [N]} \max_{x_i \geq 0} (U_i(x_i) - x_i R_{\cdot i} p) + p^T c. \tag{4.5}$$

This process decouples the coupling of the primal variables through the constraints of the original optimization problem and turns it into many small maximization problems, each of which can be handled by users with local information. The main algorithms derived from the gradient method and the Lagrangian method can be written in general form as shown below,

$$\text{Dual Law:} \begin{cases} x_i(t) = U_i'^{-1}(q_i(t)), \\ \dot{p}_j(t) = \Gamma_j(y_j(t) - c_j). \end{cases} \tag{4.6}$$

$$\text{Primal/Dual Law:} \begin{cases} \dot{x}_i(t) = K_i(U_i'(x(t)) - q_i(t)), \\ \dot{p}_j(t) = \Gamma_j(y_j(t) - c_j). \end{cases} \tag{4.7}$$

It is known that without delays both algorithms (4.6) and (4.7) converge to the optimal solution with any positive coefficients K_is and Γ_js [17]. When there are delays involved, global stability analysis of both algorithms in a heterogeneous network reveals that the stability condition depends on those coefficients in a complicated way. Although decentralized protocols exist in order to satisfy these stability conditions, they require extra communication costs and most importantly users have to reveal their own utility functions. This is illustrated by the following simple example:

Example 4.1 *A single source/link network uses the following primal/dual algo-*

rithm for its flow control

$$\dot{x}(t) = K(U'(x) - p(t - \tau)),$$

$$\dot{p}(t) = \Gamma(x - c).$$

By simple analysis it is required that $-U''(c) > 2\tau\Gamma/\pi$ *for the existence of a coefficient* K *so that the system is locally stable. In order to set the right* Γ *the router has to know the user's utility function. But it is clearly undesirable to transmit a function across the communication links, not to mention security reasons. When the user knows only her* $U(\cdot)$ *and the link knows its* c*, no one can calculate* $U''(c)$*.*

Therefore we propose the following necessary design principles for our algorithms in order to meet the needs of real-world networks

1. Equilibrium of the algorithm should solve the optimization problem (4.4);

2. The input and the parameters of user and link controllers should be obtained from local information only. For an individual user the local information is that which is accrued along the path of his flow, and for an individual link the local information is that which is aggregated from its accessing flows. Additionally each user's utility function should be only known to himself and each link's bandwidth should also be kept to itself.

3. The dynamics of the algorithm are globally asymptotically stable given heterogeneous delays.

The scalable control laws by Paganini et al [40] satisfy Principles 1)-2) and partially 3) since only linear stability is verified for their algorithm. Their algorithm takes the following form in which ζ_i is an auxiliary state variable at the user i's side:

$$\tau_i \dot{\zeta}_i = \beta_i (U_i'(x_i) - q_i),$$
$$x_i = \bar{x}_i e^{\zeta_i - \frac{\alpha_i q_i}{|r_i| \tau_i}}, \tag{4.8}$$
$$\dot{p}_j = c_j^{-1}(y_j - c_j).$$

Strictly speaking their protocol is not completely decentralized as defined in Principle 2), because their controller parameters depend on a global variable $\bar{\tau}$, which is the delay upper-bound of the whole network. Specifically in order to achieve linear stability, the following condition has to be satisfied,

$$\frac{\beta_i |r_i|}{\alpha_i} \bar{\tau} < \eta$$

for some constant η. Although this restriction might not seem to be significant, the future growth of the network may potentially require a global reset of user control coefficients and furthermore the existence of this condition on a global variable may intuitively result in slow performance due to its conservativeness. Therefore we aim at designing algorithms strictly satisfying the proposed Principles 1)-3).

4.3 General Properties of Controllers

Before we start to design a specific distributed algorithm which satisfies all the principles introduced in the previous section, we first want to understand the structural implications of controllers based on Principles 1) and 2. The reason for investigating these principles first is that to some extent they reflect the "static" characteristics our controller must possess, while Principle 3) is more relevant to its "dynamical" characteristics. It is quite difficult to make a statement about the general properties of such controllers since the controller space is a very large functional space. Therefore we resort to focus on the linearized version of both user and link controllers and the results from the linearized controllers will give us necessary conditions as well as design guidance for the full-blown nonlinear controllers in the next section. For our purpose we only consider controllers which allow a unique equilibrium state in this section.

Again consider a single user/link network with the round-trip delay τ and let $F(s)(G(s))$ be the transfer function of the user (link) controller with congestion message as the input (output) signal and flow rate as the output (input) signal. Suppose both $F(s)$ and $G(s)$ are proper rational functions. Here we made another assumption that the user (link) dynamics do not explicitly depend on her delayed value of flow rate (congestion message). This is a valid assumption since in our problem formulation delays do not bring any benefits to our goals. Then the open loop gain of the system is $e^{-\tau s}G(s)F(s)$. First we give the condition for the user controller:

Proposition 4.1 *Assume that the user dynamics (by themselves) do not involve any delays. Then the transfer function $F(s)$ of the linearized user controller is a valid user controller for the optimization problem if and only if it is stable and $F(0) = \xi^{-1}$ where $\xi = -U''(x^*)$.*

Proof: By definition we have $\delta x(s) = F(s)\delta q(s)$ where $\delta x = x - x^*$ and $\delta q = q - q^*$. Since in equilibrium $U'(x^*) = q^*$, we have $U'(x^* + \delta x) = U'(x^*) + U''(x^*)\delta x = q^* - \delta q$. Hence we conclude that $F(0) = \xi^{-1}$ is a necessary condition for

$F(s)$ to be valid. We shall show that the condition is also sufficient. Without loss of generality assume that the user controller is strictly proper and has the following form

$$F(s) = \frac{a_1 s^{n-1} + a_2 s^{n-2} + \cdots + a_{n-1}s + a_n}{s^n + b_1 s^{n-1} + \cdots + b_{n-1}s + \xi a_n}.$$

Here $a_1, \cdots, a_n$ and $b_1, \cdots, b_{n-1}$ can be functions of ξ. It is well known that this transfer function can be realized in a controller canonical form [41]

$$\begin{aligned} \begin{bmatrix} \dot{z}_1 \\ \vdots \\ \dot{z}_n \end{bmatrix} &= \begin{bmatrix} -b_1 & \cdots & -b_{n-1} & -\xi a_n \\ 1 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \\ & \quad -[a_n, 0, \cdots, 0]^T \delta q \\ \delta x &= \tfrac{a_1}{a_n} z_1 + \cdots + \tfrac{a_{n-1}}{a_n} z_{n-1} + z_n. \end{aligned}$$

This is a local version of the following nonlinear dynamics:

$$\begin{aligned} \dot{z}_1 &= -b_1(-U''(z_n))z_1 - \cdots - b_{n-1}(-U''(z_n))z_{n-1} \\ & \quad + a_n(-U''(z_n))(U'(z_n) - q), \\ \dot{z}_k &= z_{k-1}, \quad 2 \le k \le n, \\ x &= \tfrac{a_1(-U''(z_n))}{a_n(-U''(z_n))} z_1 + \cdots + \tfrac{a_{n-1}(-U''(z_n))}{a_n(-U''(z_n))} z_{n-1} + z_n \end{aligned}$$

The result then can be easily verified from the fact that the equilibrium of the system when the input is q^* is indeed $y_k = 0, 1 \le k \le n-1$ and $x = y_n = x^*$. ■

As an illustration of this proposition we can observe the correspondence between previously proposed valid user controllers and their linearized forms in Table 4.1.

Now we turn to the properties of valid link controllers.

Table 4.1: User Controllers and Their Transfer Functions

User Controller	Transfer Function
$x = U'^{-1}(q)$	$1/\xi$
$\dot{x} = K(U'(x) - q)$	$K/(s + K\xi)$
user controller by Paganini et al (4.8)	$K(s + v)/(s + Kv\xi)$

Proposition 4.2 *Assume that, like the user dynamics, the link dynamics by themselves do not involve any delays. Then the transfer function $G(s)$ of the linearized link controller is a valid link controller for the optimization problem if and only if it is stable and $G(s) = H(s)/s$ in an irreducible form where $H(s)$ is some rational transfer function.*

Proof: First we verify the sufficiency part. Suppose the equilibrium point with $F(s)$ in Proposition 4.1 and $G(s)$ given under the current form is the optimal solution of the network optimization problem. From the definition we have $\delta p(s) = G(s)\delta y(s)$ where $\delta p = p - p^*$ and $\delta y = y - y^*$. The link controller can be realized in the following form,

$$\dot{u} = y - c$$

$$\dot{v} = Av + Bu$$

$$p = Cv + Du$$

where (A, B, C, D) is a realization of the transfer function $H(s)$. Since $G(s)$ contains a pure integrator, the only input that achieves the internal stability is $\delta y = 0$, or in the realized system $y^* = c$. Along with the source controller we have the equations for the equilibrium state

$$U'(x^*) = q^* = p^*$$

$$y^* = x^* = c$$

By the KKT conditions this equilibrium point is the optimal solution of the network optimization problem.

Next we show that this integrator form is also necessary. First since only the link knows its own bandwidth c and the equilibrium point has to be $y^* = c$ for optimality, only the link controller can enforce the input δy to be zero at equilibrium. Suppose the link controller is realized as shown below,

$$\dot{z} = Az + B\delta y$$

$$\delta p = Cz + D\delta y.$$

The previous argument is equivalent to the condition $\text{rank}A < \text{rank}[A, B]$. It is sufficient to check the situation when (A, C) is observable (since (A, C) has to be detectable for stabilization, therefore the unobservable modes are asymptotically stable themselves regardless of input, so we only focus on the observable part). By a similarity transformation we can write the system in canonical observer form [41] as follows

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & \cdots & 0 & a_1 \\ 0 & \ddots & 0 & a_2 \\ 0 & \cdots & 1 & a_{n-1} \end{bmatrix},$$
$$B = [b_1, \cdots, b_n]^T,$$
$$C = [0, \cdots, 0, 1].$$

Here $A_{1n} = 0$ and $b_1 \neq 0$ due to the rank condition. Then it is straightforward to see that $G(s)$ must have a pure integrator term. ■

Remark 4.1 *The structural properties of valid user and link controllers indicated in the previous two propositions suggest that delay independent stability [20] may not be achievable given our design principles. To see this let us observe now that the open loop gain of a single user/link network can be written as* $e^{-\tau s} H(s)F(s)/s$ *where* $F(0) = \xi^{-1}$ *and* $H(0) = h$ *for some nonzero* h. *If the system is delay independent stable, then the Nyquist curve of its open loop gain should intersect the* x*-axis at points greater than -1 regardless of the value of* τ. *But it is easy to see that for sufficiently large* τ, *the Nyquist curve intersects the* x*-axis at the frequency* $\omega \approx \frac{\pi}{2\tau}$ *and the intersection point is approximately* $-\frac{2h}{\pi\xi}\tau$ *which can be made arbitrary smaller than -1. Therefore in order to achieve stability one must design the controllers based on the size of delays.*

4.4 Design of Scalable Controller

We first focus on the design of scalable controllers for a single user/link network based on previous discussions and then extend the design to arbitrary networks with heterogeneous delays.

4.4.1 The Case of Single User/Link Network

As in previous sections we denote by c the link bandwidth and by $\tau = \tau^f + \tau^b$ the round-trip delay. Similar to the user controller in Paganini et al's algorithm (4.8), we can choose the transfer function of our user controller to be

$$F(s) = \frac{s + k/\tau}{\tau s + \xi k/\tau} \tag{4.9}$$

and our link controller to be

$$G(s) = \frac{1}{s}. \tag{4.10}$$

Here ξ is defined as $-U''(c)$ as in Proposition 4.1 and k is some constant. First by direct calculation we have

Lemma 4.1 *The single user/link network with the user controller given by (4.9) and the link controller (4.10) is linearly asymptotically stable for arbitrary τ and c if $0 < k \leq k_0 \approx 0.5474$. Here $k_0 \triangleq \omega_0/\tan\omega_0$ where $\omega_0 \in (0, \pi/2)$ is the solution of the equation $\omega \sin\omega = 1$.*

From the proof of Proposition 4.1 we can realize our algorithm from its linearized form as follows:

$$\begin{aligned} \dot{z} &= \tfrac{k}{\tau^2}(U'(x) - p(t-\tau^b)), \\ x &= z - \tfrac{p(t-\tau^b)}{\tau}, \\ \dot{p} &= x(t-\tau^f) - c. \end{aligned} \tag{4.11}$$

Remark 4.2 *It is worth discussing the initial dynamics of the above system. Since there is no guarantee at the beginning from $x = z - p(t-\tau^b)/\tau$ such that x is kept positive, we have to resort to other means. A feasible solution to the initial dynamics is as follows,*

$$\begin{cases} x(t) = z - p(t-\tau^b)/\tau, & \textit{if } z > p(t-\tau^b)/\tau, \\ \dot{x}(t) = -\alpha x(t), & \textit{otherwise}, \end{cases}$$

for any positive constant α. Since from our dynamics (4.11) $p(t)$ is a continuous function of time, it is easy to see that once $x(t) > 0$, it stays positive thereafter. So the dynamics of x will be of the form $\dot{x} = -\alpha x$ for at most a finite time duration at the beginning of the algorithm. This period can be regarded as a "probing" phase of the flow dynamics. Therefore it is sufficient for us to consider only the dynamics (4.11) thereafter.

The global stability of the system (4.11) can be studied from the observation that the system is actually of Lur'e type [42] by rewriting it into an equivalent form as follows

$$\begin{aligned} \dot{x} &= -\tfrac{k}{\tau^2}(p(t-\tau^b) - p^*) - \tfrac{1}{\tau}(x(t-\tau) - x^*) + \tfrac{k}{\tau^2}u, \\ \dot{p} &= x(t-\tau^f) - c, \\ u &= U'(x) - p^*. \end{aligned}$$

Taking u as the input signal and x as the output signal, the transfer function from u to x is

$$L(s) = \frac{k}{\tau^2}\left(s + \frac{k+\tau s}{\tau^2 s}e^{-\tau s}\right)^{-1},$$

while the mapping from x to u is a $(0,\infty)$-sector nonlinear mapping. In order to obtain nonlinear stability of (4.11) by Popov's criterion [43, 44] it remains to show that there exists $\eta \in \mathbb{R}$ such that $(1+\eta s)L(s)$ is positive real.

Lemma 4.2 $(1+\tau s/2)L(s)$ *is positive real when* $0 < k \leq 1/2$.

Proof: By Lemma 4.1 we only need to check whether $\Re(1+\tau i\omega/2)L(i\omega) \geq 0$ and this in turn is equivalent to whether $\Re(1+\tau i\omega/2)^{-1}L(i\omega)^{-1} \geq 0$. Hence the proof reduces to showing that

$$\frac{1}{2}\theta(\theta^2 - k\cos\theta - \theta\sin\theta) - k\sin\theta + \theta\cos\theta \geq 0 \tag{4.12}$$

where $\theta \triangleq \omega\tau$.

When $k = 1/2$ the above inequality (4.12) becomes

$$\frac{1}{2}\theta(\theta^2 - \frac{1}{2}\cos\theta - \theta\sin\theta) - \frac{1}{2}\sin\theta + \theta\cos\theta \geq 0 \tag{4.13}$$

which is correct by checking it with numerical means.

If $0 \leq \theta \leq \theta_0 \approx 2.2889$, in which θ_0 is the smallest positive solution of the equation $\theta\cos\theta + 2\sin\theta = 0$, we have

$$\frac{1}{2}\theta\cos\theta + \sin\theta \geq 0.$$

But the left hand side of the above inequality is exactly the difference of the left hand sides of the inequalities (4.12) and (4.13) times $\frac{1}{2} - k$. Thus the inequality

(4.12) holds for $0 \le \theta \le \theta_0$. Now consider the situation when $\theta > \theta_0$. In this case the left hand side of (4.12) is lower bounded by

$$\frac{1}{2}\theta(\theta^2 - \frac{1}{2} - \theta) - \theta - \frac{1}{2}.$$

One can directly check that this cubic polynomial achieves its minimum over $\theta \ge \theta_0$ at $\theta = \theta_0$, and that the minimum is positive. Therefore we conclude that the inequality (4.12) holds for all θ and $(1 + \tau s/2)L(s)$ is positive real. ■

Then from Lemma 4.2 and Popov's criterion we immediately have:

Proposition 4.3 *With the initial dynamics discussed in Remark 4.2, the system (4.11) is globally asymptotically stable for arbitrary values of* τ *and* c *if* $k \in (0, 1/2]$.

So we obtain a scalable controller which satisfies all the design principles in Section 4.2 for a single user/link network.

4.4.2 The Case of General Network

A direct extension of the user and link controllers (4.9-4.10) from the previous subsection to the situation of a general network with heterogeneous delays is

$$F_i(s) = \frac{s + k/\tau_i}{\tau_i |r_i| s + \xi_i k/\tau_i} \tag{4.14}$$

for user i and

$$G_j(s) = \frac{1}{|f_j| s} \tag{4.15}$$

for link j.

Define a $L \times N$ matrix-valued function $\hat{R}(s)$ on the frequency domain by

$$\hat{R}_{ji}(s) = R_{ji} e^{-\tau_{ij}^f s}$$

then the relation between the flow rate vector x and the aggregate rate vector y (4.1) can be written as

$$y(s) = \hat{R}(s)x(s).$$

From the definition of the round-trip delays (4.3), the relation between the congestion message vector p and the aggregate congestion vector q (4.2) can be equivalently expressed by

$$q(s) = \mathrm{diag}(\{e^{-\tau_i s}\})\hat{R}^H(s)p(s).$$

where $\hat{R}^H$ is the Hermitian of $\hat{R}$.

Therefore combining these equations the open loop gain of the network system with tentative controllers (4.14-4.15) is given as follows

$$\begin{aligned} L(s) &= \mathrm{diag}\left(\left\{\tfrac{s+k/\tau_i}{\tau_i|r_i|s+\xi_i k/\tau_i} e^{-\tau_i s}\right\}\right) \\ &\quad \times \hat{R}^H(s)\,\mathrm{diag}\left(\left\{\tfrac{1}{|f_j|s}\right\}\right)\hat{R}(s). \end{aligned}$$

It would be desirable that this natural extension from the single user/link network (4.14-4.15) simply gives us stabilizing controllers for general networks. To examine this we need to study the eigenloci of the matrix $L(s)$, per the Generalized Nyquist Theorem [45]. Recall an elegant result by Vinnicombe [10]:

Lemma 4.3 (Vinnicombe) *Assume* $\Lambda = \mathrm{diag}(\{\lambda_i\})$ *and* $M = M^T \geq 0$ *are* $N \times N$ *matrices. Then the eigenvalues of* ΛM $\sigma(\Lambda M) \in \rho(M)\bar{co}(\{0, \lambda_1, \cdots, \lambda_N\})$. *Here* $\rho(\cdot)$ *denotes the spectral radius and* $\bar{co}(\cdot)$ *denotes the convex hull.*

Note that

$$\sigma(L(s)) = \sigma\left(\mathrm{diag}(\{l_i(s)\})M(s)\right),$$

where we define

$$\begin{aligned} l_i(s) &\triangleq \frac{s + k/\tau_i}{s(\tau_i s + \xi_i k/(\tau_i |r_i|))} e^{-\tau_i s}, \\ M(s) &\triangleq \operatorname{diag}(\{|r_i|^{-1/2}\}) \hat{R}^H(s) \operatorname{diag}(\{|f_j|^{-1}\}) \\ &\quad \times \hat{R}(s) \operatorname{diag}(\{|r_i|^{-1/2}\}). \end{aligned} \tag{4.16}$$

We first calculate the upper bound of the spectral radius $\rho(M(i\omega))$ of $M(i\omega)$:

$$\begin{aligned} &\rho(M(i\omega)) \\ = \; &\rho(\operatorname{diag}(\{|r_i|^{-1}\}) \hat{R}^H(i\omega) \operatorname{diag}(\{|f_j|^{-1}\}) \hat{R}(i\omega)) \\ \leq \; &\sup_{i,\omega} \sum_{j,n} |r_i|^{-1} |\hat{R}_{ji}(i\omega)| |f_j|^{-1} |\hat{R}_{jn}(i\omega)| \\ = \; &1 \end{aligned}$$

from the definitions of $|r_i|$ and $|f_j|$. Therefore a sufficient condition for linear stability is

$$-1 \notin \bar{\text{co}}(\{0, l_1(i\omega), \cdots, l_N(i\omega)\})$$

by Lemma 4.3 and the Generalized Nyquist Theorem. Since the Nyquist curves $l_i(i\omega)$ with arbitrary τ_i and ξ_i are bounded by a single curve $l(\theta)$ on the Nyquist plane:

$$l(\theta) \triangleq -\frac{i\theta + k}{\theta^2} e^{-i\theta},$$

we only need to check whether

$$-1 \notin \bar{\text{co}}(0 \cup \{l(\theta), \forall \theta \geq 0\}).$$

However since $\angle l(\theta) \to -180^\circ$ with $\theta \to 0$ and part of the curve $l(\theta)$ lies on the second quadrant of the Nyquist plane, the convex hull of curve $l(\theta)$ contains

-1. Therefore we cannot guarantee linear stability with controllers (4.14-4.15) from Lemma 4.3. In fact, we are able to construct a 7-user 5-link network with controllers (4.14-4.15) such that its flow dynamics is linearly unstable. Consider a network with its routing matrix R as

$$R = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix},$$

the round-trip time vector τ for each source as

$$\tau = [0.0483, 0.1155, 0.0340, 0.7009, 0.0612, 0.9030, 0.6545]^T,$$

and the forward delay matrix τ^f as

$$\tau^f = \begin{bmatrix} 0 & 0.0781 & 0 & 0.4485 & 0 & 0.6293 & 0 \\ 0.0351 & 0.1105 & 0.0145 & 0 & 0.0013 & 0 & 0.2750 \\ 0 & 0 & 0.0013 & 0 & 0 & 0 & 0.1688 \\ 0.0358 & 0.0086 & 0.0192 & 0 & 0.0466 & 0 & 0 \\ 0 & 0 & 0.0310 & 0.6816 & 0 & 0 & 0.4043 \end{bmatrix}.$$

We set the factor $k = 0.54$ in (4.14) for the source controller, which satisfies the condition in Lemma 4.1 for the linear stability of a single source/link network. The resulting Nyquist plot of the open loop system is shown in Figure 4.1. It can be observed that -1 is encircled by one of its eigenloci. Therefore we have linear instability of the network flow control system. This is the main reason why the controllers (4.8) proposed by Paganini et al have to rely on a global variable $\bar{\tau}$.

