TECHNICAL RESEARCH REPORT

Trade-offs in rate control with communication delay

by Priya Ranjan, Richard J. La, and Eyad H. Abed

TR 2003-27

Trade-offs in Rate Control with Communication Delay

Priya Ranjan, Richard J. La, and Eyad H. Abed
University of Maryland, College Park
{priya, hyongla, abed}@eng.umd.edu

Abstract

We adopt the optimization framework for rate allocation problem proposed by Kelly and characterize the stability condition with an *arbitrary* communication delay in the case of single resource. We demonstrate the existence of a fundamental trade-off between users' price elasticity of demand and the responsiveness of resource through a choice of price function as well as between system stability and resource utilization. We investigate the effects of non-responsive traffic on system stability and show that the presence of non-responsive traffic enhances the stability of system. We also investigate the system behavior after the system loses its stability.

Keyword – Economics, Control theory, Mathematical programming/optimization

1 Introduction

With the unprecedented growth and popularity of the Internet the problem of rate/congestion control is emerging as a more crucial problem. Poor management of congestion can render one part of a network inaccessible to the rest and significantly degrade the performance of networking applications. The fact that the Internet is now in the public domain, and thus in a potentially non-cooperative environment, has stimulated much work on designing rate control mechanisms based on some form of pricing mechanism to ensure that users do not misbehave.

Kelly [12] has suggested that the problem of rate allocation for elastic traffic can be posed as one of achieving maximum aggregate utility of the users and proposed an optimization framework for rate allocation in the Internet. Using the proposed framework he has shown that the system optimum is achieved at the equilibrium between the end users and resources. Based on this observation researchers have proposed various rate-based algorithms that solve the system optimization problem or its relaxation [12, 18]. However, the convergence of these algorithms has been established only in the absence of feedback delay, and the implications of feedback delay have been left open as well as any trade-off that may exist between stability and selected utility and cost functions. Modeling the communication delay is especially important when the delay is non-negligible, *e.g.,* multi-hop mobile wireless networks. Tan and Johari have studied the case with homogeneous users, *i.e.,* same round-trip delays and same form of utility functions, and provided local stability conditions in term of users' gain parameters and communication delays. In general their results state that the product of gain parameter and communication delays should be no larger than some constant. Similar results have been obtained in [4] in the context of single flow and single resource and [2] with more general utility functions in the context of single bottleneck. The stability conditions state that the product of the delay and gain parameter of end user algorithms needs to be smaller than some constant that depends on the utility function of the users. However, these results focus on characterizing *sufficient* conditions for stability and, as a result, do not point at a close relationship between system stability and the parameters at the end users and network elements as is done in this paper.

In this paper we study the problem of designing a robust rate control algorithm in the presence of a communication delay between network resources and end users. However, unlike in the previous studies where the authors give the conditions for stability of the system, we establish a *delay-independent* stability criterion for system optimization problem in the presence of an arbitrary delay. Our approach is consistent with the philosophy that network protocols must be simple and robust given the complexity and scale of the Internet. This also provides a fresh way of looking at the issue of communication delay than traditional approaches. A natural question that arises in this setting is whether or not it is possible to design a system that is stable with an arbitrary communication delay. If it is possible, what are the necessary and/or sufficient conditions for the stability? In addition, what is the impact of the non-responsive traffic on the system stability? The last question is emerging as an important issue with a growing interest of implementing real-time applications on top of User Datagram Protocol (UDP), which are not as responsive as elastic traffic.

Our analysis is based on the invariance-based global stability results for nonlinear delay-differential equations [9, 10, 19]. This kind of global stability results are different from that based on Lyapunov or Razumikhin theorems used in [2, 4, 27] or from passivity approach [28], and our set up also hints at the structure of emerging periodic orbits (such as their periodicity and amplitude) in the case of loss of stability. Generally speaking, our results can be summarized as follows. First, there is a close relationship between the stability of a delay-differential equations that describes network dynamics and an underlying discrete time map. Second, if the user and resource curves have a stable market equilibrium, which is captured by the underlying discrete time map, then corresponding dynamical equation for flow optimization will converge to the optimal point in the presence of an arbitrary delay. This result essentially shows that stability is related to utility and price curves in a fundamental way. In particular, for a given price curve, it is possible to design stable user utility functions such that the ensuing dynamical system converges to the optimal flow irrespective of communication delay. Conversely, if the underlying market equilibrium is unstable then it is possible to find a large enough delay for which the optimal point loses its stability and gives way to oscillations. In practice, this gives rise to a fundamental trade-off between the responsiveness of end users and network resources. In other words, given the responsiveness of network resource, there is a limit on how aggressively *selfish* end users can react to feedback in order to ensure delay-independent network stability. Our results reveal another trade-off between delay-independent network stability and resource utilization. These results provide an interesting perspective for designing end user algorithms and active queue management (AQM) mechanisms.

It is worth noting that in general characterizing the exact necessary and sufficient conditions for stability with a delay is difficult. Hence, our results provide a *simple* and *robust* way of dealing with the problem of widely varying feedback delay in communication networks through a clever choice of the users' utility functions and price functions. We also study the oscillatory orbits that appear when the system loses stability by explicitly giving the bounds on their amplitude. It is shown that these bounds are derived from an underlying discrete-time map that goes through a period doubling bifurcation with the loss of stability. Finally, we investigate how the presence of non-responsive traffic affects the system stability. Our analysis indicates that the presence of non-responsive traffic, which in fact can be thought of as the limiting case of elastic traffic with decreasing responsiveness, improves the system stability. This is consistent with our earlier results that the less responsive users are, the more stable system is.

This paper is organized as follows. Section 2 describes the optimization problem for rate control. Section 3 studies the single flow case, which is followed by the multiple heterogeneous users case in Section 4. We illustrate how our results can be applied to the rate control problem in networks in subsections 5. The effects of non-responsive flows are investigated in Section 6.

2 Background

In this section we briefly describe the rate control problem in the proposed optimization framework. Consider a network with a set $\mathcal{L}$ of resources or links and a set $\mathcal{I}$ of users. Let C_l denote the finite capacity of link $l \in \mathcal{L}$. Each user has a fixed route r_i, which is a non-empty subset of $\mathcal{L}$. We define a zero-one matrix A, where $A_{i,l} = 1$ if link l is in user i's route r_i and $A_{i,l} = 0$ otherwise. When the throughput of user i is x_i, user i receives utility $U_i(x_i)$. The utility $U_i(x_i)$ is an increasing, strictly concave and continuously differentiable function of x_i over the range $x_i \geq 0$.[1] Furthermore, the utilities are additive so that the aggregate utility of rate allocation $x = (x_i, i \in \mathcal{I})$ is $\sum_{i\in\mathcal{I}} U_i(x_i)$. Let $U = (U_i(\cdot), i \in \mathcal{I})$ and $C = (C_l, l \in \mathcal{L})$. The rate control problem can be formulated as the following optimization problem:
SYSTEM(U,A,C):

$$\begin{aligned} \text{maximize} \quad & \sum_{i\in I} U_i(x_i) \\ \text{subject to} \quad & A^T x \leq C \ , \ x \geq 0 \end{aligned} \tag{1}$$

The first constraint in the problem says that the total rate through a resource cannot be larger than the capacity of the resource. Instead of solving (1) directly, which is difficult for any large network, Kelly in [12] has proposed to consider the following two simpler problems.