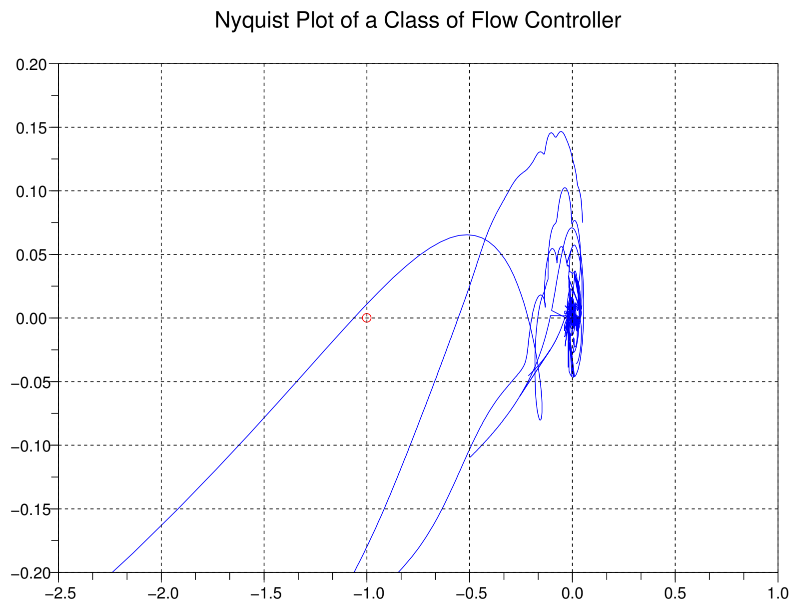

Figure 4.1: Nyquist plot of a 7-user 5-link network with source and link controllers specified in (4.14-4.15).

To this end we consider a class of source controller which is a generalization of (4.14) as follows,

$$T_i(s) = \frac{s + a(\tau_i)}{b(\tau_i)|r_i|s + a(\tau_i)\xi_i}, \tag{4.17}$$

where $a(\cdot)$ and $b(\cdot)$ are both functions from $\mathbb{R}^+$ to $\mathbb{R}^+$. We prove below that the above source controller with

$$a(\tau) = \frac{1}{4(1+\tau)}, \quad b(\tau) = 4(\tau^2 + 1), \tag{4.18}$$

along with our original link controller (4.15), provide a pair of valid stabilizing linear controllers for our optimization problem (4.4). In fact we can prove a slightly strong result,

Lemma 4.4 *The flow dynamics of a network with heterogeneous delays where each* *user controller is given by (4.17), in which the functions* $a(\cdot)$ *and* $b(\cdot)$ *both satisfy*

$$a(\tau) \leq \frac{\min\{\tau, 1\}}{2\tau}, \quad \forall \tau \geq 0$$

$$b(\tau) \geq 2\tau \max\{\tau, 1\}, \quad \forall \tau \geq 0,$$

and each link controller is given by (4.15) are linearly asymptotically stable.

Proof: See Appendix B. ∎

Clearly the condition of the above lemma holds for both functions $a(\cdot)$ and $b(\cdot)$ set in (4.18), the linear stability of the network follows immediately.

By Proposition 4.1 one can realize the linear controllers (4.17) and (4.15) by

$$\begin{aligned} \dot{z}_i &= \tfrac{1}{16(1+\tau_i)(1+\tau_i^2)|r_i|}(U_i'(x_i) - q_i), \\ x_i &= z - \tfrac{q_i}{4(1+\tau_i^2|r_i|}, \\ \dot{p}_j &= \tfrac{y_j - c_j}{|f_j|}. \end{aligned} \tag{4.19}$$

We also assume that for the above system we adopt initial dynamics similar to that in Remark 4.2, so that after an initial phase the system stays in the correct region of $x_i > 0$ with the above dynamics forever.

We proceed to show the global stability by using multiplier method. Like in the single user/link network, take the nonlinear feedback controllers of the form

$$u_i(x_i) = \frac{1}{16(1+\tau_i)(1+\tau_i^2)|r_i|}(U_i'(x_i) - q_i^*),$$

then the output transfer function of (4.19) becomes

$$W(s)^{-1} \triangleq \left(sI + \text{diag}\left(\left\{ \frac{s + \frac{1}{4(\tau_i+1)}}{4(1+\tau_i^2)|r_i|s} e^{-\tau_i s} \right\} \right) \hat{R}(s)^H \, \text{diag}(\{|f_j|^{-1}\}) \hat{R}(s) \right)^{-1}.$$

Unlike the single user/link network, we will use a stronger class of multipliers than the Popov multiplier. Since the utility functions $U_i(\cdot)$ are concave, $U_i'(\cdot)$s are monotone functions and so are the feedback controllers $u_i(\cdot)$ with respect to x_i. Therefore we should be able to use Zames-Falb multiplier [46] to prove the nonlinear stability of the system. We select the following multiplier

$$M(s)^{-1} \triangleq \text{diag}\left(\left\{\frac{s+\frac{1}{4(\tau_i+1)}}{4(1+\tau_i^2)|r_i|((\tau_i+1)s+1)}\right\}\right).$$

We argue that the multiplier $M(s)^{-1}$ is indeed a Zames-Falb multiplier. This can be observed directly from the fact that the zero of each diagonal entry of $M(s)^{-1}$, $-1/(4(\tau_i+1))$, is less than the pole, $-1/(\tau_i+1)$, in absolute value.

In order to prove global nonlinear system of the system (4.19), it is sufficient to show that

$$W(s)^{-1}M(s)^{-1} + (M(s)^{-1})^H(W(s)^{-1})^H \geq 0, \quad \forall s = i\omega.$$

Since

$$W^{-1}M^{-1} + (M^{-1})^H(W^{-1})^H = W^{-1}M^{-1}(MW + W^HM^H)(M^{-1})^H(W^{-1})^H,$$

it is sufficient to show the following holds true for all $s = i\omega$,

$$N(s) + N(s)^H \triangleq M(s)W(s) + W(s)^HM(s)^H \geq 0.$$

Lemma 4.5 *The matrix-valued function $N(s)$ defined above satisfies*

$$N(s) + N(s)^H \geq 0$$

for all $s = i\omega$.

Before we prove Lemma 4.5, let us first discuss how to make a rank 1 matrix to become positive definite by adding a diagonal matrix. It is known that the $n \times n$ matrix $A \triangleq \text{diag}(\{a_i\})\mathbf{1}\mathbf{1}^T$ is nonnegative if and only if $a_i \equiv a$ for any i. Since for any vector v, we have

$$v^H(A + A^H)v = 2\Re\left(\sum_i v_i\right)^*\left(\sum_i a_i v_i\right)$$

which is clearly not a sum of squares unless $a_i \equiv a$. For reasons to be clear later, we are interested in adding a diagonal matrix $D = \text{diag}(\{d_i\})$ where d_i depends only on a_i, so that $A + D$ is positive definite. We have the following result

Lemma 4.6 *For any $\{d_i\}$, if there exists a real constant c, such that*

$$d_i \geq \frac{1}{2}\left(ca_i - \frac{n-2}{2c}\right)^*\left(ca_i - \frac{n-2}{2c}\right) + \frac{n-1}{2c^2} - \Re a_i, \forall i = 1, \cdots, n,$$

then the $n \times n$ matrix $D + A = \text{diag}(\{d_i\}) + \text{diag}(\{a_i\})\mathbf{1}\mathbf{1}^T$ is positive definite.

Proof: It suffices to prove positive definiteness of $D + A$ when

$$d_i = \frac{1}{2}\left(ca_i - \frac{n-2}{2c}\right)^*\left(ca_i - \frac{n-2}{2c}\right) + \frac{n-1}{2c^2} - \Re a_i, \forall i = 1, \cdots, n.$$

For any vector v, let us compute the quadratic form,

$$
\begin{aligned}
& v^H(D + A + A^H + D^H)v \\
=& \sum_{i=1}^{n}(2d_i + 2\Re a_i)v_i^* v_i + \sum_{\substack{i,j=1,\\ i\neq j}}^{n}(a_i^* + a_j)v_i^* v_j \\
=& \sum_{i=1}^{n}\left(\left|ca_i - \frac{n-2}{2c}\right|^2 + \frac{n-1}{c^2}\right)v_i^* v_i + \sum_{\substack{i,j=1,\\ i\neq j}}^{n}(a_i^* + a_j)v_i^* v_j \\
=& \sum_{i=1}^{n}\left(\left|ca_i - \frac{n-2}{2c}\right|^2 + \frac{n-1}{c^2}\right)v_i^* v_i + \sum_{\substack{i,j=1,\\ i\neq j}}^{n}((ca_i^* v_i^*)(v_j/c) + (v_i^*/c)(ca_j v_j)) \\
=& \sum_{i=1}^{n}\left|\left(ca_i - \frac{n-2}{2c}\right)v_i + \sum_{j\neq i}\frac{v_j}{c}\right|^2 \\
\geq& 0.
\end{aligned}
$$

Therefore $D + A$ is positive definite and we reach the conclusion. ■

The value of d_i given in the above lemma still depends on the size of the matrix A. It is useful to obtain a size independent d_i such that $D + A$ remains positive definite. For this purpose we have to scale the entries of A down proportional to its size n. We have the following corollary of Lemma 4.6

Corollary 4.1 *Define $n \times n$ matrix $A_n \triangleq \frac{1}{n}\operatorname{diag}(\{a_i\})\mathbf{1}\mathbf{1}^T$. For any $\{d_i\}$ such that*

$$d_i \geq \frac{1}{2}|a_i|^2 - \Re a_i + \frac{1}{2}, \quad \forall i = 1, \cdots, n,$$

the matrix $D + A_n = \operatorname{diag}(\{d_i\}) + A_n$ is positive definite.

Proof: According to Lemma 4.6 we only have to show that there exist a real constant $c(n)$ such that

$$\frac{n}{2}|a_i|^2 - n\Re a_i + \frac{n}{2} \geq \frac{1}{2}\left|ca_i - \frac{n-2}{2c}\right|^2 + \frac{n-1}{2c^2} - 2\Re a_i,$$

for all a_i and n. Subtract the right-hand side from the left-hand side of the above inequality, we have

$$\text{LHS} - \text{RHS} = \frac{n - c^2}{2}|a_i|^2 - \frac{n}{2}\Re a_i + \frac{n}{2} - \frac{n^2}{4c^2} = \frac{n}{2}(|a_i|^2 - 2\Re a_i + 1) \geq 0$$

where in the second equality we replace c by $\sqrt{n/2}$. Therefore our conclusion holds.

■

Now we are ready to prove Lemma 4.5. We will use the condition given by Corollary 4.1 as a distributed test for the positive definiteness of a class of matrices which can be represented by the sum of a diagonal matrix and a rank 1 matrix.

Proof: [Proof of Lemma 4.5] From the definition we have

$$\begin{aligned} N(s) =& \operatorname{diag}\left(\left\{16(1+\tau_i)(1+\tau_i^2)|r_i|\frac{s((1+\tau_i)s+1)}{4(1+\tau_i)s+1}\right\}\right) \\ &+ \operatorname{diag}\left(\left\{\frac{e^{-\tau_i s}((1+\tau_i)s+1)}{s}\right\}\right)\hat{R}(s)^H \operatorname{diag}(\{f_j^{-1}\})\hat{R}(s). \end{aligned}$$

For each link j, we define

$$\begin{aligned} N_j(s) \triangleq& \operatorname{diag}\left(\left\{16(1+\tau_i)(1+\tau_i^2)\frac{s((1+\tau_i)s+1)}{4(1+\tau_i)s+1}R_{ji}\right\}\right) \\ &+ \operatorname{diag}\left(\left\{\frac{e^{-\tau_i s}((1+\tau_i)s+1)}{s}R_{ji}\right\}\right)\hat{R}(s)_{j\cdot}^H f_j^{-1}\hat{R}(s)_{j\cdot}. \end{aligned}$$

Apparently $N(s) = \sum_j N_j(s)$. The plan is to show $N_j(s) + N_j(s)^H \geq 0$ and then by the argument that the sum of positive Hermitians remains positive, we will reach the conclusion of the lemma.

For convenience we define

$$\begin{aligned} d_{ji} &\triangleq 16(1+\tau_i)(1+\tau_i^2)\frac{s((1+\tau_i)s+1)}{4(1+\tau_i)s+1}R_{ji}, \\ a_{ji} &\triangleq \frac{e^{-\tau_i s}((1+\tau_i)s+1)}{s}R_{ji}. \end{aligned}$$

Next we show that $N_j(s) + N_j(s)^H \geq 0$ is equivalent to $\tilde{N}_j(s) + \tilde{N}_j(s)^H \geq 0$ where $\tilde{N}_j(s)$ is defined by

$$\tilde{N}_j(s) \triangleq \text{diag}(\{d_{ji}\}) + \text{diag}(\{a_{ji}\})(R_{j\cdot})^T f_j^{-1} R_{j\cdot}.$$

This is because $N_j(s) + N_j(s)^H \geq 0$ if and only if $v^H(N_j(s) + N_j(s)^H)v \geq 0$ for all v such that $v^H v = 1$. Or equivalently

$$\sum_i (d_{ji} + d_{ji}^* + a_{ji} + a_{ji}^*)v_i^* v_i + \sum_{i \neq k} f_j^{-1}(a_{ji} + a_{ki}^*)e^{\tau_{ij}^f s} v_i^* e^{-\tau_{kj}^f s} v_k \geq 0.$$

If we define $\tilde{v}_i \triangleq v_i e^{-\tau_{ij}^f s}$, we still have $\tilde{v}^H \tilde{v} = 1$ and the left-hand side of the above inequality becomes

$$\sum_i (d_{ji} + d_{ji}^* + a_{ji} + a_{ji}^*)\tilde{v}_i^* \tilde{v}_i + \sum_{i \neq k} f_j^{-1}(a_{ji} + a_{ki}^*)\tilde{v}_i^* \tilde{v}_k$$

which is exactly $\tilde{v}^H(\tilde{N}_j(s) + \tilde{N}_j(s)^H)\tilde{v}$. Therefore we only need to focus on the proof of $\tilde{N}_j(s) + \tilde{N}_j(s)^H \geq 0$.

Since the matrices $\text{diag}(\{d_{ji}\})$ and $\text{diag}(\{a_{ji}\})$ are essentially of size $|f_j|$, by Corollary 4.1 we only need to prove the following inequality,

$$d_{ji} + d_{ji}^* \geq a_{ji}^* a_{ji} - a_{ji} - a_{ji}^* + 1. \tag{4.20}$$

The left-hand side of the above inequality (4.20) is

$$f(\omega) \triangleq \frac{96(1+\tau_i)^2(1+\tau_i^2)\omega^2}{1+16(1+\tau_i)^2\omega^2}$$

and the right-hand side of the above inequality (4.20) is

$$g(\omega) \triangleq 1 + (1+\tau_i)^2 - 2(1+\tau_i)\cos\tau_i\omega + 2\frac{1-\cos\tau_i\omega - \tau_i\omega\sin\tau_i\omega}{\omega^2}.$$

The maximally achievable value of $g(\cdot)$ is $(1+\tau_i)^2+\tau_i^2$, while the value of $f(\cdot)$ is greater than $3(\tau_i+1)^2$ when $\omega \geq (4(1+\tau_i))^{-1}$. Therefore the inequality (4.20) holds for $\omega \geq (4(1+\tau_i))^{-1}$. Now let us look at the interval $\omega \in [0,(4(1+\tau_i))^{-1}]$. The maximum derivative of ω of $g(\omega)$ is $0.62\tau_i^3+2\tau_i(1+\tau_i)$, where 0.62 is the maximum of

$$\frac{d}{d\theta}\left(2\frac{1-\cos\theta-\theta\sin\theta}{\theta^2}\right)=2\frac{2(\cos\theta+\theta\sin\theta-2)-\theta^2\cos\theta}{\theta^3}.$$

But the minimum derivative of ω of $f(\omega)$ is $6(1+\tau_i^2)(1+\tau_i)$, which is clearly greater than $0.62\tau_i^3+2\tau_i(1+\tau_i)$ for all positive τ_i. Therefore the inequality 4.20 holds. By previous discussions $N(s)$ is positive definite. ■

Now by similar arguments as [46, 43, 44] we finally reach the conclusion:

Theorem 4.1 *Along with the initial dynamics introduced in Remark 4.2, the network optimization algorithm given by (4.19) is globally asymptotically stable and thus satisfies all the design principles proposed in Section 4.2.*

4.5 Conclusions

We have succeeded in designing a scalable and distributed algorithm for the network optimization problem as promised at the beginning of the paper. We believe that our definition of the problem reflects the real meaning of the plug-and-play property for the network flow control problem and to the authors' knowledge our algorithm is the first to achieve this goal: to obtain efficient and fair bandwidth allocation for a network with the presence of delays and with minimum extra communication cost. We also believe that our approach presents a design methodology from which people can create various algorithms to meet different performance requirements while still maintaining the basic plug-and-play property, and our algorithm is just the simplest one in this class. Still many basic questions must be solved. For example we have not yet dealt with the case of time-varying delay. Also our analysis is based on fluid models of flow control mechanisms so it is interesting to see how to design real communication protocols based on our algorithm. We will address these questions in our future research.

Chapter 5

Design of Scalable Control Laws for Combined Routing and Flow Control - The Case of Multiple Path Routing

5.1 Introduction

In previous chapters we have discussed in detail the network congestion control mechanisms under the framework of network optimization in the sense of welfare maximization, which was proposed initially by Kelly et al [1]. We have emphasized our study on the stability of the various congestion control protocols, a dynamical property which indicates whether the rates of network flows will converge to the fair allocated rates set by the network welfare maximization problem. Especially, we have focused on the congestion control algorithms that are truly distributed, in a sense that for any source or link, all the parameters associated with its controller can be obtained via minimum interactions with other components of the network. This localization feature of the congestion controllers, along with the decoupling property of the network optimization, enables each source and link to act independently of each other in a large complex network to achieve global efficiency.

In many situations the network efficiency achieved by this mechanism can still be improved considerably by choosing the right paths for network flows. So far we have only considered the situation where each user has a single fixed path to send his packets from source to destination. But in many situations there are multiple paths available for the source/destination pair and a single path is selected from them by routing protocols which may not take into consideration of network congestions. For example, the prevalent interdomain routing protocol BGP is mainly policy based and the resulting route selection does not necessarily align with the status of link congestion: it can choose a low bandwidth link for a source/destination pair even though a high bandwidth link is available. In fact, this can potentially reduce the connection-level stability region of the network [47]. Therefore it is important to do combined routing and congestion control to fully utilize the existing network resources.

There have been much research efforts for designing adaptive routing protocols which adapt the route selection iteratively to a changing network to achieve better performance. But most of them is "open-loop" in a sense that the decision of the routing protocol is based on the measurements of bandwidth usage over a long duration of time and the resulting route change takes place in a much slower time scale than the actual congestion dynamics. But the problem is essentially "closed-loop", since the routing decision is based on the link congestion level and conversely the link congestion depends on the flows routed by the protocol. A routing protocol with time-scale separation will likely lead to route/rate instability. Therefore we consider combined routing and congestion control protocols in which the routing decision and the congestion control occur in the same time-scale. Specifically, we consider the situation when each user can distribute his flow over his available paths simultaneously. This is termed *multi-path* routing in literature. We assume that for each user, his available paths are known to him at the beginning of the transmission and the intermediate routers are capable of forwarding these multi-path packets using source routing. It is envisaged in [48] that the multipath routing can be implemented using IPv6 in a multi-homing environment with stepping-stone routers.

Our contribution in this chapter is to design scalable combined routing and congestion control algorithm, which is a direct application of the design techniques introduced in Chapter 4. Specifically our algorithm has the property of true decentralization just as the case of single flow algorithm in Chapter 4. Additionally it guarantees the global stability of the network instead of local stability obtained in [49, 50, 51]. Our algorithm also provides an interesting but sensible example that delay terms are artificially introduced in the control loop in order to achieve stability. The organization of the chapter follows like this. We introduce our network model in Section 5.2 and summarize the previous studies related to the subject in Section 5.3. The main result is in Section 5.4 and we conclude the Chapter in Section 5.5.

5.2 Network Model

The network model studied in this chapter is similar to those in the previous chapters. We consider a network with N users and L bottleneck links whose bandwidth will be fully utilized at equilibrium. Each user is associated with a source/destination pair between which he intends to transmit his packets. Again we only consider flows with infinite length. Different than the scenarios described in previous chapters, now each user may have more than one paths to send his traffic.

By path we mean concatenation of a set of links from source to destination. We assume that the available paths are known to each user in the beginning and the network is set up so that the user can send his traffic through all his available paths at the same time. Denote the number of paths for user i by M_i and the number of all paths by $M = \sum_i M_i$. Then each path is indexed from 1 to M and we associate each user i with set of available paths denoted by m_i. For convenience, we order path indices by user indices, that is, $m_i = \{\sum_{l=1}^{i-1} M_l + 1, \cdots, \sum_{l=1}^{i} M_l\}$. We can also represent the user/path relation by a $N \times M$ 0-1 matrix H, called multipath matrix, as follows,

$$H_{il} \triangleq \begin{cases} 1, & i \in m_i, \\ 0, & i \notin m_i. \end{cases}$$

From our path index rule, the matrix H is of the following form,

$$H = \begin{bmatrix} \mathbf{1}_{M_1}^T & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{1}_{M_N}^T \end{bmatrix},$$

where we denote M dimensional all 1 column vector as $\mathbf{1}_M$. In the case of single path routing H is an identity matrix. Each path l consists of a set $r_l \subset [L]$ of bottleneck links and each link j is accessed by a set $f_j \subset [M]$ of paths. We do not require the available paths for a single user to be disjoint from each other. The routing matrix R is a $M \times L$ 0-1 matrix with usual definition. Similar to previous chapters, we use the notations of x_l, q_l, y_j, p_j, and c_j to denote the flow rate of path l, aggregate feedback congestion information of path l, aggregate arriving rate at link j, the congestion information at link j, and the bandwidth of link j respectively.

Again we assume for each path i and link j, there is a forward propagation delay τ_{lj}^f and backward propagation delay τ_{lj}^b, if $R_{jl} = 1$. All delays are assumed to be fixed. The relations between x_l and y_j, p_j and q_l can all be represented by (4.1) and (4.2). The total transmission rate for user i is denoted by z_i which satisfies the following equation

$$z_i(t) = \sum_{l=1}^{M} H_{il} x_l(t), \forall i \in [N].$$

We are concerned with the following network optimization problem for combined routing and congestion control,

$$\begin{aligned} &\max\nolimits_{x_i \geq 0} \textstyle\sum_{i \in [N]} U_i(z_i) \\ &\text{s.t. } Hx = z, \\ &Rx \leq c. \end{aligned} \tag{5.1}$$

The utility function $U_i : \mathbb{R}^+ \to \mathbb{R}$ is a strictly concave, continuously differentiable nondecreasing function. As usual we assume that $U_i'(x) \to \infty$ as $x \to 0$. An important difference between the above optimization problem (5.1) and the single path network optimization problem (3.1) is that we no longer have strict concavity in the objective function, although U_i remains strictly concave for all i, as in (3.1). One immediate consequence of this property is that we may not have unique solutions:

Proposition 5.1 *The solution to the network optimization problem (5.1) associated with combined routing and congestion control is unique if and only if the* $(N+L) \times M$ *matrix*

$$\begin{bmatrix} H \\ R \end{bmatrix}$$

has full column rank. In addition, all optimal solutions have the same user rates $\{z_i\}$.

Proof: By the rate constraints it follows directly that the feasible set for $\{z_i\}$ is a polyhedron in $\mathbb{R}_+^N$. Since the utility function U_i is a strictly concave function of z_i for every i, the optimization problem (5.1) admits a unique optimal solution $\{z_i^*\}$. Then from the rate constraints the optimal path rates $\{x_l^*\}$ satisfies the following linear equations:

$$\begin{bmatrix} H \\ R \end{bmatrix} x^* = \begin{bmatrix} z^* \\ c \end{bmatrix},$$

which has unique solution if and only if the coefficient matrix has full column rank.

■

In addition to the non-uniqueness of the equilibrium rates, a major difficulty in deriving distributed algorithm for solving (5.1) is that the dual problem is no longer differentiable. This is the issue we will discuss in detail in the next section and we will present our design of algorithms to overcome this problem in Section 5.4.

5.3 Literature Overview

Recall the original design of decentralized algorithms in solving single path optimization problem (3.1) by [1, 17] is based on two indispensable properties of the original problem, separability and strict concavity, so that each user is allowed to use gradient method to solve his dual problem independently. The combined routing and flow control problem presented in (5.1) inherits the separability property and it is Kelly et al [1] who presented the first decentralized algorithm to solve a related optimization problem. The dual optimization problem associated with our original problem is

$$\min_{p_j \geq 0} \max_{x_l \geq 0} \sum_{i \in [N]} \left(U_i \left(\sum_{l \in [M]} H_{il} x_l \right) - \sum_{l \in m_i} \sum_{j \in [L]} R_{jl} x_l p_j \right) + \sum_{j \in [L]} p_j c_j. \tag{5.2}$$

Using the same gradient method as in the case of single path flow control, one can propose immediately [52] the following decentralized algorithm to solve (5.2) and equivalently (5.1):

$$\text{Dual Law:} \begin{cases} x_l = \text{argmax}_{x_{l:l \in m_i} \geq 0} \left(U_i \left(\sum_{l \in m_i} x_l \right) - x_l \sum_{j \in [J]} R_{jl} p_j \right), & \forall l \in [M], \\ \dot{p}_j = \Gamma_j \left(\sum_{l \in [M]} R_{jl} x_l - c_j \right), & \forall j \in [L]. \end{cases}$$

and

$$\text{Primal/Dual Law:} \begin{cases} \dot{x}_l = K_l (U_{i:l \in m_i}(z_{i:l \in m_i}) - q_l), & \forall l \in [M], \\ \dot{p}_j = \Gamma_j (y_j - c_j), & \forall j \in [L]. \end{cases}$$

However, a close examination of the dual objective function

$$L(p) \triangleq \max_{x_l \geq 0} \left(\sum_{i \in [N]} U_i \left(\sum_{l \in [M]} H_{il} x_l \right) - \sum_{j \in [L]} R_{jl} x_l p_j \right) + \sum_{j \in [L]} p_j c_j$$

reveals that $L(p)$ is not everywhere differentiable. This can be shown as follows. Define the function $V_i : \mathbb{R}_+ \to \mathbb{R}$ by $V_i(p) \triangleq \max_{x \geq 0} U_i(x) - px$. Then V_i is a strictly convex nonincreasing function. It can be computed directly that the dual objective function is

$$L(p) = \sum_{i \in [N]} V_i \left(\min_{l \in m_i} \sum_{j \in [L]} R_{jl} p_j \right) - \sum_{j \in [L]} p_j c_j.$$

Since $\min\{x, y\}$ is nondifferentiable at $x = y$, $L(p)$ is also nondifferentiable when there exist i such that

$$\exists l', l'' \in m_i, l' \neq l'', \sum_{j \in [L]} R_{jl'} p_j = \sum_{j \in [L]} R_{jl''} p_j = \min_{l \in m_i} \sum_{j \in [L]} R_{jl} p_j.$$

That is, $-p$ is not the gradient of $L(p)$ but only belongs to its subgradient at these values of p. This comes with no surprise since according to Danskin's Theorem [53] the dual objective function is differentiability if and only if the primal objective function is strictly concave. A direct consequence is that even in a delay-free environment, the flow dynamics of the dual algorithm tend to "chatter" forever, even though the total rates of all the users converge to the solution of (5.1), since every user only sends all of his traffic to the least congested path and this in turn causes that path to become the most congested and so on. The primal/dual algorithm for the system without delay does not converge either. We will use the following version of LaSalle's Invariance Principle in the process,

Theorem 5.1 (LaSalle's Principle, Proposition 5.22 in [54]) *Suppose* $x(\cdot) : \mathbb{R}_+ \to \mathbb{R}^n$ *is the solution to the differential equation* $\dot{x(t)} = f(x)$ *with initial condition* $x(0) = x_0$*. Let* $v : \mathbb{R}^n \to \mathbb{R}$ *be continuously differentiable and suppose that* $\Omega_c = \{x \in \mathbb{R}^n : v(x) \leq c\}$ *is bounded and that* $\dot{v} \triangleq f(x)dv/dx \leq 0$ *for all* $x \in \Omega_c$*. Define* $S \subset \Omega_c$ *by* $S = \{x \in \Omega_c : \dot{v}(x) = 0\}$ *and let* M *be the largest invariant set in* S*. Then, whenever* $x_0 \in \Omega_c$*,* $x(t)$ *approaches* M *as* $t \to \infty$*.*

To illustrate the oscillatory behavior of the primal/dual algorithm, let us consider a Lyapunov function P defined as

$$P(x,p) \triangleq \sum_{l\in[M]} (2K_l)^{-1}(x_l - x_l^*)^2 + \sum_{j\in[L]} (2\Gamma_j)^{-1}(p_j - p_j^*)^2.$$

Taking the time derivative of the solution path given by the primal/dual algorithm, one gets

$$\dot{P}(x,p) = \sum_{l\in[M]} (x_l - x_l^*) \left(U'_{i:l\in m_i}\left(\sum_{n\in m_i} x_n\right) - \sum_{j\in[L]} R_{jl} p_l^* \right).$$

Define the set $\Omega \triangleq \{(x,p) : \sum_l H_{il} x_l = \sum_l H_{il} x_l^*\}$. By the concavity of U_i, $\dot{P}(x,p) \leq 0$ for all $\{x_l\}$ and $\{p_j\}$ and the equality holds only when $(x,p) \in \Omega$. Therefore $\Omega \cap \{P(x,p) = C\}$ is an invariant set for the primal/dual algorithm for some C. By LaSalle's Invariance Principle the algorithm will eventually result in a periodic solution.

In the delay-free system, there have been mainly two decentralized approaches [55, 56] designed to address this issue of dual nondifferentiability. The first approach [55] is based on subgradient methods for maximizing nondifferentiable functions discussed in detail by [57]. Essentially the algorithm is a primal update algorithm with diminishing gains and binary feedback signals indicating the congestion levels of intermediate links. The use of diminishing gains ensures automatically the convergence of the algorithm at the price of slow convergence [1]. Another weak point of this method is that it is unsuitable for dynamically changing network conditions. Since once the network is changed, for example a new user initiates a file transfer, all other users have to reset their gains so that a new network equilibrium can be found. The second approach [56] avoids this shortcoming by utilizing the method of proximal optimization [58]. The idea is to transform the original nonstrictly con-

[1] It is generally possible to have exponentially diminishing gains to achieve exponential convergence [57]. But the condition for this is hard to be satisfied in a decentralized setting.

cave maximization problem (5.1) successively to a strictly concave maximization problem which can be solved by decentralized methods. The resulting decentralized algorithm effectively adds auxiliary state variables to the algorithm associated with the original optimization problem. Take the primal/dual algorithm for example, the new algorithm takes the form of the following with auxiliary variables $\{u_l\}$ for every path,

$$\begin{cases} \dot{x}_l = K_l(U_{i:l\in m_i}(z_{i:l\in m_i}) - (x_l - v_l) - q_l), & \forall l \in [M], \\ \dot{v}_l = \alpha_l(x_l - v_l), & \forall l \in [M], \\ \dot{p}_j = \Gamma_j(y_j - c_j), & \forall j \in [L]. \end{cases}$$

Let us see how the new algorithm stabilizes the system. Define the Lyapunov function $\hat{P}$ by

$$\hat{P}(x, v, p) \triangleq \sum_{l\in[M]} (2K_l)^{-1}(x_l - x_l^*)^2 + \sum_{l\in[M]} (2\alpha_l)^{-1}(v_l - x_l^*)^2 + \sum_{j\in[L]} (2\Gamma_j)^{-1}(p_j - p_j^*)^2.$$

Then its time derivative over the solution trajectory is

$$\dot{\hat{P}}(x, v, p) = \sum_{l\in[M]} (x_l - x_l^*)) \left(U'_{i:l\in m_i} \left(\sum_{n\in m_i} x_n \right) - \sum_{j\in[L]} R_{jl} p_l^* \right) - \sum_{l\in[M]} (x_l - v_l)^2.$$

One can see that $\dot{\hat{P}}(x, v, p) \leq 0$ for all $\{x_l\}$, $\{v_l\}$, and $\{p_j\}$ and the equality holds only when $(x, v, p) \in \Omega \times \{v_l = x_l, \forall l\} \triangleq \Omega'$. But this time the largest invariant set in Ω' is nothing but (x^*, x^*, p^*). Therefore again by LaSalle's Invariance Principle the algorithm converges to the equilibrium solution. The success of the introduction of auxiliary state variables for the design of a stabilizing controller indicates that it is necessary to use controllers with more states than a simple primal/dual controller to achieve our goal. This is our starting point to design scalable controllers for combined routing and flow control problem in a network with heterogeneous delays.

It is also beneficial to review the decentralized algorithms for the multipath optimization problem presented in the following form, which is distinct from (5.1),

$$\begin{aligned} &\max\nolimits_{x_i \geq 0} \textstyle\sum_{i\in[N]} U_i(z_i) - \sum_{j\in[L]} \int f_j(y_j)dy_j \\ &\text{s.t. } Hx = z, \\ &Rx = y. \end{aligned} \tag{5.3}$$

Here f_j is a convex increasing function which represents the congestion cost at link j. The dual of this optimization problem is

$$\min_{p_j} \max_{x_i \geq 0} \left(\sum_{i\in[N]} U_i \left(\sum_{l\in[M]} H_{il} x_l \right) - \sum_{j\in[L]} R_{jl} x_l p_j \right) + \left(\sum_{j\in[L]} p_j y_j - \int f_j(y_j) dy_j \right).$$

Therefore analogous to the single path flow control problem (3.4), (3.5, (3.6), the primal algorithm and the primal/dual algorithm to solve (5.3) can be written in the following form:

$$\text{Primal Law: } \begin{cases} \dot{x}_l = K_l(U_{i:l\in m_i}(z_{i:l\in m_i}) - q_l), & \forall l \in [M], \\ p_j = f_j(y_j), & \forall j \in [L]. \end{cases}$$

$$\text{Primal/Dual Law: } \begin{cases} \dot{x}_l = K_l(U_{i:l\in m_i}(z_{i:l\in m_i}) - q_l), & \forall l \in [M], \\ \dot{p}_j = \Gamma_j(f_j(y_j) - p_j), & \forall j \in [L]. \end{cases}$$

Global stability results for these algorithms have only been obtained for the delay-free case [51]. In [49, 50] authors presented decentralized algorithms which achieve local stability with delays in the network. A particularly interesting approach used in [50] to obtain per flow decentralization is to use delayed utility function instead. This idea of artificially adding delay in the system equation resembles our method. However all the decentralizing primal/dual algorithms require the end user to know the exact forms of cost function p_j of relevant links, which is not desirable in real implementations. More importantly, since the optimization problem 5.1 is equivalent to the problem (5.3) when the link cost function p_j takes the limit of $f_j(y) = \max\{0, \frac{y-c}{\delta}\}$ as $\delta \to 0$, then none of the stability conditions of any algorithms presented in [49, 50] remain valid in that limit. It can be expected that primal/dual algorithms fail to converge for our optimization problem (5.1).

5.4 Main Result

The main techniques used to derive our scalable controller for the combined routing and flow control problem follow closely from the design of scalable controllers for single path flow problem presented in Chapter 4. We use the notation of $\xi_i \triangleq -U_i''(z_i^*)$, the negative second derivative of utility function of user i at his equilibrium total rate z_i^*. Recall that the linear version of our scalable decentralized control law for single path flow control takes the form (4.17), (4.18) for the source controller

$$F_i(s) = \frac{s + \frac{1}{4(1+\tau_i)}}{4(\tau_i^2 + 1)|r_i|s + \frac{\xi_i}{4(1+\tau_i)}},$$

and the link controller

$$G_j(s) = \frac{1}{|f_j|s}.$$

The flow rate vector x and the link price vector p satisfy

$$x = \text{diag}(\{F_i\})q,$$

$$p = \text{diag}(\{G_j\})y.$$

The corresponding nonlinear dynamics can be written as

$$\begin{cases} \dot{v}_i = \frac{1}{16(1+\tau_i)(1+\tau_i^2)|r_i|}(U_i'(x_i) - q_i), \\ x_i = v_i - \frac{q_i}{4(1+\tau_i^2)|r_i|}, \\ \dot{p}_j = \frac{y_j - c_j}{|f_j|}. \end{cases}$$

In our design of scalable controllers for the combined routing and flow control problem, we make the following important assumption,

Assumption 5.1 *For every user i, the round-trip time of every available path τ_l with $l \in m_i$ is equal to each other. That is, $\tau_{l:l\in m_i} \equiv \tau_i$ for some τ_i.*

We want to remark that first this is not a constraining requirement for a network with heterogeneous delays from an algorithm design point of view. When the round-trip times of available paths of a source are different from each other, the source can hold the feedback signals q_l that have arrived early until the feedback signals from all the paths are available for him to process. As we can see later, this "artificial" introduction of delays in the control loop greatly simplifies the analysis of stability. Secondly, in reality the user may not need to add much delays to make all the round-trip times of his paths to be equal, since it has been shown that for the purpose of reliable transmission, it is necessary that all paths have the same delay in a multipath congestion control protocol. So this assumption can be regarded as a consequence of reliable transmission.

Therefore we propose our scalable controller for the combined routing and flow control problem, which is similar to the case of single path flow control, as follows,

$$\begin{cases} \dot{v}_{l:l\in m_i} = \frac{1}{16(1+\tau_i)(1+\tau_i^2)|r_i|}(U_i'(z_i) - q_l), \\ x_{l:l\in m_i} = v_l - \frac{q_l}{4(1+\tau_i^2)|r_i|}, \\ \dot{p}_j = \frac{y_j - c_j}{|f_j|}. \end{cases} \tag{5.4}$$

Here $|r_i|$ is defined as

$$|r_i| \triangleq \max_{l:l\in m_i} |r_l|.$$

It can be observed in (5.4) that the path rate controllers of the same user are the exactly the same.

We use the following notation for the repeated array in the subsequent analysis:

$$\{a_i\}_{n_i} \triangleq \{\underbrace{a_1, \cdots, a_1}_{n_1}, \underbrace{a_2, \cdots, a_2}_{n_2}, \cdots\}.$$

The linearization of scalable controller (5.4) is

$$sx_{l:l\in m_i} = -\frac{1}{16(1+\tau_i)(1+\tau_i^2)|r_i|}(\xi_i z_i + q_l) - \frac{sq_l}{4(1+\tau_i^2)|r_i|}.$$

By straightforward computations with $z = Hx$ we get the open loop gain $L(s)$ of the system as follows,

$$L(s) = \left(I + \operatorname{diag}\left(\left\{\frac{\xi_i}{16(1+\tau_i)(1+\tau_i^2)|r_i|s}\right\}_{M_i}\right) H^T H\right)^{-1} \times \operatorname{diag}\left(\left\{\frac{s + \frac{1}{4(1+\tau_i)}}{4(1+\tau_i^2)|r_i|s^2} e^{-\tau_i s}\right\}_{M_i}\right) \hat{R}(s)^H \operatorname{diag}(\{f_j^{-1}\})\hat{R}(s).$$

For notational convenience let us denote

$$l_i \triangleq \frac{\xi_i}{16(1+\tau_i)(1+\tau_i^2)|r_i|}, \forall i \in [N].$$

Then the first product term of $L(s)$ is

$$\left(I + \operatorname{diag}\left(\{l_i/s\}_{M_i}\right) H^T H\right)^{-1} = I - \operatorname{diag}\left(\left\{\frac{l_i}{l_i M_i + s}\right\}_{M_i}\right) H^T H. \tag{5.5}$$

Below we show how we reach this equality. Since from the definition,

$$H^T H = \begin{bmatrix} \mathbf{1}_{M_1}\mathbf{1}_{M_1}^T & & 0 \\ & \ddots & \\ 0 & & \mathbf{1}_{M_N}\mathbf{1}_{M_N}^T \end{bmatrix},$$

therefore the following equalities holds,

$$\mathrm{diag}(\{a_i\}_{M_i}) H^T H = H^T H \, \mathrm{diag}(\{a_i\}_{M_i}), \forall \{a_i\}$$

$$H^T H H^T H = \mathrm{diag}(\{M_i\}_{M_i}) H^T H.$$

Hence we have

$$\begin{aligned}
&(I + \mathrm{diag}(\{l_i/s\}_{M_i}) H^T H) \left(I - \mathrm{diag}\left(\left\{ \frac{l_i}{l_i M_i + s} \right\}_{M_i} \right) H^T H \right) \\
=& I + \mathrm{diag}\left(\left\{ \frac{l_i}{s} - \frac{l_i}{l_i M_i + s} \right\}_{M_i} \right) H^T H - \mathrm{diag}(\{l_i/s\}_{M_i}) H^T H \, \mathrm{diag}\left(\left\{ \frac{l_i}{l_i M_i + s} \right\}_{M_i} \right) \times H^T H \\
=& I + \mathrm{diag}\left(\left\{ \frac{l_i^2 M_i}{s(l_i M_i + s} \right\}_{M_i} \right) H^T H - \mathrm{diag}\left(\left\{ \frac{l_i^2}{s(l_i M_i + s)} \right\}_{M_i} \right) \mathrm{diag}(\{M_i\}_{M_i}) H^T H \\
=& I,
\end{aligned}$$

which shows that (5.5) holds.