Suppose that each user i is given the price per unit flow λ_i. Given λ_i, user i selects an amount to pay per unit time, w_i, and receives a flow $x_i = \frac{w_i}{\lambda_i}$.[2] Then, the user's optimization problem becomes the following [12].
$USER_i(U_i; \lambda_i)$:

$$\begin{aligned} \text{maximize} \quad & U_i\left(\frac{w_i}{\lambda_i}\right) - w_i \\ \text{over} \quad & w_i \geq 0 \end{aligned} \tag{2}$$

The network, on the other hand, given the amounts the users are willing to pay, $w = (w_i, i \in \mathcal{I})$, attempts to maximize the sum of weighted log functions $\sum_{i\in\mathcal{I}} w_i \log(x_i)$. Then the network's optimization problem can be written as follows [12].
NETWORK(A,C;w) :

$$\begin{aligned} \text{maximize} \quad & \sum_{i\in I} w_i \log(x_i) \\ \text{subject to} \quad & A^T x \leq C \ , \ x \geq 0 \end{aligned} \tag{3}$$

Note that the network does not require the true utility functions $(U_i(\cdot), i \in \mathcal{I})$, and pretends that user i's utility function is $w_i \cdot \log(x_i)$ to carry out the computation. It is shown in [12] that one can always find vectors $\lambda^* = (\lambda_i^*, i \in \mathcal{I}), w^* = (w_i^*, i \in \mathcal{I})$, and $x^* = (x_i^*, i \in \mathcal{I})$ such that w_i^* solves $USER_i(U_i; \lambda_i^*)$ for all $i \in \mathcal{I}$, x^* solves $NETWORK(A, C; w^*)$, and $w_i^* = x_i^* \cdot \lambda_i^*$ for all $i \in \mathcal{I}$. Furthermore, the rate allocation x^* is also the unique solution to $SYSTEM(U, A, C)$.

Assume that every user adopts a rate-based flow control. Let $w_i(t)$ and $x_i(t)$ denote user i's willingness to pay per unit time and rate at time t, respectively. Now suppose that at time t each resource $l \in \mathcal{L}$ charges a price per unit flow of $\mu_l(t) = p_l(\sum_{i:l\in r_i} x_i(t))$, where $p_l(\cdot)$ is an increasing function of the total rate going through it. Consider the system of differential equations

$$\frac{d}{dt} x_i(t) = \kappa_i \big(w_i(t) - x_i(t) \sum_{l\in L} \mu_l(t)\big) \tag{4}$$

[1] Such a user is said to have elastic traffic.

[2] This is equivalent to selecting its rate x_i and agreeing to pay $w_i = x_i \cdot \lambda_i$.

where $\mu_l(t) = p_l(\sum_{i:l\in r_i} x_i(t))$. These equations can be motivated as follows. Each user first computes a price per unit time it is willingness to pay, namely $w_i(t)$. Then, it adjusts its rate based on the feedback provided by the resources in the network to equalize its willing to pay and the total price. The feedback from a resource $j \in J$ can also be interpreted as a congestion indicator, requiring a reduction in the flow rates going through the resource. For more detailed explanation of (4), refer to [14].

Kelly *et al.* have shown that under some conditions on $p_l(\cdot), l \in L$, the above system of differential equations converges to a point that maximizes the following expression

$$U(x) = \sum_i U_i(x_i) - \sum_l \int_0^{\sum_{i:l\in r_i} x_i} p_l(y)dy. \tag{5}$$

Note that the first term in (5) is the objective function in our *SYSTEM(U, A, C)* problem. Thus, the algorithm proposed by Kelly *et al.* solves a relaxation of the $SYSTEM(U,A,C)$ problem.

The analysis of the convergence of the rate control algorithm, however, does not model the communication delay that is present between the resources and the end users. There has been some work done on studying the stability of the system in the presence of communication delay. Tan and Johari [11] have analyzed the case where every user has the same round-trip delay and log utility function and given the conditions on local stability in terms of the gain parameter κ_i and and communication delay D_i. Moreover, they have shown the convergence rate of the system in the case of single-user single-resource. More recently Deb and Srikant [4] have investigated the stability of the system in the context of single flow and single resource. They have provided a sufficient condition for stability. However, as will be shown in this paper, the provided sufficient condition is not necessary and can be very restrictive depending on the range in which the initial condition lies. Alpcan and Basar [2] have also studied the stability of a system with a single resource and multiple flows and provided a sufficient condition for stability.

In this paper we investigate the global stability of the system. More specifically, we are interested in characterizing the condition on the users' utility functions and resource price functions in such a way that the system is guaranteed to be stable regardless of the communication delay or users' gain parameters κ_i. We also study the trade-off between the responsiveness of resource price functions and end users' utility functions, which can be captured using the notion of price elasticity of demand.

3 Stability Condition: Single-Flow, Single-Resource

In this section we first consider a flow traversing a single resource. The rate control problem can be formulated as the following net utility optimization problem [12]:

$$\begin{aligned} \max_x \quad & U(x) - \textstyle\int_0^x p(y)dy \\ \text{s. t.} \quad & x \le C \end{aligned} \tag{6}$$

where x is the rate, $U(x)$ is the utility of the user when it receives a rate of x, $p(x)$ is the price per unit flow when the rate is x, and C is the capacity of the resource. The proposed end user algorithm in the absence of delay is given by the following differential equation [14].

$$\frac{d}{dt}x(t) = \kappa(w(t) - x(t)\mu(t)) \tag{7}$$

where $w(t)$ is the price per unit time user is willing to pay, $\mu(t) = p(x(t))$, and $\kappa, \kappa > 0$, is a gain parameter. The case where $w(t)$ is a fixed constant, *i.e.,* $U(x) = \log(x)$, is studied in [13]. In this paper we consider general utility functions that satisfy a set of assumptions to be stated shortly. An example of a family of utility functions that satisfy such assumptions is given in Section 5.