Next select $M_i \times M_i$ unitary matrix Q_i such that the first row of Q_i is

$$[M_i^{-1/2}, \cdots, M_i^{-1/2}].$$

Define $M \times M$ unitary matrix Q by

$$Q \triangleq \begin{bmatrix} Q_1 & & 0 \\ & \ddots & \\ 0 & & Q_N \end{bmatrix}.$$

Then

$$\mathrm{diag}\left(\left\{ \frac{l_i}{l_i M_i + s} \right\}_{M_i} \right) H^T H = Q^T \mathrm{diag}\left(\left\{ \left[\frac{l_i M_i}{l_i M_i + s}, \underbrace{0, \cdots, 0}_{M_i - 1} \right] \right\} \right) Q.$$

Therefore we have

$$I - \operatorname{diag}\left(\left\{\frac{l_i}{l_i M_i + s}\right\}_{M_i}\right) H^T H = Q^T \operatorname{diag}\left(\left\{\left[\frac{s}{l_i M_i + s}, \underbrace{1, \cdots, 1}_{M_i - 1}\right]\right\}\right) Q.$$

Combining (5.5) and the above equation, the open loop gain $L(s)$ can be expressed as

$$\begin{aligned} L(s) = Q^T \operatorname{diag}&\left(\left\{\left[\frac{s}{l_i M_i + s}, \underbrace{1, \cdots, 1}_{M_i - 1}\right]\right\}\right) Q \\ &\times \operatorname{diag}\left(\left\{\frac{s + \frac{1}{4(1+\tau_i)}}{4(1+\tau_i^2)|r_i| s^2} e^{-\tau_i s}\right\}_{M_i}\right) \hat{R}(s)^H \operatorname{diag}(\{f_j^{-1}\}) \hat{R}(s), \end{aligned}$$

and the spectra of $L(s)$ is equal to

$$\sigma(L(s)) = \sigma\left(\begin{array}{c} \operatorname{diag}\left(\left\{\left[\frac{s}{l_i M_i + s}, \underbrace{1, \cdots, 1}_{M_i - 1}\right]\right\}\right) Q \operatorname{diag}\left(\left\{\frac{s + \frac{1}{4(1+\tau_i)}}{4(1+\tau_i^2)|r_i| s^2} e^{-\tau_i s}\right\}_{M_i}\right) \\ \times \hat{R}(s)^H \operatorname{diag}(\{f_j^{-1}\}) \hat{R}(s) Q^T \end{array}\right).$$

Since the unitary matrix Q has the same block structure as

$$\operatorname{diag}\left(\left\{\frac{s + \frac{1}{4(1+\tau_i)}}{4(1+\tau_i^2)|r_i| s^2} e^{-\tau_i s}\right\}_{M_i}\right),$$

therefore we have

$$\begin{aligned} &Q \operatorname{diag}\left(\left\{\frac{s + \frac{1}{4(1+\tau_i)}}{4(1+\tau_i^2)|r_i| s^2} e^{-\tau_i s}\right\}_{M_i}\right) \\ &= \begin{bmatrix} Q_1 \frac{s + \frac{1}{4(1+\tau_1)}}{4(1+\tau_1^2)|r_1| s^2} e^{-\tau_1 s} I & & 0 \\ & \ddots & \\ 0 & & Q_N \frac{s + \frac{1}{4(1+\tau_N)}}{4(1+\tau_N^2)|r_N| s^2} e^{-\tau_N s} I \end{bmatrix} \\ &= \operatorname{diag}\left(\left\{\frac{s + \frac{1}{4(1+\tau_i)}}{4(1+\tau_i^2)|r_i| s^2} e^{-\tau_i s}\right\}_{M_i}\right) Q. \end{aligned}$$

So we show that the spectra of the open loop gain $L(s)$ consists of exactly the eigenvalues of the following form,

$$\operatorname{diag}\left(\left\{\left[\frac{s}{l_iM_i+s},\underbrace{1,\cdots,1}_{M_i}\right]\right\}\right)\operatorname{diag}\left(\left\{\frac{s+\frac{1}{4(1+\tau_i)}}{4(1+\tau_i^2)|r_i|s^2}e^{-\tau_i s}\right\}_{M_i}\right)$$
$$\times Q\hat{R}(s)^H\operatorname{diag}(\{f_j^{-1}\})\hat{R}(s).$$

Since the first two terms are diagonal and the last five terms constitute a Hermitian matrix, a direct application of Lemma 4.3 by Vinnicombe and Lemma 4.4 yields the following conclusion,

Lemma 5.1 *If Assumption 5.1 holds, the combined routing and flow control algorithm (5.4) has linear stability for a network with heterogeneous delays.*

Now we are going to establish the global stability of the proposed algorithm, following the same method applied to the single path flow control problem in Chapter 4. We take the nonlinear feedback controller of the form

$$u_{l:l\in m_i}(x_l)=\frac{1}{16(1+\tau_i)(1+\tau_i^2)|r_i|}(U_i'(z_i)-q_i^*),$$

as inputs to the linear system with user total rates as system output. Then the output transfer function can be expressed by

$$W(s)^{-1}=H\left(sI+\operatorname{diag}\left(\left\{\frac{s+\frac{1}{4(\tau_i+1)}}{4(1+\tau_i^2)|r_i|s}e^{-\tau_i s}\right\}_{M_i}\right)\hat{R}(s)^H\operatorname{diag}(\{f_j^{-1}\})\hat{R}(s)\right)H^T$$
$$\triangleq H\hat{W}(s)^{-1}H^T.$$

As in the single path case, we introduce a Zames-Falb multiplier

$$M(s)^{-1}=\operatorname{diag}\left(\left\{\frac{s+\frac{1}{4(\tau_i+1)}}{4(1+\tau_i^2)|r_i|((\tau_i+1)s+1)}\right\}\right)$$

and notice that

$$H^T M(s)^{-1} = \hat{M}(s)^{-1} H^T$$

where

$$\hat{M}(s)^{-1} = \text{diag}\left(\left\{\frac{s + \frac{1}{4(\tau_i+1)}}{4(1+\tau_i^2)|r_i|((\tau_i+1)s+1)}\right\}_{M_i}\right).$$

Therefore

$$\begin{aligned} & W(s)^{-1}M(s)^{-1} + (M(s)^{-1})^H (W(s)^{-1})^H \\ =& H(\hat{W}(s)^{-1}\hat{M}(s)^{-1} + (\hat{M}(s)^{-1})^H (\hat{W}(s)^{-1})^H) H^T. \end{aligned}$$

By Lemma 4.5 it holds that $\hat{W}(s)^{-1}\hat{M}(s)^{-1} + (\hat{M}(s)^{-1})^H(\hat{W}(s)^{-1})^H \geq 0$ for all $s = i\omega$. Therefore by Lemma 5.1, concavity of U_is and applications of monotone multipliers described by [46, 43, 44] we have the following result

Theorem 5.2 *Assume the initial dynamics follows Remark 4.2 and Assumption 5.1 holds, the combined routing and flow control algorithm given by (5.4) is globally asymptotically stable for any networks with heterogeneous delays.*

5.5 Conclusion

We base our design of the combined multipath routing and flow control problem on the design methodology from the last chapter and provide the first scalable algorithm for the problem which achieves the global stability for general networks. One issue with our algorithm is that the solution may eventually have negative transmission rates on individual routes for some user, even though his total transmission rate remains positive. It is certainly valid to impose the positivity constraint on the rate update, but then we no longer have the classical Lur'e system so we cannot apply Zames-Falb theorem directly to prove the system stability. Further more work needs to be done to convert the continuous-time algorithm to an implementable one in real systems.

Chapter 6

Combined Routing and Flow Control - The Case of Single Path Routing

6.1 Introduction

In the last chapter we presented a scalable combined routing and flow control algorithm which enables network users to improve their flow efficiency over fixed routing transmission by using multiple network routes at the same time. This certainly gives the maximally possible size of the flow from a source to a destination, since the resulting capacity of the network transmissions achieves the max-flow min-cut bound [48]. However, such simultaneous routing and flow control solution requires special network routers to forward multipath packets, as well as different type of end-to-end protocol to handle out-of-order packets. Given the current network infrastructure we are left with the option of combined single path routing and flow control. That is, although each user has more than one routes to send his traffic, he can only utilize the "best" one. This is in analogy with the unsplittable flow problem in the context of network flow maximization and in some sense closer to the original meaning of "routing", to decide along which path to send traffic, than the simultaneous routing and flow control problem discussed in the last chapter.

Specifically, we use the same network model as in the last chapter in which there are N users and L bottleneck links. Each user i has a set of paths m_i with cardinality M_i available to send his traffic. The total number of paths is $M = \sum_{l=1}^{N} M_l$. Recall we index all the paths by the order of users so that $m_i = \{\sum_{l=1}^{i-1} M_l + 1, \cdots, \sum_{l=1}^{i} M_l\}$. We do not require the available paths for a single user to be disjoint from each other. As usual we use the notation x_l, c_j, z_i to represent the flow rate of path l, the bandwidth of link j, and the flow rate of user i, respectively, as well as $M \times L$ 0-1 matrix R as the routing matrix of the network to indicate the path/link relation. Distinct from the simultaneous routing and flow control problem, now each user cannot simply take the aggregation of traffic flows of all his available paths. Instead the effective user flow rate takes the following form

$$z_i(t) = \max_{l \in m_i} x_l(t).$$

Therefore the network optimization problem for combined single path routing and congestion control can be written as,

$$\begin{aligned} &\max\nolimits_{x_l \geq 0} \textstyle\sum_{i \in [N]} U_i(z_i) \\ \text{s.t. } \; &z_i = \max\nolimits_{l \in m_i} x_l, \forall i \in [N], \\ &Rx \leq c. \end{aligned} \tag{6.1}$$

The only difference between the above formulation and the multipath routing and flow control problem (5.1) is the effective user flow rate z_i. Using maximum instead of sum over path rates introduces nonconcavity into this problem, and consequently strong duality, which is fundamental to all the algorithms derived in previous chapters, does not hold any more. Therefore in this chapter we are not mainly concerned of distributed algorithms to solve the corresponding optimization problem, instead we will focus on studying the properties of (6.1).

This problem is first studied Wang, et al [59]. As the unsplittable flow problem, the problem of combined single path routing and flow control is NP hard, which means in the worst case scenario it is unlikely to have an efficient algorithm to obtain the optimal solution. For expository purpose, we briefly describe a special type of network by which Wang, et al in [59] showed the NP hardness. The network is shown in Figure 6.1. There are $N+3$ nodes in the network, among which there is 1 server node, 2 intermediate router nodes, and N edge router nodes from which users can access the network. Each edge router i has two identical links with bandwidth c_i to each of the intermediate routers. And each intermediate router has one direct link to the server with bandwidth equal to the half of the sum of its incoming bandwidths from all the edge routers, that is, $\frac{1}{2}\sum_{i=1}^{N} c_i$. Assume each user wants to establish a link to the server and then each edge router has to decide which one of the two outgoing links it should select. It is straightforward to see that solving the network optimization (6.1) is equivalent to solving the number partitioning problem, that is to minimize $|\sum_{i\in S} c_i - \sum_{i\notin S} c_i|$ over all possible sets S. The latter is known as one of the NP-complete problems [60]. Therefore the combined single path routing and flow control problem is NP-hard.

Although the above result presents a somewhat pessimistic perspective for complete algorithmic solution of general (6.1), we can nonetheless proceed to in the following directions which are still relevant to the real world scenario. First, since the exact optimal solution may be difficult to obtain, user may be content with a good approximation close to the true optimum. Second, the above NP-

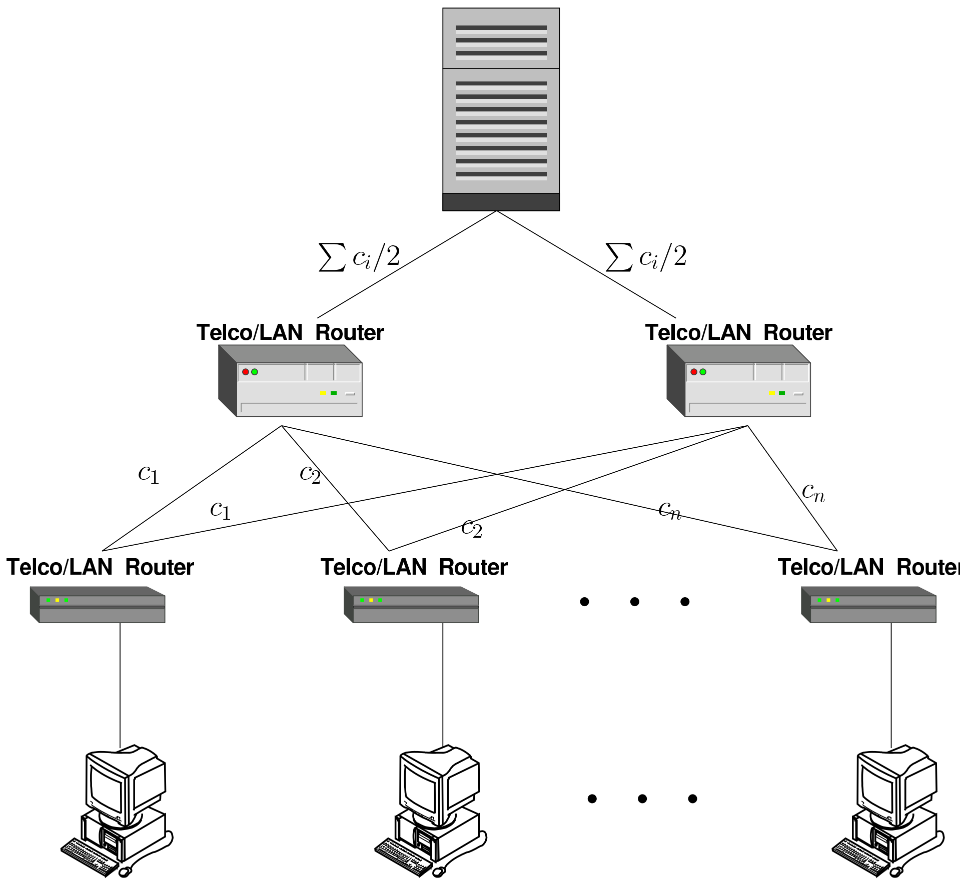

Figure 6.1: A simple network showing the joint single path routing and flow optimization is NP-hard

hardness proof relies on the assumption that the network size, i.e. the number of links, grows at a similar speed as the number of users. In reality there are much more network users than network links so that it does not deviate from the reality much to only consider the number of users grow while keeping the number of links fixed. In the case of the network in Figure 6.1 this is in fact similar to the scenario considered by Mertens [61] in his treatment of the number partitioning problem. He showed that when the ratio of the number resolution and the problem size is below a certain threshold, the number partitioning problem becomes easy to solve. Third, an important concern is whether the algorithm is local or not and how the algorithm uses local information. This is because improper use of local information, for example path selection decisions purely based on aggregate link prices, generally leads to route instability, as shown in [59].

Based on the above three observations, we will focus ourselves first on a simplistic local algorithm - Nash dynamics and see how it performs when the number of users grows large. For exactly the same network in Figure 6.1 we will show the route stability and bounded price of anarchy, that is, the gap between the result by simple Nash dynamics and the global optimal is small on average. Next we will show in Section 6.3 that in general networks all the Nash equilibrium solutions are close to the optimal solution when the number of users is sufficiently large. The final discussion is in Section 6.4.

6.2 Price of Anarchy - A Case Study

We consider the following type of noncooperative routing game with its normal form representation $([N], \{m_i\}, \{V_i\})$. The set of players $[N] = \{1, \cdots, N\}$ coincides with the set of users in the combined single path routing and flow control optimization problem (6.1). Each player/user i has a finite number of strategies - its available routes - m_i. A pure strategy profile is then represented by a N-tuple $\sigma = (\sigma_1, \cdots, \sigma_N)$ where $\sigma_i \in m_i$ is the strategy chosen by player i. The set of pure strategy profile is denoted by Σ. Player i's payoff function $V_i : \Sigma \to \mathbb{R}$ is defined by $V_i(\sigma) = U_i(z_i(\sigma))$ where $z_i(\sigma)$ is the optimal rate of player i in the following network optimal flow problem,

$$\begin{aligned} &\max\nolimits_{x_l \geq 0} \textstyle\sum_{i \in [N]} U_i(z_i) \\ &\text{s.t. } z_i = x_{\sigma_i}, \quad \sigma_i \in m_i, \forall i \in [N], \\ &Rx \leq c. \end{aligned} \tag{6.2}$$

Since the strategy set Σ is finite, there exists a pure strategy profile such that the resulting aggregate payoff in the above game achieves the maximum among all the possible strategies. This particular routing strategy is one of the Pareto optima of the game and along with the associated optimal flow rates, they are exactly the optimal solution of combined single path routing and flow control problem (6.1). To arrive at this Pareto optimum of the game requires global coordination among players in general. An alternative way is to look at a solution concept of the game in which only local interaction is needed. A simple and also most well-known such solution is the Nash equilibrium of the game, which is defined as a set of strategy

σ^{NE} such that

$$V_i(\sigma^{NE}) \geq V_i(\sigma_i, \sigma_{-i}^{NE}), \quad \forall \sigma_i \in m_i \text{ and } i \in [N].$$

Here σ_{-i}^{NE} denotes the $(N-1)$-tuple $(\sigma_1^{NE}, \cdots, \sigma_{i-1}^{NE}, \sigma_{i+1}^{NE}, \cdots, \sigma_N^{NE})$. We denote the set of Nash equilibria of the routing game by Σ^{NE}. The route update procedure to reach the Nash equilibrium can be described as follows. At each discrete time step t only a randomly selected player $p(t)$ switches its route to achieve a better resulting flow rate after the flow control mechanism is stabilized. This is known as Nash dynamics. Since this is a finite game, it is well known that there may not exist pure Nash equilibrium in general, and if that is the case, the prescribed route update procedure will never finish. We show below that in our routing game for the special network in Figure 6.1, this routing instability will never occur.

Proposition 6.1 *The routing game $(N, \{m_i\}, \{V_i\})$ for the network in Figure 6.1 has pure Nash equilibrium and consequently every Nash dynamics of the network converges after a finite number of steps.*

Proof: Let us define the set of strategy profile $\Sigma = \{0,1\}^N$ where 0 represents the left route and 1 represents the right route. Recall we use the notations $\sigma_i(t)$ and $z_i(t)$ for the route selection and actual bandwidth assigned to client i at the step t, respectively. Also define $P^0(t) = \{i \in [N] : \sigma_i(t-) = 0\}$ and $P^1(t) = \{i \in [N] : \sigma_i(t-) = 1\}$. Since every Nash dynamics can be decomposed into "rounds", during which the chosen players select the same route, let us denote T_k as the set of the time steps spent at round k. It follows as long as there is routing instability, or equivalently there are infinitely many rounds, the following holds,

$$c_{p(t_k)} > \sum_{i \in P^0(t_k)} c_i - \bar{c},$$

where $t_k = \max T_k$ and we assume without loss of generality hereafter the switch of route at round k is from left to right. This is because $p(t_k)$ is the last player in round k to change to his route from left to right.

We will prove that $c_{p(t_k)}$ is strictly decreasing in k therefore we conclude the number of round is finite. From the condition of changing route, we know that

$$z_{p(t_k)}(t_k) < z_{p(t_k)}(t_k + 1).$$

But in the round $k+1$, we know that

$$\bar{c} - \sum_{i \in P^0(t_{k+1}-1)} c_i \leq \bar{c} - \sum_{i \in P^0(t_k+1)} c_i,$$

since at $t_{k+1} - 1$ there are more players choosing the left route than at $t_k + 1$, and

$$\sum_{i \in P^1(t_{k+1}-1)} c_i - \bar{c} \leq \sum_{i \in P^1(t_k+1)} c_i - \bar{c},$$

since for the same reason there are less players choosing the right route than at $t_k + 1$.

Therefore $\forall i \in \sigma^{1,t_{k+1}-1}$, $z_i(t_{k+1} - 1) \geq z_i(t_k + 1)$. But on the other hand, the left route offers less "free" bandwidth than the beginning of this round $k+1$, thus any player $i \in P^1(t_{k+1} - 1) \cap \{j : c_j \geq c_{p(t_k)}\}$ will not have bandwidth gain when switching over to the left route. Thus we conclude $c_{p(t_{k+1})} < c_{p(t_k)}$ and the stability follows. ■

Therefore in contrast to the NP-hardness of the "social optimum solution" for the network in Figure 6.1, we have shown above that there always exist Nash equilibrium which can be reached in finite time by a simple local algorithm. It is natural to ask how far the Nash solution is from the network optimum. So next we should be concerned with ourselves the problem of the price of anarchy of this routing game, which measures this gap of the aggregate utility between the worst case Nash equilibrium and the social optimum. It is important to note that our routing game is significantly different from those studied in [62, 63, 64], since each player's strategy has non-negligible effects over other players' payoff and each player's payoff is not an explicit function of aggregate strategies. We adopt the definition of the price of anarchy as the *difference*, rather than the ratio as in most literatures, of the aggregated utility function of the worst case Nash equilibrium to the global optimal value. We will first study the case when each user has logarithm utility function, which corresponds to the proportional fairness allocation of network resources. Then we will consider more general utility functions which correspond to the α proportional fairness.

For N players in the network in Figure 6.1, without loss of generality we assume that $0 < c_1 \leq c_2 \leq \cdots \leq c_N$. Every strategy profile σ corresponds to a routing matrix $R(\sigma)$ and the rate allocations of the players are the solution to the following optimization problem,

$$\max_{z_i \geq 0} \sum_{i=1}^{N} \log z_i, \quad \text{s.t.} \quad R(\sigma)z \leq c,$$

where c is an appropriate column vector of link bandwidths.

It is easy to show that the solution satisfies the following property.

1. If $\sum_{i \in P^0} c_i \leq \bar{c}$, $z_i = c_i$, $\forall i \in \sigma^0$. Same applies to P^1.

2. If $\sum_{i \in P^0} c_i > \bar{c}$,

$$z_i = \min \left\{ c_i, \max_{k<i} \frac{\bar{c} - \sum_{j \in P^0, j<k} c_j}{|\{j \in P^0, j \geq k\}|} \right\}, \tag{6.3}$$

$\forall i \in P^0$. Same applies to P^1.