The model used for design of end user rate control algorithm described here [13] does not explicitly address the case where the total demand of the users exceeds the link capacity. In practice total rate of the users (or at least the feedback from the resource) is limited by the link capacity. We prevent the resource feedback signal from exceeding the link capacity by making the following assumption:

Assumption 1 *We assume that the rate of the flow is bounded from above by the link capacity C.*

Therefore, throughout the rest of the paper we implicitly assume that when the rate of the flow reaches the link capacity, the time derivative is given by $\min(\dot{x}(t), 0)$. This assumption can be lifted if the solution to the optimization problem in (6) is smaller than the link capacity and communication delay is sufficiently small.

Using the end user algorithm given in [14] we assume that $w(t) = x(t) \cdot U^{'}(x(t))$. Now, suppose that congestion signal generated at the resource, *i.e.*, $p(x(t))$, is returned to the user after a fixed round trip delay T. In the presence of delay the interaction is given by the following delayed differential equation

$$\frac{d}{dt}x(t) = \kappa\left(x(t)U^{'}(x(t)) - x(t-T)p(x(t-T)\right) \tag{8}$$

After normalizing time by T and replacing $t = s \cdot T$, (8) becomes

$$\begin{aligned} \frac{d}{Tds}x(s) &= \kappa\left(x(s)U^{'}(x(s)) - x(s-1)p(x(s-1))\right) \\ \nu\frac{d}{ds}x(s) &= x(s)U^{'}(x(s)) - x(s-1)p(x(s-1)) \end{aligned} \tag{9}$$

where $\nu = \frac{1}{T\kappa}$. It is (9) we are interested in studying from the stability point of view. For $T >> 1$, this equation can be seen as following singular perturbation

$$\nu\frac{d}{dt}x(t) = g(x(t)) - f(x(t-1)) \tag{10}$$

of general nonlinear difference equation with continuous argument given by $g(x(t)) = f(x(t-1))$, $t \geq 0$, where $g(x) = x \cdot U^{'}(x)$ and $f(y) = y \cdot p(y)$ in the context of (9). Under certain natural invertibility conditions on $g(\cdot)$, it leads to a much studied equation [24]

$$x(t) = F(x(t-1)), \;\; t \geq 0 \tag{11}$$

where $F(\cdot) = g^{-1}(f(\cdot))$. For the solution of (11) to be continuous for $t \geq -1$, along with the continuity of F and initial condition $\phi(\cdot)$, a so-called consistency condition $\lim_{t\to -0}\phi(t) = F(\phi(-1))$ is required [10, 24]. It turns out that a great deal about the asymptotic stability of (10) can be learned from the asymptotic behavior of following difference equation, with Z_+ denoting the set of positive integers:

$$x_{n+1} = F(x_n), \;\; n \in Z_+ \tag{12}$$

The concave utility functions and resource price functions assumed in [12, 14] do not satisfy the assumptions in [9], and hence we cannot directly apply their results. However, the general approach used in the paper for establishing the stability can be extended to study the convergence of (8). In the following subsection we first establish general convergence results for one dimensional case described here and a bound on the range of system when the system loses its stability. In Section 5 we illustrate how our results in subsection 3.1 can be applied to study the convergence of the system described in this section to the optimum with utility and resource price functions given in the section.

3.1 Convergence Results

In this subsection we establish the conditions for convergence of the system in (9) regardless of the communication delay T. Consider the following substitution:

$$y(t) = x(t)U^{'}(x(t)) := g(x(t)) \text{ and } f(x(t)) := x(t)p(x(t)).$$

We first make the following assumptions on the functions $g(x)$ and $f(x)$.

Assumption 2 *(i) The function $g(x)$ is strictly decreasing with $-g^{'}(x) > 0$ for all $x > 0$, (ii) the function $f(x)$ is strictly increasing for all $x > 0$, and (iii) both $g(x)$ and $f(x)$ are Lipschitz continuous on $[\varepsilon\ ,\ \infty)$, where ε is an arbitrarily small positive constant.*

This allows us the following change of coordinate:

$$x(t) = g^{-1}(y(t)) \Rightarrow \dot{x}(t) = \frac{\dot{y}(t)}{g^{'}(g^{-1}(y(t)))} \tag{13}$$
$$\nu\dot{y}(t) = g^{'}(g^{-1}(y(t)))(y(t) - f(g^{-1}(y(t-1))))$$

where the inverse $g^{-1}(\cdot)$ exists from Assumption 2. Let $\kappa(y(t)) := -g^{'}(g^{-1}(y(t)))$. Clearly, $\kappa(y(t)) > 0$ under Assumption 2. Using this substitution in (13) we get the following form.

$$\nu\dot{y}(t) = \kappa(y(t))\left(f(g^{-1}(y(t-1))) - y(t)\right) \tag{14}$$

We study (14) and show that there is a close correspondence between invariance and global stability properties of the discrete-time map

$$y_{n+1} = f(g^{-1}(y_n)) := F(y_n) \tag{15}$$

and those of (14). In particular, we will prove that if $y_{n+1} = F(y_n)$ has a fixed point then (14) will have a uniformly constant solution for all possible time delays $T \geq 0$ if the initial function's range is contained in the immediate basin of attraction of this fixed point. The proofs are based on the invariance property of the underlying map $F(\cdot)$ and the monotonicity of function $g(\cdot)$. The map $F(y)$ is strictly decreasing because $g^{-1}(y)$ is strictly decreasing under Assumption 2 and a composition of a strictly increasing function and a strictly decreasing function is a strictly decreasing function.

Assumption 3 *Suppose now that $I \subset \{x : x \geq \varepsilon\}$ is a closed invariant interval under F. In particular let $I = [a,\ b]$ be compact.*

Let $X := C([-1,0],\Re_+)$, and $X_I := \{\phi \in X : \phi(s) \in I\ \ \forall\ s \in [-1,0]\}$. Under this assumption, we have invariance for the solution of (14) for all time $t \geq 0$ and for all $\nu \geq 0$. Since the functions involved in (14) are Lipschitz continuous by assumption, solutions do exist for all $t \geq 0$ and are unique for any initial function $\phi \in X_I$, where I is the assumed closed invariant interval under F. Furthermore, the invariance property of the solutions, which is stated below (Theorem 1), ensures that they stay positive and bounded by the initial set they start in, which is assumed to be invariant under map F. For the proofs of the results in this paper refer to [21].

Theorem 1 *(Invariance) If $\phi \in X_I$, the corresponding solution $y(t) = y(t;\phi)$ satisfies $y(t) \in I$ for all $t \geq 0$. It means that set I is invariant under (14).*

Proof: : Let t_0 be the first time when solution $y(t;\phi)$ leaves I with $\phi \in X_I$. In particular, we can assume that $y(t_0) = b$ and every right hand neighborhood of t_0 will have a $t_1 > t_0$ such that $y(t_1) > b$. Then, we can find a point $t_2, t_0 < t_2 < t_0 + 1$, such that $y(t_2) > b$ and $\dot{y}(t_2) > 0$. Since $y(t_2 - 1) \le b$, we have $\dot{y}(t_2) = \frac{\kappa(y(t_2))(f(g^{-1}(y(t_2-1)))-y(t_2))}{\nu} < 0$ from (14) and Assumption 3 that I is invariant under F, *i.e.*, $f(g^{-1}(y(t_2 - 1))) \le b$. This contradicts with the earlier assumption that $\dot{y}(t_2) > 0$.

Similarly, suppose that $y(t_0) = a$ and the trajectory exits from left end of the interval. Then, every right hand neighborhood of t_0 will have a $t_1 > t_0$ such that $0 < y(t_1) < a$ due to the smoothness of solutions, and we can find $t_2, t_0 < t_2 < t_0 + 1$, such that $0 < y(t_2) < a$ and $\dot{y}(t_2) < 0$. From that $y(t_2 - 1) \ge a$, we have $\dot{y}(t_2) = \frac{\kappa(y(t_2))(f(g^{-1}(y(t_2-1)))-y(t_2))}{\nu} > 0$ from (14) and Assumption 3. This, however, contradicts the assumed negativity of $\dot{y}(t_2) < 0$. Hence, the theorem follows. ∎

Next theorem considers the case when map F has an attracting fixed point y^* with immediate basin of attraction $J_0 : F^n y_0 \to y^*$ for any $y_0 \in J_0$. Let $X_{J_0} = C([-1,0], J_0)$. Then, the following theorem holds.

Theorem 2 *(Stability) For any* $\nu > 0$ *and* $\phi \in X_{J_0}$, $\lim_{t\to\infty} y^{\nu}_{\phi}(t) = y^*$.

Proof: The proof is given in Appendix A ∎

The above theorem tells us that if the initial function lies in X_{J_0}, then the rate $x(t)$ converges to the solution of (6) regardless of the value of T or κ. Hence, it establishes a strong convergence result in the presence of a communication delay.

3.2 Linear Instability

In the previous subsection we have demonstrated that the system in (9) converges with an arbitrary delay under Assumptions 2 - 3 if the initial condition lies in the specified invariant set. In this subsection we study the case where the map defined by (15) loses stability and goes through a period doubling bifurcation with its eigenvalue $\lambda := \frac{dF}{dx}|_{x=x^*} < -1$, where x^* is an unstable fixed point of map F. We describe how the instability of underlying discrete-time map is translated to the instability of delay-differential equation in (14).

Assuming that the map F given by (35) is locally smooth, it is possible to find conditions for linear instability of the fixed point of the map y^* and that of constant function $y(t) = y^*$ for the delay-differential equation in (14). In order for $y(t) = y^*$ to be locally asymptotically stable for all $T \ge 0$, following variational equation should have its zero solution stable.

$$\begin{aligned}
z'(t) &= \kappa(y(t))F'(y(t-T))\Big|_{y=y^*} z(t-T) + \Big(\kappa'(y(t))\,[F(y(t-T)) - y] - \kappa(y(t))\Big)\Big|_{y=y^*} z(t) \\
&= \kappa(y(t))F'(y(t-T))\Big|_{y=y^*} z(t-T) - \kappa(y(t))|_{y=y^*}\, z(t) \qquad \text{(because } F(y^*) = y^* \text{)} \\
&:= Bz(t-T) + Az(t) \qquad (16)
\end{aligned}$$

where $B = \kappa(y^*)F'(y^*)$ and $A = -\kappa(y^*)$. Now in order to determine the stability of $y(t) = y^*$, we can apply the following well known results [20].

(1) Eq. (16) is stable for all $T \ge 0$ only if:

$$A \le 0 \quad \text{and} -A \ge |B| \qquad (17)$$

(2) In case when the above condition in (17) is violated we have partial stability for some values of time delays:

$$-B > |A| \;\text{ and }\; T \le T^* := \frac{cos^{-1}(-\frac{A}{B})}{B^2 - A^2}. \qquad (18)$$

For our case it follows from (16) that $A := -\kappa(y^*)$ is always negative, which holds due to the fact that $\kappa(\cdot)$ is always positive. The second condition $\kappa \geq \kappa|F^{'}|$ is crucial to stability of (14). Clearly, for the case when $F^{'} < -1$ (period doubling condition for the map F) the linear stability condition given by (17) is violated and for a large enough T the constant solution $y(t) = y^*$ will not be stable. Thus, we know that in unstable situation solutions will be more complex than a constant function and will stay within the interval they initially start from due to the invariance results given by Theorem 1.

Theorem 3 *Let $I := [a,b]$ be a closed interval such that $F(I) := [a_1, b_1] \subset I$. Let the initial condition $\phi(t) \in X_I$ be the solution of (14). Now, if the points a_1 and b_1 are fixed points of F, then for all sufficiently small $\varepsilon \geq 0$ there exists a finite $T = T(\phi, \varepsilon, \kappa)$ such that $y(t) \in [a_1 - \varepsilon, b_1 + \varepsilon]$ for all $t \geq T$.*

Proof: The proof follows the same arguments used in the proof of Theorem 1 except that boundary considered here is $b_1 + \varepsilon$ from right. In the interval $[b_1 + \varepsilon, b]$ solution will be strictly decreasing until it reaches the point $b_1 + \varepsilon$. Afterwards from invariance theorem it stays bounded by $b_1 + \varepsilon$ from above. Similar reasoning follows for lower bound. ■

Above theorem essentially gives bounds for the interval which will contain the solution asymptotically. Now suppose that $y(t)$ is a solution of (14) under the instability condition that map $F^{'}(y^*) < -1$. Due to invariance theorem we know that $0 < \liminf_{t\to\infty} y(t) = m \leq \limsup_{t\to\infty} = M < +\infty$.

Now based on the theory developed in [8] we make following observations:

(1) If the solution $y(t)$ is strictly monotone then $m = M = y^*$ because of the boundedness of solutions.

(2) If $m \neq M$ then the solution $y(t)$ is oscillating. In particular, the solutions will have a sequence of maxima at times $\{t_n\}$ and minima at times $\{s_n\}$. Clearly, $y^{'}(t_n) = y^{'}(s_n) = 0$ for all t_n and s_n. This implies from (14) that $y(t_n) = F(y(t_n - 1))$ and $y(s_n) = F(y(s_n - 1))$. This shows interesting discrete time map structures in the solution of delay-differential equation (14).