Suppose we fix the optimal aggregated utility function as

$$\sum_{i=1}^{N} \log c_i = 0,$$

or

$$\prod_{i=1}^{N} c_i = 1.$$

Here the "optimum" includes the situation when multipath routing is allowed. Therefore, our problem becomes,

$$\min_{\sigma \in \Sigma^{NE}} J_0 = \max_{z} \sum_{i=1}^{N} \log z_i, \quad \text{s.t.} \quad R(\sigma) z \leq c, \quad \prod_{i=1}^{N} c_i = 1. \tag{6.4}$$

Recall Σ^{NE} is the set of pure Nash equilibrium profiles.

Proposition 6.2 *The optimization problem (6.4) achieves its lower bound -1 when $N \to \infty$.*

Proof: For any Nash equilibrium profile σ define $\{p_1^0, \cdots, p_{N_1}^0\} \triangleq P^0$ and $\{P_1^1, \cdots, P_{N_2}^1\} \triangleq P^1$ where $N_1 + N_2 = N$. The players in each set are ordered such that $c_1^0 \leq \cdots \leq c_{N_1}^0$ and $c_1^1 \leq \cdots \leq c_{N_1}^1$. Without loss of generality assume

$$\sum_{i=1}^{N_2} c_i^1 > \bar{c}.$$

There exists an integer k_1, $0 \le k_1 \le N_1$, for the player set P_0 such that when a player in P_1 changes his route to the left for larger bandwidth, there will be k_1 players in P_0 that change their bandwidth due to the rule (6.3). It is easy to see that these k_1 players are the ones with the largest bandwidths in P_0. Define $s_1^0 \triangleq \sum_{i=1}^{N_1-k1} c_i^0$ and $s_2^0 \triangleq \sum_{i=N_1-k_1+1}^{N_1} c_i^0$. From the rule (6.3), the following inequality is satisfied,

$$s_1^0 - c_{N_1-k_1}^0 + (k_1+2)c_{N_1-k_1}^0 \le \bar{c} \le s_1^0 + (k_1+1)c_{N_1-k_1+1}^0.$$

But by definition $s_1^0 \le (N_1-k_1)c_{N_1-k_1}^0$ and $s_2^0 \ge k_1 c_{N_1-k_1+1}^0$, we have,

$$\frac{N_1+1}{N_1-k_1} s_1^0 \le \bar{c} \le s_1^0 + \frac{k_1+1}{k_1} s_2^0. \tag{6.5}$$

Also there exists an integer k_2, $0 \le k_2 < N_2$, such that there are k_2 players in P_1 whose allocated bandwidths are equal to their maximally possible bandwidths. It is straightforward to see that such k_2 players are the ones with the smallest bandwidths in P_1. Define $s_1^1 \triangleq \sum_{i=1}^{k_2} c_i^1$ and $s_2^1 \triangleq \sum_{i=k_2+1}^{N_2} c_i^1$. Again from the rule (6.3), the following inequality holds,

$$s_1^1 - c_{k_2}^1 + (N_2-k_2+1)c_{k_2}^1 \le \bar{c}.$$

By definition $s_1^1 \le k_2 c_{k_2}^1$, therefore,

$$\frac{N_2}{k_2} s_1^1 \le \bar{c}. \tag{6.6}$$

Since the current strategy profile σ is a Nash equilibrium, we have the following inequality from the allocation rule (6.3),

$$\frac{\bar{c}-s_1^1}{N_2-k_2} \ge \frac{\bar{c}-s_1^0}{k_1+1} \tag{6.7}$$

From (6.5) there exists $\delta_1 \geq 0$ such that

$$s_1^0 = \frac{N_1 - k_1}{N_1 + 1}(\bar{c} - \delta_0). \tag{6.8}$$

Again from (6.5) we have

$$s_2^0 = \frac{k_1}{N_1 + 1}\bar{c} + \frac{N_1 - k_1}{N_1 + 1}\frac{k_1}{k_1 + 1}\delta_0.$$

Add the above two equations together,

$$s_1^0 + s_2^0 = \frac{N_1}{N_1 + 1}\bar{c} - \frac{N_1 - k_1}{N_1 + 1}\frac{1}{k_1 + 1}\delta_0.$$

But since $s_1^0 + s_2^0 + s_1^1 + s_2^1 = 2\bar{c}$, we have

$$s_1^1 + s_2^1 = \frac{N_1 + 2}{N_1 + 1}\bar{c} + \frac{N_1 - k_1}{N_1 + 1}\frac{1}{k_1 + 1}\delta_0.$$

From (6.6) there exists $\delta_1 \geq 1$ such that

$$s_1^1 = \frac{k_2}{N_2}\bar{c} - \delta_1. \tag{6.9}$$

Hence

$$s_2^1 = \frac{N_1 + 2}{N_1 + 1}\bar{c} - \frac{k_2}{N_2}\bar{c} + \frac{N_1 - k_1}{N_1 + 1}\frac{1}{k_1 + 1}\delta_0 + \delta_1.$$

We can have $\delta_2 \geq 0$ such that

$$s_1^0 - s_1^1 = \frac{N_2 - k_2 - k_1 - 1}{N_2 - k_2}(\bar{c} - c_1^1) + \delta_2,$$

or equivalently,

$$s_1^0 = \frac{N_2 - k_1 - 1}{N_2}\bar{c} - \frac{k_1 + 1}{N_2 - k_2}\delta_1 + \delta_2$$

from the Nash equilibrium condition (6.7). Together with (6.8), it follows,

$$\left(\frac{1}{N_2} - \frac{1}{N_1 + 1}\right)\bar{c} + \frac{1}{N_2 - k_2}\delta_1 - \delta_2 = \frac{N_1 - k_1}{N_1 + 1}\delta_0.$$

Therefore

$$s_2^1 = \frac{N_2 - k_2 + 1}{N_2}\bar{c} + \frac{1}{N_2 - k_2}\delta_1 - \delta_2 \tag{6.10}$$

We return back to the optimization problem (6.4). It is clear that the final rates z_is from the current strategy profile σ satisfy

$$z_i = \begin{cases} c_i, & p_i \notin \{p^1_{k_2+1}, \cdots, p^1_{N_2}\}, \\ (\bar{c} - s_1^1)/(N_2 - k_2), & \text{otherwise.} \end{cases}$$

Therefore we have

$$\exp(J_0) = \prod_{p_i \notin \{p^1_{k_2+1}, \cdots, p^1_{N_2}\}} c_i \left(\frac{\bar{c} - s_1^1}{N_2 - k_2}\right)^{N_2 - k_2} = \left(\prod_{i=k_2+1}^{N_2} c_i^1\right)^{-1} \left(\frac{\bar{c} - s_1^1}{N_2 - k_2}\right)^{N_2 - k_2}.$$

The second inequality is due to the bandwidth constraint. But the above is equal to

$$\begin{aligned} &\left(\prod_{i=k_2+1}^{N_2} c_i^1\right)^{-1} \left(\frac{\frac{N_2-k_2}{N_2}\bar{c} + \delta_1}{N_2 - k_2}\right)^{N_2-k_2} \\ =&\frac{\left(\frac{s_2^1}{N_2-k_2}\right)^{N_2-k_2}}{\prod_{i=k_2+1}^{N_2} c_i^1} \left(\frac{\frac{N_2-k_2}{N_2}\bar{c} + \delta_1}{s_2^1}\right)^{N_2-k_2} \\ \geq &\left(\frac{\frac{N_2-k_2}{N_2}\bar{c} + \delta_1}{s_2^1}\right)^{N_2-k_2} \end{aligned}$$

The inequality is due to the fact that arithmetic average is greater than geometric average with equality when all the summands are equal to each other. Substitute s_2^1 with that in (6.10), it follows

$$\begin{aligned} &\left(\frac{\frac{N_2-k_2}{N_2}\bar{c} + \delta_1}{s_2^1}\right)^{N_2-k_2} \\ =&\left(\frac{\frac{N_2-k_2}{N_2}\bar{c} + \delta_1}{\frac{N_2-k_2+1}{N_2}\bar{c} + \frac{1}{N_2-k_2}\delta_1 - \delta_2}\right)^{N_2-k_2} \\ \geq &\left(\frac{N_2 - k_2}{N_2 - k_2 + 1}\right)^{N_2-k_2}. \end{aligned}$$

The inequality becomes equality when $\delta_1 = \delta_2 = 0$. Since it always holds

$$\left(\frac{N}{N+1}\right)^N \searrow e^{-1}, N \to \infty,$$

we conclude that $J_0 \geq -1$. A concrete example to achieve this bound is that there are $2N-1$ players and among those N players on the right route have bandwidths $\lambda(N+1)/N^2$ and $N-1$ players on the left route have bandwidths λ/N, where λ is the appropriate scaling constant so that the bandwidth constraint holds. The bound is achieved when $N \to \infty$. ■

We can generalize the result on logarithm utility function to power utility function, whose optimal solution corresponds to α proportional fairness. To this end, the network optimization problem is

$$\min_{z_i \geq 0} \sum_{i=1}^{N} -z_i^{-\gamma} \quad \text{s.t.} \quad R(\sigma)z \leq b,$$

where $R(\sigma)$ is the routing matrix corresponding to the strategy σ. The same as in the logarithm utility case, we try to obtain the value of worst case Nash equilibrium in the condition of fixed optimum value of multipath routing. That is the value of c_is must satisfy

$$\sum_{i=1}^{N} -c_i^{-\gamma} = -1.$$

Therefore the selfish routing problem of the network in Figure 6.1 for the power utility functions can be expressed as the following minimization problem,

$$\min_{\sigma \in \Sigma^{NE}} J_\gamma = \max_z \sum_{i=1}^{N} -z_i^{-\gamma}, \quad \text{s.t.} \quad R(\sigma)z \leq c, \quad \sum_{i=1}^{N} -c_i^{-\gamma} = -1. \tag{6.11}$$

Following the steps in the case of logarithm utility function, we obtain for a specific Nash equilibrium strategy σ, the aggregate utility in (6.11)

$$\begin{aligned}
J_\gamma = & \sum_{p_i \notin \{p^1_{k_2+1}, \cdots, p^1_{N_2}\}} -c_i^{-\gamma} - (N_2 - k_2)\left(\frac{\bar{c} - s^1_1}{N_2 - k_2}\right)^{-\gamma} \\
= & -1 + \sum_{i=1}^{N_2 - k_2} (c^1_i)^{-\gamma} - (N_2 - k_2)\left(\frac{\frac{N_2 - k_2}{N_2}\bar{c} + \delta_1}{N_2 - k_2}\right)^{-\gamma} \\
\geq & -1 + (N_2 - k_2)^{\gamma+1}(s^1_2)^{-\gamma} - (N_2 - k_2)^{\gamma+1}\left(\frac{N_2 - k_2}{N_2}\bar{c} + \delta_1\right)^{-\gamma} \\
= & -1 - (N_2 - k_2)^{\gamma+1}\left(\left(\frac{N_2 - k_2}{N_2}\bar{c} + \delta_1\right)^{-\gamma} - \left(\frac{N_2 - k_2 + 1}{N_2}\bar{c} + \frac{1}{N_2 - k_2}\delta_1 - \delta_2\right)^{-\gamma}\right).
\end{aligned}$$

Here we use the relations (6.9) and (6.10). In addition, the inequality follows from the following result

$$\frac{\sum_{i=1}^N x_i}{N} \geq \left(\frac{\sum_{i=1}^N x_i^{-\gamma}}{N}\right)^{-1/\gamma}$$

for all $x_i > 0$ and the equality holds if and only if $x_i \equiv x$, $\forall i$.

Again based on the above inequality result, as well as the definition of s^0_1, s^0_2, s^1_1, and s^1_2, we have the following inequalities,

$$\begin{aligned}
k_2^{\gamma+1}\left(\frac{k_2}{N_2}\bar{c} - \delta_1\right)^{-\gamma} &= k_2^{\gamma+1}(s^1_1)^{-\gamma} \leq \sum_{i=1}^{k_2}(c^1_i)^{-\gamma}, \\
(N_2 - k_2)^{\gamma+1}\left(\frac{N_2 - k_2 + 1}{N_2}\bar{c} + \frac{1}{N_2 - k_2}\delta_1 - \delta_2\right)^{-\gamma} &= (N_2 - k_2)^{\gamma+1}(s^1_2)^{-\gamma} \\
&\leq \sum_{i=k_2+1}^{N_2}(c^1_i)^{-\gamma}, \\
N_1^{\gamma+1}\left(\frac{N_2 - 1}{N_2}\bar{c} + \frac{N_2 - k_2 - 1}{N_2 - k_2}\delta_1 + \delta_2\right)^{-\gamma} &= N_1^{\gamma+1}(s^0_1 + s^0_2)^{-\gamma} \leq \sum_{i=1}^{N_1}(c^0_i)^{-\gamma}.
\end{aligned}$$

Add the above together,

$$\bar{c}^{-\gamma} \leq \left(\begin{array}{c} k_2^{\gamma+1}\left(\frac{k_2}{N_2} - \delta'_1\right)^{-\gamma} + (N_2 - k_2)^{\gamma+1}\left(\frac{N_2 - k_2 + 1}{N_2} + \frac{1}{N_2 - k_2}\delta'_1 - \delta'_2\right)^{-\gamma} \\ + N_1^{\gamma+1}\left(\frac{N_2 - 1}{N_2} + \frac{N_2 - k_2 - 1}{N_2 - k_2}\delta'_1 + \delta'_2\right)^{-\gamma} \end{array}\right)^{-1},$$

where we define $\delta'_i = \delta/\bar{c}$ for $i = 1, 2$. Hence

$$\begin{aligned} J_\gamma \geq & -1 - (N_2 - k_2)^{\gamma+1} \frac{\left(\frac{N_2-k_2}{N_2} + \delta'_1\right)^{-\gamma} - \left(\frac{N_2-k_2+1}{N_2} + \frac{1}{N_2-k_2}\delta'_1 - \delta'_2\right)^{-\gamma}}{\left(\begin{array}{c} k_2^{\gamma+1}\left(\frac{k_2}{N_2} - \delta'_1\right)^{-\gamma} + (N_2-k_2)^{\gamma+1}\left(\frac{N_2-k_2+1}{N_2} + \frac{1}{N_2-k_2}\delta'_1 - \delta'_2\right)^{-\gamma} \\ +N_1^{\gamma+1}\left(\frac{N_2-1}{N_2} + \frac{N_2-k_2-1}{N_2-k_2}\delta'_1 + \delta'_2\right)^{-\gamma} \end{array} \right)} \\ \geq & -1 - \frac{1 - \left(\frac{N_2-k_2}{N_2-k_2+1}\right)^\gamma}{\frac{k_2}{N_2-k_2} + \left(\frac{N_2-k_2}{N_2-k_2+1}\right)^\gamma + \frac{N_1}{N_2-k_2}\left(\frac{N_1}{N_2-1}\right)^\gamma}. \end{aligned}$$

The second inequality holds when $\delta'_1 = \delta'_2 = 0$. By detailed calculations it can be shown that the right hand side of the above inequality reaches its minimum when $k_2 = 0$ and $N_1 = N_2 - 1$ if N is odd and $N_1 = N_2$ if N is even. That is, the lower bound of the aggregate utility J_γ in the case of odd N is

$$J_\gamma \geq -1 - \frac{1 - \left(\frac{N+1}{N+3}\right)^\gamma}{1 - \frac{2}{N+1} + \left(\frac{N+1}{N+3}\right)^\gamma}$$

It is worth to notice that for large N, the lower bound can be approximated by $-1 - \gamma/N$. Therefore now the utility loss due to selfish routing becomes zero when the number of players grows to infinity, as oppose to the case of logarithm utility. Therefore we conclude that

Proposition 6.3 *The optimal value of the problem (6.11) converges to -1 when $N \to \infty$. Therefore the price of anarchy of the routing game with power utility is arbitrarily small for sufficiently large number of players.*

We extend the network in Figure 6.1 to the one with multiple links to the server instead of just two. The network consists of $N+M+1$ nodes, where there are N edge router nodes, M intermediate router nodes, and one server nodes. Each edge router node has links to every intermediate router nodes and all those links have the same bandwidth. Each intermediate router node has a direct link to the server. So there are altogether $(N+1)M$ links and each edge router has M different choices of route to the server. Again we denote the link bandwidth between client i and a router as c_i. The link bandwidth between a router and the server is $\bar{c} = \sum_{i=1}^{N} c_i/M$. Now the strategy set $\Sigma = \{1, \cdots, M\}^N$. This is similar to multi-processor scheduling problem, which is an extension to the number partitioning problem. We will show that the Nash dynamics also converges in this case.

Proposition 6.4 *The routing game $(N, \{m_i\}, \{V_i\})$ defined for the extended network has pure Nash equilibrium and consequently every Nash dynamics of the network converges after a finite number of steps.*

Proof: Clearly, the approach we use for the 2-link scenario can not be directly applied here. This is because the variable we considered in that case (the bandwidth of the last changing client in a round) is not monotonically decreasing due to the fact that we have more than two routing options for every player. So now we design a new bounded variable which is strictly increasing for each step in the Nash dynamics and therefore the convergence follows. From (6.7) we know that when a server link does not match the aggregated incoming bandwidths, or $\sum_{i \in P^j} c_i > \bar{c}$, the allocated bandwidth z_i for the player i, $i \in P^j \triangleq \{i \in [N] : \sigma_i = j\}$, is

$$z_i = \begin{cases} c_i, & i < k_j \\ \frac{\bar{c} - \sum_{l \in P^j, l < k_j} c_l}{|\{l \in P^j, l \geq k_j\}|}, & i \geq k_j, \end{cases}$$

where

$$k_j = \min\{i : c_i \geq \frac{\bar{b} - \sum_{l \in P^j, l < k_j} c_l}{|\{l \in P^j, l \geq k_j\}|}, i \in P^j\}.$$

In other words, those players whose bandwidth is less than a certain threshold are able to keep their bandwidth but the clients with higher bandwidth than that threshold have to divide the remaining bandwidth of the server link by equal shares. We try to look at the following variable for the player i at the route j which we define as y_i^j,

$$y_i^j \triangleq \frac{\bar{c} - \sum_{l \in P^j, l < i} c_l}{|\{l \in P^j, l \geq i\}|}.$$

This variable can be interpreted as the *worst projected allocated bandwidth* for the player i had the player i and all the players with higher bandwidth not joined the route j. We show that for the route j, y_i^j achieves maximum at $i = k_j$. Without loss of generality, we assume that both the player i and $i+1$ use the router j. We have,

$$\begin{aligned} y_i^j > y_{i+1}^j \quad &\Leftrightarrow \quad \frac{\bar{c} - \sum_{l \in P^j, l < i} c_l}{|\{l \in P^j, l \geq i\}|} > \frac{\bar{c} - \sum_{l \in P^j, l < i+1} c_l}{|\{l \in P^j, l \geq i+1\}|} \\ &\Leftrightarrow \quad \big(|\{l \in P^j, l \geq i\}| - |\{l \in P^j, l \geq i+1\}|\big) \sum_{l \in P^j, l < i} c_l + |\{l \in P^j, l \geq i\}| c_i \\ &\qquad > \big(|\{l \in P^j, l \geq i\}| - |\{l \in P^j, l \geq i+1\}|\big)\, \bar{c} \\ &\Leftrightarrow \quad \sum_{l \in P^j, l < i} c_l + |\{l \in P^j, l \geq i\}| c_i > \bar{c} \\ &\Leftrightarrow \quad c_i > \frac{\bar{c} - \sum_{l \in P^j, l < i} c_l}{|\{l \in P^j, l \geq i\}|}. \end{aligned}$$

So it follows directly that the y_i^j peaks at k_j and $y_{k_j}^j$ is the allocated bandwidth for players $i \geq k_j$. If the server link can contain the aggregated incoming bandwidths, y_i^j has the maximum at the player of the largest bandwidth. Define $y_j = \max_{i \in P^j} y_i^j$. We can write the allocated bandwidth for the player i using the route j (without the condition whether $\sum_{i \in P^j} c_i > \bar{c}$) as

$$z_i = \min\{c_i, y_j\}. \tag{6.12}$$

Now we define a vector $y(n) = [y_{(1)}, \cdots, y_{(M)}]$ which is a list of the lexicographically ordered y_j's at the nth step of the Nash dynamics and $y_{(i)}$ is the ith smallest y_j's. Certainly y_n is bounded. We will show that it is strictly increasing at each Nash step. Consider that at the nth step player i switches his route from j to k. Since it is a Nash step, we have that $z_i(n-1) < z_i(n)$. But from (6.12) $z_i(n-1) = y_j(n-1)$ and $z_i(n) = \min\{c_i, y_k(n)\}$. Therefore $y_j(n-1) < \min\{c_i, y_k(n)\} \leq y_k(n)$. Also after the removal of player i in the route j, it is obvious that $y_j(n-1) < y_j(n)$. Therefore $\min\{y_j(n-1), y_k(n-1)\} < \min\{y_j(n), y_k(n)\}$ which concludes that $y(n)$ is strictly increasing. So we conclude that the Nash dynamics converge. ■

To conclude this section we will demonstrate that the utility loss of the routing game for this extended network is greater than the case of 2 server links, although we will not derive the precise lower bounds. We can show that it is at least $1 - M$ when the logarithm utility function is used. Assume that the number of players N is a multiple of M. The link bandwidth for players $1, \cdots, n(1-1/M)$ is

$$\left(1 + \frac{M^2}{N-M}\right)^{-\frac{1}{M}}$$

and the link bandwidth for players $N(1-1/M)+1, \cdots, n$ is

$$\left(1 + \frac{M^2}{N-M}\right)^{\frac{M-1}{M}}.$$

The routing strategy for those N players is as follows. We group players $(k-1)N/M + 1, \cdots, kN/M$ to the server link k for $k = 1, \cdots, M$. It is easy to show that this routing scheme results in a Nash equilibrium. In this case, the aggregate utility function becomes

$$-\frac{N}{M}\log\left(1+\frac{M(M-1)}{N}\right)$$

which converges to the minimum of $1-M$ when $N\to\infty$. In the case of γ proportional fairness utility functions, we can also construct a Nash equilibrium which offers worse utility lower bound than that of the network of 2 server links. Suppose we have $N=Mr$ players and each server link serves r players. There are $M-K$ over-provisioned server links and K under-provisioned server links. As in the logarithm utility situation, in order to have Nash equilibrium, the link bandwidth c_1 of any players who use the over-provisioned server links must satisfy $c_1 \geq (r-1)/r\bar{c}$. Therefore we let

$$c_1=\left((M-K)r+Kr\left(1+\frac{M}{K(1-r)}\right)^{-\gamma}\right)^{1/\gamma}$$

and let the link bandwidth of any players who use the under-provisioned server links be

$$c_2=\left(1+\frac{M}{K(r-1)}\right)c_1.$$

We obtain the utility loss as

$$-\frac{M-K+K\left(\frac{r}{r-1}\right)^{-\gamma}}{M-K+K\left(1+\frac{M}{K(r-1)}\right)^{-\gamma}}.$$

It attains minimum when $K=1$ and $r\sim\lambda\gamma$ as $\gamma\to\infty$ where γ is the real root of the equation

$$\lambda^{\frac{M}{M-1}}+\frac{M}{M-1}\lambda^{\frac{1}{M-1}}-1=0.$$

The maximum utility loss is then

$$\frac{1+(M-1)\lambda}{1+(M-1)\lambda^{\frac{M}{M-1}}}.$$

6.3 Nash Equilibrium and Optimality - Asymptotic Results

We have shown in the previous section that for a special type of network, the value of Nash equilibrium solutions of the routing game are asymptotically close to the optimal value of combined single routing and flow control problem. In this section our intention is to show that in some sense this argument is valid for general networks when the number of users become large compared to the number of bottleneck links. This formulation of the problem in the many users region instead of many links region is motivated by the fact that in the real world there are usually much more users than bottleneck links in the Internet.