(3) If $y(t)$ does not converge to y^* then it oscillate around it. This holds due to the fact that image of the interval $[m, M]$ under F contains $[m, M]$. Hence, it will have a fixed point y^*.

From (1) through (3) above we conclude that solutions either converge to the fixed point or oscillate around it.

3.3 Note on Lyapunov Function for Single User with Single Resource

In this subsection, we study the stability of the underlying discrete time model which in turn determines the stability of the delay-differential equation, using Lyapunov theory for discrete time maps [22]. Interestingly, the class of Lyapunov functions that were originally proposed by Kelly [12] for the systems with no delay appears to be useful for underlying discrete time maps of the delay-differential equations.

Consider the following (Lyapunov) functions for discrete time maps of interest:

$$L(x) = -\left[U(x) - c(x) - (U(x^*) - c(x^*))\right] \tag{19}$$

where x^* is the maximizer of $U(x) - c(x)$ and $c(x) = \int_0^x p(y)dy$, which exists for utility and cost functions in [12]. Clearly, function in (19) is strictly positive everywhere except for $x = x^*$, where it is zero, and convex. We also know that x^* is the only fixed point of the discrete time map $F(x^*) = x^*$ given by (15), and is also the unique solution of $L^{'}(x) = 0$ due to extremality condition.

In order to argue that (19) is a Lyapunov function for the underlying discrete time system we need to show that change in $L(\cdot)$ is strictly negative along the discrete time map $F(\cdot)$ starting from any initial condition except for $x = x^*$ or

$$\Delta L(x) := L(F(x)) - L(x) < 0, \ \forall x > 0, \ x \neq x^* \tag{20}$$

If we look at the differential of $\Delta L(x)$, it is clear that $\Delta L'(x)$ has at least one zero at $x = x^*$.

$$\begin{aligned}\Delta L'(x)\big|_{x=x^*} &= L'(F(x))F'(x)\big|_{x=x^*} - L'(x)\big|_{x=x^*} \\ &= L'(x^*)F'(x^*) - L'(x^*) = 0\end{aligned}$$

In order to study the nature of x^* we will need the second derivative of $\Delta L(x)$ evaluated at $x = x^*$.

$$\begin{aligned}\Delta L''(x)\big|_{x=x^*} &= L''(F(x))\left(F'(x)\right)^2\Big|_{x=x^*} + L'(F(x))F''(x)\big|_{x=x^*} - L''(x)\big|_{x=x^*} \\ &= L''(x^*)(\left(F'(x^*)\right)^2 - 1)\end{aligned} \tag{21}$$

because $F(x^*) = x^*$ and $L'(x^*) = 0$ due to the fact that x^* is the unique minimizer of L. Depending on the magnitude of $F'(x^*)$ we have three cases to consider: (i) $|F'(x^*)| < 1$, (ii) $|F'(x^*)| > 1$, and (iii) $|F'(x^*)| = 1$. In the first two cases, we assume that $\Delta L'(x)$ has a unique zero at $x = x^*$ (*i.e.*, $\Delta L'(x)$ is monotonic) and show under this assumption that in the first two cases x^* is either a global maximizer or a global minimizer of $\Delta L(x)$, leading to either global stability or the lack thereof for the discrete time system $F(\cdot)$. One can show numerically that the utility and resource price functions used in Section 5 satisfy this assumption under a mild condition. Also, we conjecture that this assumption holds for a large set of utility and resource price functions. We discuss the aforementioned three cases:

(i) For $|F'(x^*)| < 1$, $\Delta L''(x^*) < 0$ as $L''(x^*) > 0$ due to the convexity of $L(\cdot)$. This implies that x^* is a maximizer of $\Delta L(\cdot)$ and since $\Delta L(x^*) = 0$, $\Delta L(x)$ uniformly negative over positive real axis except for $x = x^*$. This gives us uniform asymptotic stability for the map $F(\cdot)$.

(ii) For $|F'(x^*)| > 1$, $\Delta L''(x^*) > 0$ as $L''(x^*) > 0$ due to the convexity of $L(\cdot)$. This means that x^* is a minimizer of $\Delta L(\cdot)$, and since $\Delta L(x^*) = 0$, its uniformly positive over positive real axis except for $x = x^*$. This shows instability for the map $F(\cdot)$.

(iii) In the case where $|F'(x^*)| = 1$, x^* becomes neutrally stable. This is essentially the bifurcation point from stable to unstable behavior. Hence, parameters ρ at which $|F'(x^*(\rho), \rho)| = 1$ are critical from stability point of view and give the region of stable operation.

Hence, we see that $|F'(x^*)| < 1$ is crucial to the stability of the whole system. In particular the case when $F'(\cdot) \leq -1$ is more interesting as it indicates the birth of oscillations through period doubling bifurcation.

Next we describe a sufficient condition that guarantees the existence of a unique global attractor under the above Lyapunov function. In order for x^* to be the global maximum it suffices to show that $\Delta L'(x) > 0$ for all $0 < x < x^*$ and $\Delta L'(x) < 0$ for all $x > x^*$. To this end we rewrite $\Delta L'(x)$ as follows:

$$\begin{aligned}\Delta L'(x) &= L'(F(x))F'(x) - L'(x) \\ &= \frac{-F'(x)}{F(x)}\left(g(F(x)) - f(F(x))\right) + \frac{1}{x}\left(g(x) - f(x)\right) \\ &= \frac{-1}{x}\left[\frac{-xF'(x)}{F(x)}\left(f(F(x)) - f(x)\right) - \left(g(x) - f(x)\right)\right]\end{aligned} \tag{22}$$

where the last equality follows from that $F(x) = g^{-1}(f(x))$.

Assumption 4 *Suppose that* $0 < \frac{-xF'(x)}{F(x)} < 1$ *for all* $x > 0$.

Under the above assumption one can see from (22) that a sufficient condition to have $\Delta L'(x) > 0$ for all $0 < x < x^*$ is

$$f(F(x)) - f(x) < g(x) - f(x)$$

since both $f(F(x) - f(x)$ and $g(x) - f(x)$ are positive. This is equivalent to

$$g^{-1}(f(F(x))) > g^{-1}(g(x)) = x$$

because $g(\cdot)$ is assumed to be monotonically decreasing. From the definition of the mapping $F(\cdot)$ one can see that $g^{-1}(f(F(x))) = F(F(x))$. Therefore, a sufficient condition is that $F^2(x) > x$ for all $0 < x < x^*$. Similarly, one can show that a sufficient condition for the other case, *i.e.,* $\Delta L'(x) < 0$ for all $x > x^*$, is that $F^2(x) < x$.

Proposition 1 *Suppose that Assumption 4 holds. Then, a sufficient condition for the existence of a unique global stable equilibrium under the above Lyapunov function is that $F^2(x) < x$ for all $x > x^*$ and $F^2(x) > x$ for all $0 < x < x^*$.*

One can easily show that Assumption 4 and the above conditions in Proposition 1 hold with the utility and price functions used in Section 5 under a stability condition.

4 Stability Conditions: Multiple-Flow, Multiple-Resource

Let I_l be the set of users traversing resource $l \in L$, *i.e.,* $I_l = \{i \in L \mid l \in r_i\}$. We assume that $p_l(\cdot), l \in L$, are strictly increasing and continuously differentiable. Suppose that the feedback information from the resources to the end users is delayed by $T > 0$. Following Kelly's rate control formulation [12], the rate of i-th user now evolves according to the following delay-differential equation

$$\frac{d}{dt}x_i(t) = \kappa_i(x_i(t)U'(x_i(t)) - x_i(t-T)(\sum_{l \in r_i} \mu_l(t-T))) \tag{23}$$

where $\mu_l(t-T) = p_l(\sum_{j \in I_l} x_j(t-T))$. After normalizing time in (23) by T and using the substitution $y_i = x_i U'(x_i) := g_i(x_i)$, the dynamic equation for i-th user can be rewritten as

$$\begin{aligned} x_i(t) = g_i^{-1}(y_i(t)),\ \dot{x}_i(t) &= \frac{\dot{y}_i(t)}{g_i'(g_i^{-1}(y_i(t)))} \\ \nu \dot{y}_i(t) &= \kappa_i g_i'(g_i^{-1}(y_i(t)))(y_i(t) - f_i(\overline{g}^{-1}(\overline{y}(t-1)))) \end{aligned}$$

where $\nu = \frac{1}{T}$, $\overline{y}(t-1) = (y_1(t-1), \cdots, y_N(t-1))$, and

$$f_i(\overline{g}^{-1}(\overline{y}(t-1))) = g_i^{-1}(y_i(t-1)) \left(\sum_{l \in r_i} p_l(\sum_{j \in I_l} g_j^{-1}(y_j(t-1))) \right) .$$

We can write the above in the following matrix form:

$$\nu \dot{\overline{y}}(t) = \kappa(\overline{y}(t))\,(F(\overline{y}(t-1)) - \overline{y}(t)) \tag{24}$$

where $\kappa(\cdot)$ is state dependent diagonal gain matrix with $\kappa_{ii} = -\kappa_i g_i'(g_i^{-1}(y_i(t)))$. Clearly, this decomposition is possible due to the fact that the utility of a user is a function only of its own rate and does not depend on those of other users. The map $F(\cdot)$ given by (25) is a multidimensional one step nonlinear map which is crucial for understanding the stability of the system. We note that this system of differential equations has a natural underlying difference map structure similarly as in the single-flow, single-resource case.

$$\bar{y}_{n+1} = F(\bar{y}_n),\; n \in Z_+, y_n \in \Re^n_+ \;\text{ and }\; F_i(\bar{y}) = f_i(g_1^{-1}(y_1), ..., g_N^{-1}(y_N)) \tag{25}$$

The importance of this multidimensional one step map will be evident when it will be shown that global stability of this map is a sufficient condition for the global stability of the delay-differential system given by (24).

4.1 Convergence Results

For the case of multiple users with heterogenous utility functions we will need to prove the global convergence in a multidimensional space. Our approach builds upon the approach used by Verriest and Ivanov [26]. The basic idea behind this approach is to use invariance and continuity properties of the underlying map for the differential equation and find a sequence of bounds using convex sets, which converge to the singleton with the solution to (5). Hence, the convergence is derived from the underlying map, which provides the bounds for the trajectories of the delay-differential system. Following this plan, we will first prove the invariance of system given by (24) when the underlying map given by (25) has a convex invariance set that is a product space. The assumption of existence of a convex invariance set is natural in rate control problem as will be illustrated in the next section using a family of utility and resource price functions.

Before we present the convergence results, we state an assumption that we make on $g(\cdot)$ and $f_i(\cdot)$, $i \in I$.

Assumption 5 *(i) The function $g_i(x_i)$ is strictly decreasing with $-g_i'(x_i) > 0$ for all $x_i > 0$, (ii) the function $f_i(\underline{x})$ is strictly increasing in each component for all $\underline{x} > 0$, and (iii) both $g(x_i)$ and $f_i(\underline{x})$ are Lipschitz continuous on $\Re_+$ and $\Re^N_+$, respectively.*

It can be seen that under this assumption $\kappa(\cdot)$ is strictly positive definite matrix, which turns out to be an important property to prove the convergence results for the system given by (24).

Our first result states that the set $C([-1,0],D)$ is invariant under the action generated by (24), provided that D is closed, convex and invariant under F in (25).

Theorem 4 *(Invariance) Suppose that D is a closed, convex, invariant domain under $F(\cdot)$ given as a product space, i.e., $D = \prod_{i=1}^N proj_i(D)$ where $proj_i(\cdot)$ denotes the i-th component projection operator. Then, for any initial function $\phi \in C([-1,0],D) := X_D$ the resulting $\bar{y}^\nu(t)$ from (24) belongs to the domain D for all $t \geq 0$ and $\nu \geq 0$, where the superscript ν is used to denote the dependence on $\nu = \frac{1}{T}$.*

Proof: The proof is given in Appendix B. ■

Now, using the invariance property, we turn to the asymptotic property of delay-differential equation under the natural assumption that the underlying map is stable. The lack of convexity of an image $F(D)$ of a convex set D forbids the direct application of techniques developed by Verriest and Ivanov [26]. Instead our approach is to construct a series of convex coverings of image $F(D_n)$ and look at their asymptotic behavior. In particular, we construct a sequence of product spaces which seem to be the most reasonable choice for the networking problems. These product spaces are obviously convex and if we can prove that they contain their images, then the invariance follows. Also, due to the monotonicity property of map $F(\cdot)$, which follows from Assumption 5, the coordinates of these product spaces and their images can be computed explicitly. To recapitulate we want to construct a series of convex closed domains $\{D_n\}$ such that under certain stability conditions that $F(D_n) \subset D_{n+1} \subset int(D_n)$, where $int(D_n)$ denotes the interior of D_n, and $\{y^*\} = \cap_{n\geq 0} D_i$, all the solutions of the map given by (25) will converge to y^* asymptotically.