It is well known that Nash equilibrium causes efficiency loss in exchange economy as oppose to the competitive equilibrium where each player acts like price-taker. It is plausible that in large economies this price-taking behavior is justifiable, since each player's ability to influence the price formation and consequently his gain to deviate from his true demand is diminished when the number of players becomes large. This limit behavior of Nash equilibrium, either in a continuum economy or in an asymptotic sequence of economies, has been studied extensively (see for example, [65, 66, 67] and references therein) and many indicate the convergence to the competitive equilibrium. However, there are also many situations (see for example [68]) which show that in some economies the competitive equilibrium is not an asymptotic limit of any Nash equilibria, or even in a continuum setting Nash equilibrium can be far away from the competitive equilibrium. Note that our network flow optimization problem is a special type of pure exchange economy and the competitive equilibrium corresponds to the optimal (price-taking) solution. It is demonstrated as well in [24] that for a particular network setting, the aggregate utility of a Nash equilibrium is 3/4 of the optimal solution when the number of users goes infinite. Nonetheless, we will show that Nash equilibrium of our routing game (6.2) converges to the optimal solution in the many players region.

First we will motivate our intuition by showing the relation between the combined single path routing and flow control problem (6.1) and the multipath routing and flow control problem (5.1). Recall the definition of the conjugate function $f^* : X^* \to \mathbb{R}$ of $f : X \to \mathbb{R}$ as $f^*(y) = \inf_{x \in X}\{\langle x, y\rangle - f(x)\}$, where X and X^* is a pair of dual vector spaces defined by a bilinear operator $\langle \cdot, \cdot \rangle$. Then the bipolar function f^{**} of f is the conjugate of the conjugate of f, that is, $f^{**}(x) = \inf_{y \in X^*}\{\langle x, y\rangle - f^*(y)\}$. The following theorem by Falk states that the dual of a nonconvex optimization with linear constraints is equivalent to the "convex envelope" of the primal optimization,

Theorem 6.1 ([69]) *For a compact set X and $f : X \to \mathbb{R}$ a lower semicontinuous function over X, consider the optimization problem*

$$(P) \quad \max f(x), \ \textit{subject to } Ax \leq b, x \in X,$$

with its dual problem

$$(P^*) \quad \min_{y \geq 0} \max_{x \in X} f(x) - y^T Ax + y^T b.$$

The dual problem P^* is also the dual problem associated with the following problem

$$(P') \quad \max f^{**}(x), \ \textit{subject to } Ax \leq b, x \in X.$$

Here f^{**} is the bipolar function of f. Further, if the Slater condition is satisfied, then the strong duality between P' and P^* holds in which the maximum value of P' is equal to the minimum value of P^*.

The above theorem can be applied immediately to the combined single path and flow control problem (6.1. The only thing left to be calculated is the bipolar function of $f_i(x_1, \cdots, x_{M_i}) \triangleq U_i(\max\{x_1, \cdots, x_{M_i}\})$,

$$f_i^*(y_1, \cdots, y_{M_i}) = \inf_{x_1, \cdots, x_{M_i}} \left\{ \sum_{j=1}^{M_i} x_j y_j - U_i(\max\{x_1, \cdots, x_{M_i}\}) \right\}$$
$$= U_i^*(\min\{y_1, \cdots, y_{M_i}\}),$$

and

$$f_i^{**}(x_1, \cdots, x_{M_i}) = \inf_{y_1, \cdots, y_{M_i}} \left\{ \sum_{j=1}^{M_i} x_j y_j - U_i^*(\min\{y_1, \cdots, y_{M_i}\}) \right\}$$
$$= U_i \left(\sum_{j=1}^{M_i} x_j \right).$$

The above derivation uses the assumption that $U_i(\cdot)$ is a concave increasing function. Notice this utility function is exactly the utility function used in the combined multipath routing and flow control problem (5.1), therefore we conclude,

Proposition 6.5 *The dual problem of the combined single path routing and flow control optimization (6.1) is equivalent to the combined multipath routing and flow control optimization (5.1). Therefore the duality gap is nonzero if and only if the optimal value of (5.1) is strictly larger than that of (6.1).*

So the duality gap of the nonconvex optimization problem (6.1) can be interpreted as the "social" inefficiency caused by restricting every player to use only one path to route his traffic. It is then interesting to see what the minimum relaxation of this restriction one should make in order to eliminate this gap. We will show that one only needs to make negligible modifications of this single path rule to achieve multipath optimality in the many-player region. The derivation relies on a theorem by Shapley and Folkman (see [70] Appendix I) whose statement is as follows,

Theorem 6.2 (Shapley-Folkman) *Given a finite family of sets* $X_i \subset \mathbb{R}^m$, $i \in I$, *for any* $x \in \operatorname{co} \sum_{i\in I} X_i$, *there exists a subset* $I(x) \subset I$, *whose cardinality* $|I(x)| \le m$, *such that* $x \in \operatorname{co} \sum_{i\in I(x)} X_i + \sum_{i\in I\setminus I(x)} X_i$.

Intuitively, the Shapley-Folkman Theorem says that the sum of a large number of nonconvex sets in a finite dimensional space is close to a convex set. Let us proceed to see the implication of this "smoothing" effect on our problem. First let us introduce an indicator function $\chi_l : \mathbb{R} \to \{0, -\infty\}$ for each link l such that

$$\chi_l(y) = \begin{cases} 0, & y \ge -c_l, \\ -\infty, & y < -c_l. \end{cases}$$

We will consider the following perturbed function

$$\Phi(x_1, \cdots, x_M; d_1, \cdots, d_L) \triangleq \sum_{i=1}^{N} U_i(\max_{j\in m_i} x_j) + \sum_{l=1}^{L} \chi_l \left(d_l - \sum_{j=1}^{M} R_{lj} x_j \right),$$

and the perturbed maximization problem

$$V(d_1, \cdots, d_L) = \max_{\{x_i\}\in\mathbb{R}_+^M} \Phi(x_1, \cdots, x_M; d_1, \cdots, d_L).$$

It is clear that the optimal value of the combined single path routing and flow control optimization (6.1) is the same as $V(0, \cdots, 0)$. For notational convenience define $N+L$ functions $g_i : \mathbb{R}^{M+L} \to \mathbb{R} \cup \{-\infty\}$ as follows,

$$g_i(\{x_i\}) \triangleq \begin{cases} U_i(\max_{j \in M_i} x_j), & 1 \le i \le N, \\ \chi_{i-N}(x_i), & N+1 \le i \le N+L. \end{cases}$$

Also define a $L \times (M+L)$ 0-1 matrix E such that

$$\sum_{j=1}^{M+L} E_{lj} x_j = \sum_{j=1}^{M} R_{lj} x_j + x_{N+l}.$$

Then we can rewrite the optimization problem for $V(d_1, \cdots, d_L)$ as

$$\begin{aligned} & V(d_1, \cdots, d_L) \\ = & \max_{\{x_i\} \in \mathbb{R}_+^{M+L}} \left\{ \sum_{i=1}^{N} g_i(\{x_{j \in m_i}\}) : \sum_{j=1}^{M} R_{lj} x_j + x_{M+l} = d_l \right\} \\ = & \max_{\{x_i\} \in \mathbb{R}_+^{M+L}} \left\{ \alpha : (d_1, \cdots, d_L, \alpha) = \sum_{i=1}^{N+L} \left(\sum_{j \in m_i} E_{\cdot j} x_j, \alpha_i \right), \left(\sum_{j \in m_i} E_{\cdot i} x_i, \alpha_i \right) \in W_i \right\}, \end{aligned}$$

where the set $W_i \in \mathbb{R}^{L+1}$ is defined by

$$W_i \triangleq \left\{ (y, \alpha) : y = \sum_{j \in m_i} E_{\cdot j} x_j, \alpha \le g_i(\{x_{j \in m_i}\}) \right\}.$$

Define hypograph of a function f as $\operatorname{hypo} f \triangleq \{(x,y) : y \le f(x)\}$. Since for any function f its bipolar f^{**} satisfies $\operatorname{hypo} f^{**} = \bar{\operatorname{co}} \operatorname{hypo} f$, where co is the convex hull, it follows

$$\operatorname{graph} V^{**} \subset \bar{\operatorname{co}} \operatorname{graph} V \subset \bar{\operatorname{co}} \sum_{i=1}^{N+L} W_i.$$

Since the dual optimal is finite in our network optimization problem, by the upper-semicontinuity of the sets W_i, each of the set $\operatorname{co} W_i$ is closed, and we have

$$(0, \cdots, 0, V^{**}(0, \cdots, 0)) \in \operatorname{co} \sum_{i=1}^{N+L} W_i = \operatorname{co} \sum_{i=1}^{N+L} \left\{ \left(\sum_{j \in m_i} E_{\cdot j} x_j, g_i(\{x_{j \in m_i}\}) \right) \right\}.$$

So $(0, \cdots, 0, V^{**}(0, \cdots, 0))$ is in the sum of $N + L$ convex sets. By the Shapley-Folkman Theorem 6.2, there exists an index set I with cardinality at most $L + 1$ such that

$$(0, \cdots, 0, V^{**}(0, \cdots, 0)) \in \sum_{i \notin I} W_i + \text{co} \sum_{i \in I} W_i.$$

Equivalently, there exist $\tilde{x}_i$s for $i \notin I$ and two $|I| \times (L+2)$ matrices, $\{\tilde{x}_{ij}\}$ and $\{\gamma_{ij}\}$, such that

$$\sum_{i \notin I} E_{li} \tilde{x}_i + \sum_{i \in I} E_{li} \sum_{j=1}^{L+2} \gamma_{ij} \tilde{x}_{ij} = 0, \quad \forall l \in [L],$$
$$\sum_{i \notin I} g_i(\tilde{x}_i) + \sum_{i \in I} \sum_{j=1}^{L+2} \gamma_{ij} g_i(\tilde{x}_{ij}) = V^{**}(0, \cdots, 0).$$

Here $\gamma_{ij} \geq 0$ and $\sum_j \gamma_{ij} = 1$ by the representation of the convex hull from the Caratheodory theorem. Denote $\tilde{x}_i = \sum_j \gamma_{ij} \tilde{x}_{ij}$ for $i \in I$, we have

$$\sum_{j=1}^{L+2} \gamma_{ij} g_i(\tilde{x}_{ij}) \leq g^{**}(\tilde{x}_i).$$

Therefore we obtain

$$\sum_{i \notin I} E_{li} \tilde{x}_i + \sum_{i \in I} E_{li} \tilde{x}_i = 0, \quad l \in [L],$$
$$\sum_{i \notin I} g_i(\tilde{x}_i) + \sum_{i \in I} g_i^{**}(\tilde{x}_i) \geq V^{**}(0, \cdots, 0).$$

But the above really is

$$\begin{aligned} V^{**}(0, \cdots, 0) \leq \tilde{V} \triangleq \max_{x_l \geq 0} & \sum_{i \in [N]} U_i(z_i) \\ \text{s.t. } z_i &= \max_{l \in m_i} x_l, \forall i \in I, \\ z_i &= \sum_{l \in m_i} x_l, \forall i \in [N] \setminus I, \\ Rx &\leq c. \end{aligned}$$

It is straightforward to see that the above non-strict inequality should be equality, that is, $V^{**}(0, \cdots, 0) = \tilde{V}$. Therefore we have the following conclusion

Proposition 6.6 *The combined single path routing and flow control problem (6.1) can achieve the same efficiency as the combined multipath routing and flow control problem (5.1) by allowing at most $L+1$ users to use multipath routing to transmit their traffic.*

Hence in the many-user region, the percentage of users that needs to be changed in order to transform the hard problem (6.1) into the easy problem (5.1) is vanishingly small given the fixed number of links. Also we can now see that the reason for the problem (6.1) to be difficult to solve for the network on Figure 6.1 is that there are at least as many bottleneck links as the number of users. Therefore, intuitively, in the many-user region, the problem of combined single path routing and flow control becomes close to its multipath counterpart and thus easier to solve. However for this section we will not discuss the exact solution of the problem (6.1) in the many-user region. Instead we will show below that in the many-user asymptotics, the approximate Nash equilibrium should be able to achieve the approximate optimal solution.

We have shown in the last section that for the network on Figure 6.1 or its extension there exists pure Nash equilibrium for our routing game $([N], \{m_i\}, \{V_i\})$. But whether a pure Nash equilibrium of the routing game exists for a general network is an open problem. To circumvent this issue we introduce a more general ϵ-Nash equilibrium: a strategy $\sigma \in \Sigma$ is a ϵ-Nash equilibrium if and only if $V_i(s, \sigma_{-i}) \leq V_i(\sigma) + \epsilon$ for all $s \in m_i$ and all $i \in [N]$. Recall the notation $\sigma_{-i} = (\sigma_1, \cdots, \sigma_{i-1}, \sigma_{i+1}, \cdots, \sigma_N)$. We denote the set of ϵ-Nash equilibria by $\Sigma^{NE(\epsilon)}$. Clearly $\Sigma^{NE} = \Sigma^{NE(0)}$ and $\Sigma^{NE(\epsilon)} \subset \Sigma^{NE(\epsilon')}$ if $\epsilon \leq \epsilon'$.

Now we construct a concrete example of many-user network. Here we use the term "network" for the network topology along with users' characteristics, that is, their utility functions and their available routes. Denote $\mathcal{U}$ as the set of concave strictly increasing functions defined on $\mathbb{R}_+$ and M as the set of all possible routes in a given network topology. For simplicity we consider the following "type sequence" of users for the network topology. There is a finite set $[T]$ of types and each $t \in [T]$ corresponds to a utility/strategy set pair $(U_t(\cdot), m_t) \in \mathcal{U} \times 2^M$. Denote the Nth network by $\mathcal{N}_N$ and its consists of $n_N \in [N, N+T)$ users, among whom there are $\lceil w_t N \rceil$ users of type t for each $t \in T$. Here $w_t > 0$ can be considered as the percentage of type t users in the entire population and we have $\sum_{t \in T} w_t = 1$. In addition, the bandwidth of each link l of $\mathcal{N}_N$ is equal to Nc_l. The type sequence method offers a simple model similar to the real world scenario and its use is popular as the first step towards the study of the limiting behavior of large number of users in economic theory (see for example in the case of core equivalence [16]).

We will use a parametric dependence theorem by Wets frequently in the rest of this section. First let us introduce the concept of "uniform level boundedness". A function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ with values $f(x, u)$ is level-bounded in x locally uniformly in u if for each $\bar{u} \in \mathbb{R}^m$ and $\alpha \in \mathbb{R}$ there is a neighborhood V of $\bar{u}$ along with a bounded set $B \subset \mathbb{R}^n$ such that $\{x : f(x, u) \geq \alpha\} \subset B$ for all $u \in V$. Now the theorem is stated as follows

Theorem 6.3 (Theorem 1.17 of [71]) *Consider*

$$p(u) = \sup_x f(x,u), \quad P(u) = \operatorname*{argmax}_x f(x,u),$$

in the case of a proper, upper semicontinuous function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{\pm\infty\}$ *such that* $f(x,u)$ *is level-bounded in* x *locally uniformly in* u*. Then*

(a) The function p *is proper and upper semicontinuous on* $\mathbb{R}$*, and for* p *to be continuous at a point* $\bar{u}$ *relative to a set* U *containing* $\bar{u}$*, a sufficient condition is the existence of some* $\bar{x} \in P(\bar{u})$ *such that* $f(\bar{x}, u)$ *is continuous in* u *at* $\bar{u}$ *relative to* U*.*

(b) If $x_n \in P(u_n)$*, and if* $u_n \to \bar{u} \in \operatorname{dom} p$ *in such a way that* $p(u_n) \to p(\bar{u})$*, then the sequence* $\{x_n\}$ *is bounded, and all its cluster points lies in* $P(\bar{u})$*.*

In the network $\mathcal{N}_N$, for type t users we introduce a M_t-dim vector $\{v^N_{t,\tau}\}$ in which $v^N_{t,\tau}$ represents the number of type t users choosing the τth route within m_t. Apparently $\sum_\tau v^N_{t,\tau} = \lceil w_t N \rceil$ and all these $v^N_{t,\tau}$ users will have the same rate allocation. Therefore each $\{v^N_{t,\tau}\}$ corresponds to an equivalent class of strategy profile σ in the sense of rate distribution. For a fixed $\{v^N_{t,\tau}\}$ the combined single path routing and flow control problem for the network $\mathcal{N}_N$ can be written as

$$\begin{aligned} \max_{x \geq 0} & \sum_{t \in [T]} \sum_{\tau=1}^{M_t} v^N_{t,\tau} U_t(x_{t,\tau}) \\ \text{s.t.} & \sum_{t \in [T]} \sum_{\tau=1}^{M_t} R^l_{t,\tau} v^N_{t,\tau} x_{t,\tau} \leq N c_l, \quad \forall l \in [L]. \end{aligned}$$

Here R is the routing matrix of the network which is invariant in N. Define $\bar{v}_{t,\tau} \triangleq v^N_{t,\tau}/N$ to be the scaled down version of $v_{t,\tau}$. Further denote the finite set $\bar{\mathcal{V}}_N$ to be all the possible $\{\bar{v}_{t,\tau}\}$ in which each component can be expressed by $\bar{v}_{t,\tau} = v_{t,\tau}/N$ for some $v^N_{t,\tau}$ for all t and τ. Then the "scaled down" version of the above optimization problem can be rewritten as

$$\begin{aligned} V(\{\bar{v}_{t,\tau}\}) &\triangleq \max_{x\geq 0} \textstyle\sum_{t\in[T]} \sum_{\tau=1}^{M_t} \bar{v}_{t,\tau} U_t(x_{t,\tau}) \\ \text{s.t. } & \textstyle\sum_{t\in[T]} \sum_{\tau=1}^{M_t} R^l_{t,\tau} \bar{v}_{t,\tau} x_{t,\tau} \leq c_l, \quad \forall l \in [L]. \end{aligned} \tag{6.13}$$

Note that for $\mathcal{N}_N$, $\sum_\tau \bar{v}_{t,\tau} = n_N/N \in [w_t, w_t + L/N)$. The dual problem of the above optimization is

$$V^*(\{\bar{v}_{t,\tau}\}) \triangleq \min_{p\geq 0} D(p, \{\bar{v}_{t,\tau}\}) - \sum_{l=1}^{L} c_l p_l,$$

where

$$D(p, \{\bar{v}_{t,\tau}\}) \triangleq \max_{x\geq 0} \sum_{t\in[T]} \sum_{\tau=1}^{M_t} \bar{v}_{t,\tau} \left(U_t(x_{t,\tau}) - \sum_{l=1}^{L} R^l_{t,\tau} x_{t,\tau} p^l \right).$$

Since $U_t \in \mathcal{U}$, $\bar{v}_{t,\tau} \left(U_t(x_{t,\tau}) - \sum_{l=1}^{L} R^l_{t,\tau} x_{t,\tau} \right)$ is everywhere continuous for $x_{t,\tau} \in \mathbb{R}_+$. Therefore by Theorem 6.3 $D(p, \{\bar{v}_{t,\tau}\})$ is everywhere continuous for $p \in \mathbb{R}^L_+$. Since for any $\{\bar{v}_{t,\tau}\}$, the optimal solution $p(\{\bar{v}_{t,\tau}\})$ is unique. Then again by Theorem 6.3, $p(\{\bar{v}_{t,\tau}\})$ is continuous for all $\{\bar{v}_{t,\tau}\}$.