Assumption 6 *Multidimensional map $F : \Re^N \to \Re^N$ has an arbitrary fixed point y^* and there exists an open convex neighborhood $int(D_0)$, which is an open product space. Also, assume that there is a sequence of closed convex domains $D_n, n \geq 0$, that are product spaces, such that $F(D_n) \subset D_{n+1} \subset int(D_n)$ and $\{y^*\} = \cap_{n\geq 0} D_n$.*

Let $Y_{D_0} = C([-1,0], int(D_0))$ be a subset of initial functions and $\bar{y}^\nu_\phi$ a solution of (24) constructed through $\phi \in Y_{D_0}$.

Theorem 5 *(Stability) All solutions starting with initial functions $\phi \in Y_{D_0}$ converge to y^* for all $\nu > 0$.*

Proof: The proof is given in Appendix C. ■

Theorem 5 establishes that the attracting fixed point is stable in set D_0. The basic tool these theorems give us is to look at a delay-differential equation as a discrete map which is much more convenient to study and intuitive from implementation point of view due to the discrete nature of computation elements. The study of underlying maps give us more insight than the equations themselves as shown in the next section.

4.2 Note on Lyapunov Function for a General Network

In this subsection, we study the stability of the underlying discrete time model using Lyapunov theory for discrete time maps [22] similarly as in Section 4.2 for one flow with single resource case. Consider the following (Lyapunov) function for a discrete time map of interest:

$$L(x) = -\left[\sum_i U_i(x_i) - \sum_l \int_0^{\sum_{i:l\in r_i} x_i} p_l(y)dy. - \sum_i U_i(x_i^*) - \sum_l \int_0^{\sum_{i:l\in r_i} x_i^*} p_l(y)dy.\right] \tag{26}$$

where x^* is the unique maximizer of $\sum_i U_i(x_i) - \sum_l \int_0^{\sum_{i:l\in r_i} x_i} p_l(y)dy$, which exists for utility and cost functions used in [12]. Clearly, $L(x)$ is convex and strictly positive everywhere except for at $x = x^*$, where it is zero.

In order to show the stability of the underlying discrete time system we need to show that the change in $L(\cdot)$

$$\Delta L(x) := L(F(x)) - L(x)$$

is strictly negative along the discrete time map $F(\cdot)$, starting from any initial condition except for at $x = x^*$. In other words,

$$\Delta L(x) < 0 \text{ for all } x > 0,\ x \neq x^*.$$

Before we analyze the nature of change in our Lyapunov function, let us define the following. This notation will help us carry over the intuition of one dimensional case in Section 4.2.

$$\begin{aligned} L'(\cdot) &= \nabla L(\cdot) \text{ or Gradient of } L(\cdot) \\ L''(\cdot) &= \text{Hessian of } L(\cdot) \\ F'(\cdot) &= \text{Jacobian of multidimensional map } F(\cdot) \\ F'(\cdot)^T &= \text{Transpose of } F'(\cdot) \end{aligned}$$

If we look at the gradient of $\Delta L(x)$, it is clear that $\Delta L'(x)$ has at least one zero at $x = x^*$.

$$\begin{aligned} \Delta L'(x)\big|_{x=x^*} &= F'(x)^T L'(F(x))\big|_{x=x^*} - L'(x)\big|_{x=x^*} \\ &= F'(x^*)^T L'(x^*) - L'(x^*) = 0 \end{aligned}$$

In order to study the nature of the extremal $\Delta L(x^*)$ we need the second-order derivative or Hessian of $\Delta L(x)$ evaluated at $x = x^*$.

$$\begin{aligned}\Delta L''(x)\big|_{x=x^*} &= F'(x)^T L''(F(x))F'(x)\big|_{x=x^*} - L''(x)\big|_{x=x^*} + \Lambda(x)|_{x=x^*} \\ &= F'(x^*)^T L''(x^*)F'(x^*) - L''(x^*) \end{aligned} \tag{27}$$

because $F(x^*) = x^*$. The last term $\Lambda(x^*)$ goes away because it consists of the second-order derivative of map $F(\cdot)$ multiplied by $L'(x^*)(= 0)$. The stability of this system is dependent on the magnitude of eigenvalues of $F'^T(x^*)F'(x^*)$. Clearly, in order for x^* to be a (local) maximizer of $\Delta L(\cdot)$ (27) must be negative definite.

A simple example of a sufficient condition for (27) to be negative definite is

$$\lambda_{max} F'(x^*)^T F'(x^*) < \lambda_{min} I \tag{28}$$

where λ_{min} and λ_{max} are minimum and maximum eigenvalues of $L''(x^*)$, respectively. Note that $L''(x^*) > 0$ as $L(\cdot)$ is a convex function, and hence both λ_{min} and λ_{max} are positive. Here I denotes the $N \times N$ identity matrix. The condition in (28) implies that the singular value of matrix $F'(x)$ should be less than $\frac{\lambda_{min}}{\lambda_{max}}$.

5 Application to Rate Control Example

In this section we apply the results in the previous sections to investigate the stability of the rate control problem described in Section 2 with a single resource. We consider the following class of users' utility functions:

$$U_a(x) = -\frac{1}{a}\frac{1}{x^a}\ , \ a > 0. \tag{29}$$

In particular, $a = 1$ has been found useful for modeling the utility function of Transmission Control Protocol (TCP) algorithms [15]. This class of utility functions in (29) has been used extensively in engineering literature [1, 12, 15]. We say that a user u_1 with utility function $U_{a_1}(x)$ is greedier than another user u_2 with utility function $U_{a_2}(x)$ if $a_2 > a_1$. One can interpret the notion of greed here using the notion of elasticity of demand [25]. With the utility functions of the form in (29) one can easily show that the elasticity of the demand decreases with a as follows. Given a price p, the optimal rate $x^*(p)$ of the user that maximizes the net utility $U_a(x) - \int_0^x p(y)dy$ is given by $p^{-\frac{1}{1+a}}$. The price elasticity of the demand, which measures how responsive the demand is to a change in price, is defined to be the percent change in demand divided by the percent change in price [25]. In our case the price elasticity of demand is given by

$$\begin{aligned}\frac{p}{x^*(p)}\frac{dx^*(p)}{dp} &= \frac{p}{p^{-\frac{1}{1+a}}} \cdot \frac{-1}{1+a} p^{-\frac{1}{1+a}-1} \\ &= \frac{-1}{1+a}\ . \end{aligned} \tag{30}$$

Therefore, one can see that the price elasticity of demand decreases with a,[3] *i.e.,* the larger a is, the less responsive the demand is.

The class of resource price functions that we consider is of the form:

$$p(y) = \left(\frac{y}{C}\right)^b, \text{ where } b > 0 \tag{31}$$

This kind of marking function arises if the resource is modeled as $M/M/1$ queue with a service rate C packet per unit time and a packet receives a mark with a congestion indication signal if it arrives at the queue to find at least b packets in the queue. One can easily verify that these utility functions and resource price functions satisfy the assumption in Sections 3 and 4.