We consider the multipath version of the problem,

$$\begin{aligned} & \max_{x\geq 0} \textstyle\sum_{t\in[T]} w_t U_t(z_t) \\ \text{s.t. } & \textstyle\sum_{t\in[T]} \sum_{\tau=1}^{M_t} R^l_{t,\tau} w_t x_{t,\tau} \leq c_l, \quad \forall l \in [L], \ c \\ & z_t = \textstyle\sum_{\tau=1}^{M_t} x_{t,\tau}. \end{aligned} \tag{6.14}$$

Denote the solution of the above multipath optimization to be z^*_t and $x^*_{t,\tau}$ (if there are multiple solutions, we just pick one of them) and the solution to its dual problem to be p^*_l (it is unique according to Proposition 5.1). Define $\bar{v}^*_{t,\tau} \triangleq w_t x^*_{t,\tau}/z^*_t$ and the aggregate price along the route $q^*_{t,\tau} \triangleq \sum_l R^l_{t,\tau} p^*_l$. From optimality we have for fixed t, $q^*_{t,\tau} = q^*_{t,\tau'}$ if $\bar{v}^*_{t,\tau} > 0$ and $\bar{v}^*_{t,\tau'} > 0$ and $q^*_{t,\tau} \le q^*_{t,\tau'}$ if $\bar{v}^*_{t,\tau} > 0$ and $\bar{v}^*_{t,\tau'} = 0$. Further, $x_{t,\tau} = z^*_t$ if $\bar{v}^*_{t,\tau} > 0$ and $x_{t,\tau} = 0$ if $\bar{v}^*_{t,\tau} = 0$ is the optimal solution of (6.13) with parameters $\bar{v}_{t,\tau} = \bar{v}^*_{t,\tau}$, and along with $\bar{v}^*_{t,\tau}$, the multipath solution is the optimal solution of $\max V(\{\bar{v}_{t,\tau}\})$ with the constraint $\sum_\tau \bar{v}_{t,\tau} = w_t$ for all $t \in [T]$.

Define $\{\bar{v}^{*N}_{t,\tau}\}$ to be the closest element in $\bar{\mathcal{V}}_N$ to $\{\bar{v}^*_{t,\tau}\}$ in l_∞. We argue that for any $\epsilon > 0$, there exists an integer $N(\epsilon)$, such that for any $N > N(\epsilon)$, any strategy profile corresponds to $\{\bar{v}^{*N}_{t,\tau}\}$ is a ϵ-Nash equilibrium. To see this, first notice that we have

$$U'_t(z^*_t) = q^*_{t,\tau}, \quad \forall \tau \in m_t, \text{ such that } x^*_{t,\tau} > 0.$$

Since $U_t \in \mathcal{U}$, for fixed ϵ, there exists $\eta(\epsilon) > 0$, such that for any $0 \le \eta < \eta(\epsilon)$, $|U_t(z^*_t) - U_t(U'^{-1}_t(q))| < \epsilon/2$ for all $|q - q^*_{t,\tau}| < \eta$. Next recall from the above discussion $p(\{\bar{v}_{t,\tau}\})$ is continuous, then there exists $\delta(\eta) > 0$, such that for any $0 \le \delta < \delta(\eta)$,

$$\|p(\{\bar{v}^*_{t,\tau}\}) - p(\{\bar{v}_{t,\tau}\})\|_\infty < \eta/L, \quad \forall \|\{\bar{v}^*_{t,\tau}\} - \{\bar{v}_{t,\tau}\}\| < \delta.$$

We know that from our construction there exists $N(\delta)$ such that for all $N > N(\delta)$

$$\|\{\bar{v}^{*N}_{t,\tau}\} - \{\bar{v}^*_{t,\tau}\}\|_\infty < \delta,$$

and

$$\|\{\bar{v}^{*N}_{t,\tau}\}' - \{\bar{v}^*_{t,\tau}\}\|_\infty < \delta,$$

where $\{\bar{v}^{*N}_{t,\tau}\}'$ is the route distribution when a single user changes his strategy from that in $\{\bar{v}^{*N}_{t,\tau}\}$ while the rest of users keep the same strategies. Suppose the user changes his strategy from τ to τ' in m_t. Then we can readily conclude that

$$\|p(\{\bar{v}^{*N}_{t,\tau}\}) - p(\{\bar{v}^{*}_{t,\tau}\})\|_\infty < \eta/L,$$

and

$$\|p(\{\bar{v}^{*N}_{t,\tau}\}') - p(\{\bar{v}^{*}_{t,\tau}\})\|_\infty < \eta/L.$$

Define $q_{t,\tau} \triangleq \sum_{l=1}^{L} R^l_{t,\tau} p_l(\{\bar{v}^{*N}_{t,\tau}\})$ and $q'_{t,\tau} \triangleq \sum_{l=1}^{L} R^l_{t,\tau} p_l(\{\bar{v}^{*N}_{t,\tau}\}')$. Then for every t and τ,

$$|q^*_{t,\tau} - q_{t,\tau}| < \eta,$$

and

$$|q^*_{t,\tau} - q'_{t,\tau}| < \eta.$$

Therefore we have

$$U_t(U_t'^{-1}(q_{t,\tau})) \geq U_t(z^*_t) - \epsilon/2 \geq U_t(U_t'^{-1}(q^*_{t,\tau'})) - \epsilon/2 \geq U_t(U_t'^{-1}(q'_{t,\tau'})) - \epsilon.$$

Since t and τ is arbitrary, we conclude that $\{\bar{v}^{*N}_{t,\tau}\}$ is a ϵ-Nash equilibrium and we have the following proposition,

Proposition 6.7 *For any $\epsilon > 0$, there exists an integer $N(\epsilon)$ such that for any $N > N(\epsilon)$ there exists ϵ-Nash equilibrium of the routing game of $\mathcal{N}_N$.*

We use the notation $\mathbb{R}_{++}$ for the open set of strictly positive real numbers. We will need the following fact before studying the property of ϵ-Nash equilibrium of the routing game,

Proposition 6.8 *For any $\epsilon > 0$, there exists a compact set $K \subset \mathbb{R}^T_{++}$ such that for all N, $q^N_{t,\tau}(\{\bar{v}^N_{t,\tau}\}) \in K$ where $\{\bar{v}^N_{t,\tau}\}$ is any ϵ-Nash equilibrium distribution of $\mathcal{N}$.*

Proof: Suppose there does not exist such K. We will only consider the situation where there exists a sequence of ϵ-Nash equilibrium $\{\bar{v}^N_{t,\tau}\}$ such that its corresponding $\min_{t,\tau} q^N_{t,\tau}(\{\bar{v}^N_{t,\tau}\}) \to 0$ as $N \to \infty$. Since there are only finite number of ts and τs, there exists a subsequence such that $\liminf_{N\to\infty} q^N_{t,\tau} = 0$ for a fixed pair (t,τ) and we will identify the subsequence as the original sequence for convenience. We now argue that there exists an integer $N(\epsilon)$, such that for all $N > N(\epsilon)$, $\{\bar{v}^N_{t,\tau}\}$ is not an ϵ-Nash equilibrium and thus follows contradiction. It is easy to see with the bandwidth constraint, there exists $\tau' \in m_t$ such that $q^N_{t,\tau'} \to q_{t,\tau'} > 0$. Let us consider one type t user choosing route τ' to switch to route τ. Denoting the route distribution after this switch by $\{\bar{v}^N_{t,\tau}\}'$ and corresponding aggregate price by $q'^N_{t,\tau}$. Then for N sufficiently large by the continuity argument $q'^N_{t,\tau} \approx q^N_{t,\tau} \ll q^N_{t,\tau'} \approx q'^N_{t,\tau'}$ and therefore $U_t(U'^{-1}_t(q'^N_{t,\tau})) \gg U_t(U'^{-1}_t(q^N_{t,\tau'}))$ so the conclusion follows. ■

Now we state the main result of this section,

Theorem 6.4 *For any $\delta > 0$, there exists a $\epsilon(\delta) > 0$ such that for any $0 < \epsilon < \epsilon(\delta)$ there is a $N(\epsilon,\delta)$ and for all $N > N(\epsilon,\delta)$ any ϵ-Nash equilibrium utilities $U^N_{t,\tau}$s for $\mathcal{N}_N$ satisfies $\max_{t,\tau} |U^N_{t,\tau} - U^{N*}_{t,\tau}| \le \delta$. Here $U^{N*}_{t,\tau}$ is the optimal utility for the type t user using route τ.*

Proof: It suffices to show that the argument $\max_{t,\tau} |U^N_{t,\tau} - U^*_{t,\tau}| < \delta$ where $U^*_{t,\tau}$ is the optimal utility of the multipath version (6.14), since $U^{N*}_{t,\tau}$ converges to $U^*_{t,\tau}$ when $N \to \infty$. Suppose the statement does not hold. That is, for any $\varepsilon > 0$, and any $N > 0$, there exists $n > N$ and $0 < \epsilon < \varepsilon$ such that $\max_{t,\tau} |U^n_{t,\tau} - U^*_{t,\tau}| > \delta$ for a ϵ-Nash equilibrium $\{\bar{v}^N_{t,\tau}\}$. Therefore we can have two infinite sequences $\{\epsilon_n\}$

and $\{N_n\}$, such that $\epsilon_n > \epsilon_{n+1}$ and $N_{n+1} > N_n$ for all n, $\epsilon_n \to 0$ and $N_n \to \infty$ as $n \to \infty$, $|U_{t,\tau}^{N_n} - U_{t,\tau}^*| > \delta$ where $U_{t,\tau}^{N_n}$ corresponds to a ϵ_n-Nash equilibrium $\{\bar{v}_{t,\tau}^{N_n}\}$ of $\mathcal{N}_{N_n}$ for a fixed (t, τ) (since we can always take subsequence due to finiteness of t and τ). By Proposition 6.8, all the aggregate price $q_{t,\tau}^{N_n}$ of ϵ_n-Nash equilibrium belongs to a compact set K. Then by uniform continuity of a continuous function over a compact set, there exists an infinite sequence $\{\eta_n\}$ such that $\eta_n > \eta_{n+1} > 0$ for all n, $\eta_n \to 0$ as $n \to \infty$, and 1) $|q_{t,\tau}^{N_n} - q_{t,\tau'}^{N_n}| < \eta_n$ for all $\tau, \tau' \in m_t$ such that $\bar{v}_{t,\tau}^{N_n} > 0$ and $\bar{v}_{t,\tau'}^{N_n} > 0$, and 2) $q_{t,\tau}^{N_n} < q_{t,\tau'}^{N_n} - \eta_n$ for all $\tau, \tau' \in m_t$ such that $\bar{v}_{t,\tau}^{N_n} > 0$ and $\bar{v}_{t,\tau}^{N_n} = 0$. Since $\bar{v}_{t,\tau}^{N_n}$ also belongs to a compact set, we can assume by taking subsequence if necessary that $\bar{v}_{t,\tau}^{N_n} \to \bar{v}_{t,\tau}$ as $n \to \infty$ for all t and τ. Therefore in asymptotics we have the Nash equilibrium strategy $\bar{v}_{t,\tau}$ with 1) $q_{t,\tau} = q_{t,\tau'}$ for all $\tau, \tau' \in m_t$ such that $\bar{v}_{t,\tau} > 0$ and $\bar{v}_{t,\tau'} > 0$, and 2) $q_{t,\tau} \leq q_{t,\tau'}$ for all $\tau, \tau' \in m_t$ such that $\bar{v}_{t,\tau} > 0$ and $\bar{v}_{t,\tau} = 0$. However, we still have $|U_{t,\tau} - U_{t,\tau}^*| > \delta$. This is in contradiction with the fact that only the optimal solution of the multipath problem (6.14) has this property with the aggregate prices. ■

Therefore we conclude that the simplistic local routing algorithm, Nash dynamics, leads to the optimal solution of the combined single path routing and flow control problem (6.1) when the number of users becomes sufficiently large in a general network.

6.4 Discussion

Our focus on the combined single path routing and flow control problem in this section is mostly on the descriptive side. That is, we have shown that approximate Nash equilibria are sufficiently close to the social optimal in many-user region, although we have not provided a precise rate or bound of this convergence. It is intuitive from Proposition 6.6 that it will be helpful if we can exactly find those $L+1$ users who causes the difference between the multipath problem and the single path problem. But since the Shapley-Folkman theorem is nonconstructive, it is difficult to go in that direction. It is unknown whether the computational complexity of the combined single path routing and flow control problem for a fixed number of bottleneck links is still NP hard. These are topics for future study.

Appendix A

Lyapunov-Krasovskii Functional for Delay-Independent Rate Controlled Network

For the network optimization problem

$$\begin{aligned} &\max_{x_i \geq 0} \textstyle\sum_{i=1}^N U_i(x_i) - \sum_{j=1}^L \int_0^{y_j} p_j(y)dy \\ &\text{s.t. } \textstyle\sum_{i=1}^N R_{ji} x_i = y_j, \forall j \in [L], \end{aligned} \tag{A.1}$$

Ranjan and La [20, 32] demonstrated the global delay-independent stability condition for the primal algorithm

$$\begin{aligned} \dot{x}_i(t) &= k_i \left(x_i(t) U_i'(x_i(t)) - x_i(t-\tau_i) \textstyle\sum_{j=1}^L R_{ji} p_j(y(t-\tau_{ij}^b)) \right), \\ y(t) &= \textstyle\sum_{j=1}^N R_{ji} x_i(t-\tau_{ij}^f). \end{aligned} \tag{A.2}$$

In particular, for the power utility $U_i(x) = -1/(a_i x^{a_i})$ and the power marking function $p_j(x) = x^{b_j}$, the result shows that the primal algorithm is globally stable if for every i, $a_i > \max\{b_j | R_{ji} = 1\} + 1$. Their method is based on contraction mapping by bounding the solution trajectory. In the section we introduce an alternative way by using Lyapunov-Krasovskii method to reach the same result.

The approach for designing the Lyapunov-Krasovskii functional is first we search for a Lyapunov-Razumikhin function for the system and then rely on the equivalence relation between Lyapunov-Razumikhin function and ISS small gain property pointed out by [31] to obtain the final Lyapunov-Krasovskii functional.

First let us consider the simple single source/link version of the above algorithm,

$$\dot{x}(t) = k\left(\frac{1}{x(t)^a} - x(t-\tau)^{b+1}\right). \tag{A.3}$$

By changing of the variable $y = \log x$ we have

$$\dot{y}(t) = ke^{-y(t)}(e^{-ay(t)} - e^{(b+1)y(t-\tau)}).$$

Consider the function $V(y) = y^2/2$, then the following always holds when $a > b+1$

$$\frac{d}{dt}V(y) \leq ky(t)e^{-y(t)}(e^{-ay(t)} - e^{-(b+1)qy(t)}) \leq 0$$

for all $y(t-\tau)$ such that $q^2V(y) \leq V(y(t-\tau))$ and $q \in (0,1)$. Therefore $V(y)$ is a Lyapunov-Razumikhin function and by Theorem 3.2 the system is globally asymptotically stable. The way of changing the state variable from x to y can be extended to the general utility/marking functions. To this end let us consider a single source/link version of (A.2)

$$\dot{x}(t) = k(x(t)U'(x(t)) - x(t-\tau)p(x(t-\tau))). \tag{A.4}$$

By changing of the variable with $y(t) = f(x(t)) \triangleq x(t)U'(x)$ and denoting $F(y) \triangleq f^{-1}(y)p(f^{-1}(y))$, the above system dynamics can be written as

$$\dot{y}(t) = k(y(t))(y(t) - F(y(t-\tau))),$$

where $k(y)$ is the according scaling function which is nonnegative. From the results of [20], if the mapping F has only one unique fixed point which is globally attracting, then the system is delay-independent stable. We are going to show that this can also be done by Razumikhin approach. Let us denote the globally attracting fixed point of F as $\bar{y}$, and define the following functions $F_l(y) \triangleq F(y)$ for all $y \in (0, \bar{y})$ and $F_r(y) \triangleq F(y)$ for all $y \in [\bar{y}, \infty)$. By global attraction of F we know that $F_r^{-1}(y) > F_l(y)$ for all $y \in (0, \bar{y})$. Then there exist a monotonically decreasing function $\psi(y)$ defined on $(0, \bar{y})$ such that $F_r^{-1}(y) > \psi(y) > F_l(y)$ for all $y \in (0, \bar{y})$ and $\psi(\bar{y}) = \bar{y}$ and $\psi(\bar{y})' = -1$. Next define the extension of $\psi(y)$ as $\phi(y)$ such that $\phi(y) = \psi^{-1}(y)$ for $y \in (\bar{y}, \infty)$. Therefore we have $\phi(y) > F_l(y)$ for $y \in (0, \bar{y})$ and $\phi(y) < F_r(y)$ for $y \in [\bar{y}, \infty)$. Define function $h : \mathbb{R} \to \mathbb{R}^+$ a strictly monotonically increasing function which satisfies $h(0) = 0$, $\lim_{u \to -\infty} h(u) = -\bar{y}$, $\lim_{u \to \infty} h(u) = \infty$, and $\phi(h(u) + \bar{y}) = h(-u) + \bar{y}$. Such h always exists given the construction of ϕ. Therefore by changing of variable again with $y = h(u(t)) + \bar{y}$ we arrive at

$$\dot{u}(t) = l(u(t))(h(u(t)) + \bar{y} - F(h(u(t-\tau)) + \bar{y}))$$

where $l(u)$ is an appropriate scaling function which is nonnegative. Consider the case when $|u(t)| > |u(t-\tau)|$. From the previous discussion the following inequalities hold

$$h(u) > F_l(h(-u) + \bar{y}) - \bar{y} \geq F(h(v) + \bar{y}) - \bar{y}, \forall u > 0, |u| > |v|,$$

$$h(u) < F_r(h(-u) + \bar{y}) - \bar{y} \leq F(h(v) + \bar{y}) - \bar{y}, \forall u < 0, |u| > |v|.$$

Therefore it can be verified directly that the function $V(u) = u^2/2$ is a Lyapunov-Razumikhin function and by Theorem 3.2 the global stability follows. We just demonstrated that the global delay-independent stability of any scalar system described in [20] can be shown using Razumikhin type argument.

Now that we have the Lyapunov-Razumikhin function for the system (A.4), we should be able to derive the Lyapunov-Krasovskii functional from the links between ISS small gain property, dissaptive form, and Lyapunov-Razumikhin function. Consider a system with input as follows,

$$\dot{x}(t) = f(x, u)$$

where the state variable $x \in \mathbb{R}^n$, the input variable $u \in \mathbb{R}^m$, and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$. Recall that a smooth function $V : \mathbb{R}^n \to \mathbb{R}^+$ is an ISS Lyapunov function is there exist K_∞-functions α_1, α_2, and K-functions α_3 and χ such that

$$\alpha_1(|\xi|) \leq V(\xi) \leq \alpha_2(|\xi|)$$

for any $\xi \in \mathbb{R}^n$ and $\nabla V(\xi) \cdot f(\xi, \mu) \leq -\alpha_3(|\xi|)$ for any $\xi \in \mathbb{R}^n$ and any $\mu \in \mathbb{R}^m$ such that $|\xi| \geq \chi(|\mu|)$. It is known [34] that the system is ISS with ISS gain γ if and only if there exist an ISS Lyapunov function and the gain can be represented by $\gamma = \alpha_1^{-1} \circ \alpha_2 \circ \chi$. Also from [34] previous ISS Lyapunov function exists if and only if there exist a smooth function $U : \mathbb{R}^n \to \mathbb{R}^+$ and K_∞-functions α_1, α_2, α, and σ such that

$$\alpha_1(|\xi|) \leq U(\xi) \leq \alpha_2(|\xi|)$$

for any $\xi \in \mathbb{R}^n$ and $\nabla U(\xi) \cdot f(\xi, \mu) \leq -\alpha(|\xi|) + \sigma(|\mu|)$. Here the corresponding ISS gain $\gamma = \alpha_1^{-1} \circ \alpha_2 \circ \alpha^{-1} \circ \sigma$. This second representation is of dissipative form and it is easy to use this form to derive Lyapunov function from the small-gain stability. Since the system (A.4) is proved to be stable by Lyapunov-Razumikhin type arguments, it is equivalent to say that the system is ISS with an appropriate gain if we take the delayed term as input. For illustrative purposes, let us again consider the case of (A.3). We argue that in the coordinate $y = \log x$, the system is ISS with gain $(b+1)/a$. To see this, define $V(x) = x^2/2$ as a ISS-Lyapunov candidate for the system with input u in the place of the delayed term,

$$\dot{x} = ke^{-x}(e^{-ax} - e^{(b+1)u}).$$

We have immediately

$$\frac{d}{dt}V(x) = ke^{-x}y(e^{-ax} - e^{(b+1)u}) \leq 0$$

if $|x| \geq \frac{b+1}{a}|u|$. So ISS holds with the ISS gain as $\gamma = (b+1)/a$. Since the delayed term can be viewed as ISS with unity gain, by an argument of small-gain theorem the stability follows immediately. In order to construct Lyapunov-Krasovskii functional, we need to use the dissipative representation of the ISS system. In other words, it is necessary to obtain K_∞ functions α and σ. Let us denote $g(x) \triangleq U'(x)$ and assume $\alpha(0) = \sigma(0) = 0$. Then it follows directly that these conditions have to be satisfied:

1. $\alpha(|x|) \leq -g(x)f(x,0)$.

2. $\sigma(|u|) \geq \max_x\{g(x)f(x,u) + \alpha(|x|)\}$

3. $\alpha(|\gamma x|) \geq \sigma(|x|)$.

In our system (A.3), $g(x)f(x,u)$ is positive only when $x \in (0, -\frac{b+1}{a}u)$ if $u < 0$ or $x \in (-\frac{b+1}{a}u, 0)$ if $u > 0$. In order to let the third inequality to hold, the maximum value of the right-hand side of the second inequality has to be achieved at $x = -\frac{b+1}{a}u$. Equivalently,

$$\frac{d}{dx}g(x)f(x,u)\bigg|_{u=-\frac{a}{b+1}x} + \frac{d}{dx}\alpha(|x|) = 0,$$

which is the same as

$$\frac{d}{dx}\alpha(|x|) = -g(x)\,\frac{d}{dx}f(x,u)\bigg|_{u=-\frac{a}{b+1}x}.$$

In the system (A.3), this equality can be rewritten as

$$\frac{d}{dx} = ah(x)e^{ax}$$

where $h(x) = kg(x)e^{-x}$. The first condition gives

$$\alpha(x) \leq h(x)(1 - e^{-ax}).$$

Let us write $\alpha(x) = h(x)\beta(x)(1 - e^{-ax})$ where $0 \leq \beta(x) \leq 1$, $\forall x \in \mathbb{R}$. Then we obtain

$$\frac{d}{dx}\alpha(x) = \frac{a\alpha(x)}{(e^{ax} - 1)\beta(x)}$$

But $\alpha(x) \equiv \alpha(-x)$. Therefore

$$\frac{a\alpha(x)}{(e^{ax} - 1)\beta(x)} \equiv \frac{-a\alpha(-x)}{(e^{-ax} - 1)\beta(-x)},$$

or $e^{ax}\beta(x) \equiv \beta(-x)$. For simplicity define $\beta(x) = 1$, $\forall x < 0$ and $\beta(x) = e^{-ax}$, $\forall x \geq 0$. We can then solve the above equation by

$$\alpha(x) = e^{a|x|} - 1.$$

Therefore we obtain the dissipative form for the system (A.3) in the changed coordinate and the Lyapunov-Krasovskii functional $W'(\cdot)$ can be shown easily from the dissipative form as

$$W'(x_t) = U(x_t(0)) + \int_{-\tau}^{0} (\exp((b+1)|x_t(s)|))\, ds$$

where

$$U(x) = \begin{cases} (2a+1)^{-1}k^{-1}(\exp((2a+1)x) - 1), & x \geq 0, \\ k^{-1}(1 - \exp(x)), & x < 0. \end{cases}$$

In the original coordinate, the Lyapunov-Krasovskii functional is

$$W(x_t) = \int_{-\tau}^{0} \left(x_t(s)^{(b+1)\mathrm{sgn}(x_t(s)-1)} - 1 \right) ds + \begin{cases} (2a+1)^{-1}k^{-1}\left(x^{2a+1} - 1\right), & x \geq 1, \\ k^{-1}(1 - x), & x < 1. \end{cases}$$

From the above single source/link system. we can also build a Lyapunov-Krasovskii functional for the network system. For example, consider two rate controlled flows with one bottleneck link. Its system dynamics is

$$\dot{x}_1(t) = x_1(t)^{-a_1} - x_1(t-\tau)(x_1(t-\tau) + x_2(t-\tau))^b,$$

$$\dot{x}_2(t) = x_2(t)^{-a_2} - x_2(t-\tau)(x_1(t-\tau) + x_2(t-\tau))^b.$$

We can use the following Lyapunov-Krasovskii functional

$$W = \frac{\bar{x}_1}{\bar{x}_1 + \bar{x}_2}(W_1(x_1) + V_1(x_{1t}, x_{2t})) + \frac{\bar{x}_2}{\bar{x}_1 + \bar{x}_2}(W_2(x_2) + V_2(x_{1t}, x_{2t}))$$

where the intermediate functions are defined as

$$W_j(x) \triangleq \begin{cases} (2a_j+1)^{-1}\bar{x}_j^{a_j+1}\left(\left(\frac{x_j}{\bar{x}_j}\right)^{2a_j+1} - 1\right), & x_j \geq \bar{x}_j, \\ \bar{x}_j^{a_j+1}\left(1 - \frac{x_j}{\bar{x}_j}\right), & x_j < \bar{x}_j. \end{cases} \qquad j = 1, 2.$$

and

$$V_j(x_{1t}, x_{2t}) = \int_{-\tau}^{0} \left(\left(\frac{x_j(s)}{\bar{x}_j}\right)^{\mathrm{sgn}(x_j(s) - \bar{x}_j} \left(\frac{x_1(s) + x_2(s)}{\bar{x}_1 + \bar{x}_2}\right)^{b\mathrm{sgn}(x_1(s) + x_2(s) - \bar{x}_1 - \bar{x}_2)} - 1 \right) ds,$$

$$j = 1, 2.$$

Appendix B

Proof of Lemma 4.4

Following the discussion in Section 4.4.2 it is sufficient to check whether

$$-1 \notin \bar{\text{co}}(0 \cup \{g(i\omega, \tau), \forall \tau \geq 0\})$$

holds for every $\omega \geq 0$. Here we define

$$g(s, \tau) \triangleq \frac{e^{-\tau s}}{\tau^2 s^2} \frac{\tau s + \frac{1}{2}\min\{\tau, 1\}}{2\max\{\tau, 1\}}.$$

The proof then breaks into the examination of the 3 parts of the curve $g(i\omega, \tau)$ on the Nyquist plane for any fixed ω. First we study the part of the curve after it crosses the real axis for the first time. Then we study the situation of the curve before it crosses the real axis, where the cases when $\tau \leq 1$ and $\tau > 1$ are studied separately.

First by direct calculation we obtain that for fixed ω the first intersection of $g(i\omega, \tau)$ with the real axis takes place at $\tau_0 = \omega_1/\omega$ if $\omega < \omega_1$, where $\omega_1 \approx 1.1656$ is the solution of the equation $2\omega = \tan\omega$, and $\tau_0 = (\arctan 2\omega)/\omega$ if $\omega \geq \omega_1$. The location of the intersection is

$$\begin{aligned} & -1/(2\omega\tau_0 \sin\omega\tau_0 \max\{\tau_0, 1\}) \\ \geq \quad & -1/(2\omega_1 \sin\omega_1) \approx -0.4668. \end{aligned}$$

The maximum value of imaginary part attained by the curve $g(i\omega, \tau)\max\{\tau, 1\}$ with fixed ω is obtained by maximizing

$$\text{Im}g(i\omega, \tau)\max\{\tau, 1\} = \frac{\sin\omega\tau}{4\omega^2\tau^2} - \frac{\cos\omega\tau}{2\omega\tau}.$$

By numerical calculation the maximum value is $v_{\max} \approx 0.1824$ when $\omega\tau \approx 2.5288$. We can then show that the part of the curve $g(i\omega, \tau)$ at which $\tau \geq \tau_0$, or equivalently the part after passing the real axis, lies below the affine line $\mathcal{L}$ defined by $\mathrm{Im}z = \omega(\mathrm{Re}z + 1)$, since the slope of the line passing through -1 under which our curve lies is less than

$$
\begin{aligned}
& \frac{v_{\max}}{\left(1-\frac{1}{2\omega_1 \sin\omega_1}\right)\max\{\tau,1\}} \\
< \quad & \frac{v_{\max}}{\left(1-\frac{1}{2\omega_1 \sin\omega_1}\right)\frac{\omega_1}{\omega}} \approx 0.2964\omega.
\end{aligned}
$$

Therefore the argument is valid.

Now let us inspect the part of the curve before passing the real axis. There are 2 situations. When $\tau \leq 1$ the curve can be written as

$$
g(i\omega, \tau) = -\frac{e^{-i\tau\omega}}{\tau^2\omega^2}\frac{i\tau\omega + \tau/2}{2} = -\frac{e^{-i\tau\omega}}{\tau\omega}\left(i\frac{1}{2} + \frac{1}{4\omega}\right).
$$

We will show that in this situation the curve lies below $\mathcal{L}$. By some manipulations this is equivalent to the following inequality

$$
2\tau\omega^3 > (\omega^2+1)\sin\tau\omega - \frac{1}{2}\cos\tau\omega.
$$

One can verify that

$$
h_1(\omega) \triangleq 2\tau\omega^3 - (\omega^2+1)\sin\tau\omega + \frac{1}{2}\cos\tau\omega
$$

is the integral of the following

$$
\begin{aligned}
h_1'(\omega) = \quad & 6\tau\omega^2 - \tau\omega^2\cos\tau\omega - \tfrac{\tau-1}{2}\cos\tau\omega \\
& -(\tau/2+2)\omega\sin\tau\omega
\end{aligned}
$$

with respect to ω. The above is greater than zero for $\tau \in [0, 1]$ since

$$
h_1'(\omega) > 5\tau\omega^2 - (\tau/2+2)\omega\sin\tau\omega > \frac{5}{2}\omega(\tau\omega - \sin\tau\omega) > 0.
$$

Since $h_1(0) = 1/2 > 0$, $h_1(\omega) = h_1(0) + \int_0^\omega h_1'(u)du > 0$. We then conclude that when $\tau \in [0, 1]$ the curve lies below the line $\mathcal{L}$.

In the other situation when $\tau > 1$ the curve can be expressed as

$$g(i\omega, \tau) = -\frac{e^{-i\tau\omega}}{\tau^2\omega^2}\frac{i\tau\omega + 1/2}{2\tau}.$$

We will just consider the curve

$$\tilde{g}(i\omega, \tau) = -\frac{e^{-i\tau\omega}}{\tau^2\omega^2}\frac{i\tau\omega + 1/2}{2}.$$

since the curve $\tilde{g}$ lies above the curve g. Again by using the simplifying notation $\theta = \tau\omega > \omega$ and after some algebraic manipulations, it suffices to show the validity of the following inequality

$$4\omega\theta^2 > \omega(2\theta \sin\theta + \cos\theta) + (\sin\theta - 2\theta\cos\theta).$$

It is actually sufficient to check the above inequality when $\theta = \omega$. So we will only need to show

$$4\theta^3 > (2\theta^2 + 1)\sin\theta - \theta\cos\theta.$$

Similarly to the situation when $\tau \leq 1$, we define a function

$$h_2(\theta) \triangleq 4\theta^3 - (2\theta^2 + 1)\sin\theta + \theta\cos\theta.$$

Its derivative is

$$h_2'(\theta) = 12\theta^2 - 2\theta^2\cos\theta - 5\theta\sin\theta.$$

But

$$h_2'(\theta) > 10\theta^2 - 5\theta\sin\theta > 5\theta(\theta - \sin\theta) > 0,$$

and by $h_2(0) = 0$, one obtains $h_2(\theta) = h_2(0) + \int_0^\theta h_2'(u)du > 0$. Therefore we have shown that the curve $\tilde{g}$ lies below the line $\mathcal{L}$ when $\tau > 1$.

Combining all these results we have confirmed that for any fixed ω, the curve $g(i\omega, \tau)$ lies below the line $\mathcal{L}$ for all $\tau \geq 0$ and therefore the convex hull of $g(i\omega, \tau)$ and 0 cannot contain the point -1 on the Nyquist plane.

[1] F.P. Kelly, A. Maulloo, and D. Tan. Rate control for communication networks: Shadow prices, proportional fairness and stability. *J. Oper. Res. Soc.*, 49:237–252, March 1998.

[2] P. Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1968.

[3] H. J. Kushner. *Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory*. The MIT Press, 1984.

[4] M. Vojnovic, J.Le Boudec, and C. Boutremans. Global fairness of additive-increase and multiplicative-decrease with heterogeneous round-trip times. *Proceedings of IEEE Infocom*, March 2000.

[5] S. Shakkottai and R. Srikant. Mean FDE models for Internet congestion control under a many-flows regime. *IEEE Transactions on Information Theory*, 50:1050–1072, 2004.

[6] P. Tinnakornsrisuphap and A. Makowski. Limit behavior of ECN/RED gateways under a large number of tcp flows. *Proceedings of IEEE Infocom*, April 2003.

[7] S. Deb and R. Srikant. Rate-based versus queue-based models of congestion control. *IEEE Transactions on Automatic Control*, 51:606–619, 2006.

[8] H. Chen and D. D. Yao. *Fundamentals of Queueing Networks*. New York: Springer-Verlag, 2000.

[9] S. Low, F. Paganini, J. Wang, and J. Doyle. Linear stability of TCP/RED and a scalable control. *Computer Networks Journal*, 43(5):633–647, December 2003.

[10] G. Vinnicombe. On the stability of networks operating TCP-like congestion control. *Proc. IFAC*, 2002.

[11] C. Hollot and Y. Chait. Nonlinear stability analysis for a class of TCP/AQM. *IEEE Conference on Decision and Control*, December 2001.

[12] C. Hollot, V. Misra, D. Towsley, and W.B. Gong. A control theoretic analysis of RED. *Proceedings of IEEE Infocom*, April 2001.

[13] F. P. Kelly. Models for a self-managed Internet. *Phil. Trans. Royal. Soc.*, A358:2335–2348, 2000.

[14] D. Schmeidler and K. Vind. Fair net trades. *Econometrica*, 40(4):637–642, 1972.

[15] J. Mo and J. Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Trans. on Networking*, 8(5):556–567, 2000.

[16] W. Hildenbrand. *Core and Equilibria of a Large Economy*. Princeton University Press, Princeton, 1974.

[17] S.H. Low and D.E. Lapsley. Optimization flow control I: Basic algorithm and convergence. *IEEE/ACM Trans. Networking*, 7:861–874, December 1999.

[18] M. Johari and D. Tan. End-to-end congestion control for the Internet: Delays and stability. *IEEE/ACM Trans. Networking*, 9(6):818–832, December 2001.

[19] X. Fan, M. Arcak, and J.T. Wen. Robustness of network flow control against disturbances and time-delay. *System and Control Letters*, 53(1):13–29, 2004.

[20] R. Ranjan, R.J. La, and E.H. Abed. Global stability conditions for rate control with arbitrary communication delays. *IEEE Conference on Decision and Control*, December 2004.

[21] J.T. Wen and M. Arcak. A unifying passivity framework for network flow control. *IEEE Trans. Automatic Control*, 49(2):162–174, 2004.

[22] A. Papachristodoulou. Global stability analysis of a TCP/AQM protocol for arbitrary networks with delay. *IEEE Conference on Decision and Control*, December 2004.

[23] V. Jacobson, R. Braden, and D. Borman. TCP extenstions for high performance, May 1993.

[24] R. Johari and J.N. Tsitsiklis. Efficiency loss in a network resource allocation game: The case of elastic supply. *IEEE Trans. on Automatic Control*, June 2004. in press.

[25] D. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, 1982.

[26] C. Hollot, V. Misra, D. Towsley, and W.B. Gong. On designing improved controllers for aqm routers supporting TCP flows. *Proceedings of IEEE Infocom*, April 2001.

[27] H.-O. Walther. Contracting return maps for some delay differential equations. In T. Faria and P. Freitas, editors, *Functional Differential and Difference Equations*, Fields Inst. Commun. 29, pages 349–360. American Mathematical Society, 2001.

[28] M. Peet and S. Lall. On global stability of Internet congestion control. *IEEE Conference on Decision and Control*, December 2004.

[29] J. Hale and S.M. Verduyn Lunel. *Introduction to Functional Differential Equations*. Springer-Verlag New York, Inc., 1993.

[30] S. Deb and R. Srikant. Global stability of congestion controllers for the Internet. *IEEE Trans. Automatic Control*, 48(6):1055–1060, June 2003.

[31] A. Teel. Connections between razumikhin-type theorems and the ISS nonlinear small gain theorem. *IEEE Trans. Automatic Control*, 43(7):960–964, July 1998.

[32] P. Ranjan, R.J. La, and E.H. Abed. Global stability with a state-dependent delay in rate control. *IFAC TDS*, 2004.

[33] J-H Kim. Delay and its time-derivative dependent robust stability of time-delayed linear systems with uncertainty. *IEEE Trans. on Automatic Control*, 46(5):789–792, May 2001.

[34] E.D. Sontag and Y. Wang. On characterizations of the input-to-state stability property. *Systems and Control Letters*, 24:351–359, 1995.

[35] Z.P. Jiang, I.M. Mareels, and Y. Wang. A Lyapunov formulation of the nonlinear small-gain theorem for interconnected iss systems. *Automatica*, 32:1211–1215, 1996.

[36] H. Ito. A constructive proof of ISS small-gain theorem using generalized scaling. *IEEE Conference Decision and Control*, December 2002.

[37] Z. Wang and F. Paganini. Global stability with time-delay in network congestion control. *IEEE Conference on Decision and Control*, December 2002.

[38] E. Liz, V. Tkachenko, and S. Trofimchuk. York and Wright 3/2-stability theorems from a unified point of view. *Proc of the 14th International Conference on Dynamical Systems and Differential Equations*, May 2002.

[39] J. C. Lillo. Oscillatory solutions of the equation $y'(x) = m(x)y(x - n(x))$. *Journal of Differential Equations*, 6:1–36, 1969.

[40] F. Paganini, Z. Wang, J. Doyle, and S.H. Low. Congestion control for high performance, stability and fairness in general networks. *IEEE/ACM Trans. Networking*, 2004.

[41] W.J. Rugh. *Linear System Theory*. Prentice Hall, second edition, 1999.

[42] K.S. Narendra and J.H. Taylor. *Frequency Domain Criteria for Absolute Stability*. Academic Press, 1973.

[43] A. Halanay. *Differential Equations: Stability, Oscillations, Time Lags*. Academic Press, 1966.

[44] X. Li. On the absolute stability of systems with time lags. *Chinese Mathematics*, 4:609–626, 1963.

[45] C.A. Desoer and Y.T. Yang. On the generalized Nyquist stability criterion. *IEEE Transactions on Automatic Control*, 25:187–196, 1980.

[46] G. Zames and P. L. Falb. Stability conditions for systems with monotone and slope-restricted nonlinearities. *SIAM Journal on Control*, 6:89–108, 1968.

[47] T. Bonald and L. Massoulie. Impact of fairness on Internet performance. In *Proceedings of ACM Sigmetrics*, 2001.

[48] P. Key, L. Massoulie, and D. Towsley. Combining multipath routing and congestion control for robustness. In *40th Conference on Information Sciences and Systems*, Princeton, March 2006.

[49] H. Han, S. Shakkottai, C. Hollot, R. Srikant, and D. Towsley. Overlay TCP for multi-path routing and congestion control. In *ENS-INRIA ARC-TCP Workshop*, Paris, France, 2003.

[50] F. Kelly and T. Voice. Stability of end-to-end algorithms for joint routing and rate control. *Computer Communication Review*, 35(2):5–12, 2005.

[51] T. Voice. A global stability result for primal-dual congestion control algorithms with routing. *Computer Communication Review*, 34(3):35–41, 2004.

[52] W. H. Wang, M. Palaniswammi, and S. H. Low. Optimal flow control and routing in multi-path networks. *Performance Evaluation*, 52(2-3):119–132, 2003.

[53] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 1995.

[54] S. Sastry. *Nonlinear Systems*. Springer, New York, 1999.

[55] K. Kar, S. Sarkar, and L. Tassiulas. Optimization based rate control for multipath sessions. Technical Report 2001-1, Institute for Systems Research, University of Maryland, 2001.

[56] X. Lin and N. B. Shroff. Utility maximization for communication networks with multipath routing. *IEEE Transactions on Automatic Control*, 51(5):766–781, 2006.

[57] N. Z. Shor. *Minimization Methods for Non-Differentiable Functions and Applications*. Springer-Verlag, Berlin, 1985.

[58] D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Prentice-Hall, Upper Saddle River, NJ, 1989.

[59] J. Wang, L. Li, S. H. Low, and J. C. Doyle. Cross-layer optimization in TCP/IP networks. *IEEE/ACM Transactions on Networking*, 13(1):43–56, February 2005.

[60] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of Np-Completeness*. W. H. Freeman, New York, 1979.

[61] S. Mertens. A physicist's approach to number partitioning. *Theoretical Computer Science*, 2001.

[62] E. Koutsoupias and C. H. Papadimitriou. Worst-case equilibrium. *Lecture Notes in Computer Science*, 1999.

[63] T. Roughgarden. *Selfish Routing*. PhD thesis, Cornell University, 2002.

[64] A. Czumaj and B. Vocking. Tight bounds for worst-case equilibria. *SODA 02*, 2002.

[65] P. Dubey, A. Mas-Colell, and M. Shubik. Efficiency properties of strategic market games: an axiomatic approach. *J. Econ. Theory*, pages 339–363, April 1980.

[66] L. Shapley and M. Shubik. Trade using one commodity as a means of payment. *J. Pol. Econ.*, 85:937–968, 1977.

[67] A. Mas-Colell. The Cournotian foundations of Walrasian equilibrium theory: an exposition of recent theory. In W. Hildenbrand, editor, *Advances in Economic Theory*. Cambridge University Press, 1982.

[68] J. Roberts and A. Postlewaite. The incentive of price-taking behavior in large exchange economies. *Econometrica*, 44:115–127, 1976.

[69] J. Falk. Lagrange multipliers and nonconvex programs. *SIAM J. Control*, 7:534–545, 1969.

[70] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems*. North Holland Publishing Company, Amsterdam, 1976.

[71] R. Rockafellar and R. Wets. *Variational Analysis*. Springer, Berlin, 1998.