[3]When comparing the price elasticity, typically the absolute value of (30) is used.

5.1 Homogeneous Users

Using the utility function and resource price function of (29) and (31), respectively, (8) can be rewritten as

$$\frac{d}{dt}x(t) = \kappa \left(\frac{1}{x(t)^a} - x(t-T) \left(\frac{x(t-T)}{C} \right)^b \right) . \tag{32}$$

The theorems in subsection 3.1 can be directly applied to study the dynamical behavior of (32) which is essentially described by the underlying discrete time difference equation

$$y_{n+1} = F(y_n) \tag{33}$$
$$\frac{1}{x_{n+1}^a} = x_n \left(\frac{x_n}{C} \right)^b , \quad x_n > 0 \tag{34}$$
$$x_{n+1} = \left(\frac{C^b}{x_n{}^{b+1}} \right)^{\frac{1}{a}} \tag{35}$$

Consider the dynamical behavior of map given by (35). It has a fixed point

$$x^* = C^{\frac{b}{a+b+1}} , \tag{36}$$

and the market equilibrium price is given by $p^* = C^{-\frac{b(1+a)}{1+a+b}}$. The market equilibrium price can be obtained from that $x^* = p^{*-\frac{1}{1+a}}$. This expression of equilibrium flow shows that x^* increases with decreasing a. This is the reason that we characterize the user with smaller a greedier. The eigenvalue at this fixed point, which is interestingly independent of the fixed point, is

$$\lambda(x^*) = -\frac{b+1}{a} . \tag{37}$$

Suppose that $a > b+1$. Then, the fixed point x^* is locally attracting. In fact, x^* can be shown to be *globally* stable as follows. According to the Sharkovsky cycle coexistence ordering [23] the most general condition for the fixed point x^* to be globally attracting is that the second iteration F^2 of the map F does not have a fixed point in relevant state interval other than x^*, and x^* is locally attracting. This in turn implies the global delay independent stability of (32). It is interesting to note that when the utility function of user is given by $U_1(x) = -\frac{1}{x}$ as has been suggested for TCP algorithms [15], the delay independent stability of the system cannot be ensured by a price function of the form in (31).

Our results have the following interpretation. If the functions $x \cdot U'(x)$ and $x \cdot p(x)$ have an intersecting point that is a stable fixed point, then the communication delays are irrelevant for system stability, and user rate and resource price converge to the system optimum. Furthermore, our results tell us that the stability of system depends critically on the user utility functions, more specifically on the parameter a, for a given price function. This can be seen from the eigenvalue $\lambda(x^*) = -\frac{b+1}{a}$. Larger values of b mean that the slope of the price function is steeper, which in turn implies that the price varies more widely in response to a change in rate x. Hence, in order to maintain the stability of system, user demand should be less elastic, *i.e.*, the response of user to a change in price should be less dramatic. Thus, this presents a fundamental trade-off between the elasticity of user demand and responsiveness of price function. In other words, in order to keep the stability of system, if one wants to increase the responsiveness of one, then the responsiveness of the other must be sacrificed. This trade-off can also be seen from the definition of the discrete time map $F(y) = f(g^{-1}(y))$. The derivative of $F(\cdot)$ evaluated at the equilibrium is given by

$$\frac{d}{dy}F(y)\bigg|_{y^*} = \frac{df}{dg^{-1}(y)}\bigg|_{y^*} \frac{dg^{-1}(y)}{dy}\bigg|_{y^*} \tag{38}$$

$$
\begin{aligned}
&= (b+1)\,\frac{\left(g^{-1}(y)\right)^b}{C^b}\cdot\left(-\frac{1}{a}\right)y^{-\frac{1+a}{a}}\Bigg|_{y^*=C^{\frac{-ab}{a+b+1}}} \\
&= -\frac{b+1}{a}
\end{aligned}
$$

by the chain rule, where $g^{-1}(y) = y^{-\frac{1}{a}}$. Here the first term in the right hand side gives the slope of the resource price function as a function of the flow rate, while the second term is a function only of user's utility function, namely $\frac{d}{dy}g^{-1}(y) = -\frac{1}{a}y^{-\frac{1+a}{a}}$. Therefore, (38) clearly describes the trade-off between user's price elasticity of demand and the responsiveness of resource price function for stability.

The above results have the following practical implications. Characterizing the exact stability conditions of the system with a given choice of utility and price functions is not easy. In addition, the round-trip delays of connections tend to vary widely. Therefore, one approach to designing a stable system is to select a pair of user utility and price functions in such a way the communication delay does not affect the stability of the system. This is, however, not to say that the dynamics of the system do not depend on the delay.

Our results also provide us with the following design guideline for the AQM mechanism and end user algorithms for efficient use of network resources. Note that from (36) the fixed point x^* is strictly increasing in b and is strictly decreasing in a. Therefore, in order to increase the utilization at the fixed point, we should increase the ratio $\frac{b}{a}$. However, this ratio cannot be increased arbitrarily without losing the stability from (37). Therefore, in order to achieve high utilization of the resource and maintain the stability of the system, the parameter b should be selected as large as possible and the parameter a should be selected just large enough so that $|\lambda(x^*)|$ is smaller than one. However, having the eigenvalue close to -1 comes at the price of a larger settling time. In order to reduce the settling time, the ratio of $\frac{b}{a}$ should be lowered. Therefore, the selection of parameters a and b presents a fundamental trade-off between stability, settling time, and utilization of the system. This is numerically demonstrated in the following section.

We now study what effects the load of the system, *i.e.,* the number of users in the system, has on the stability of the system. Since the load on a resource is beyond the control of a network manager, ideally the stability of the system should not depend on the load. Suppose that there are $N, N \geq 1$, homogeneous users in the system. Since users are assumed to be homogeneous, we denote the rate of a user by $x(t)$. We assume that utility function of the users is of the form in (29) and the price function used at the resource is that of (31). Then, the end user algorithm is given by

$$
\begin{aligned}
\dot{x}^{(N)}(t) &= \kappa\left(x^{(N)}(t)U_a'(t) - x^{(N)}(t-T)\cdot p(N\cdot x^{(N)}(t-T)\right) \\
&= \kappa\left(\frac{1}{x^{(N)}(t)^a} - x^{(N)}(t-T)\left(\frac{N\cdot x^{(N)}(t-T)}{C}\right)^b\right),
\end{aligned}
$$

where a superscript (N) is used to denote the dependence on N. Following similar steps as in the single flow case above, the discrete time difference equation corresponding to (33) - (35) of single flow case yields

$$
x_{n+1}^{(N)} = \left(\frac{(C/N)^b}{x_n^{(N)\,b+1}}\right)^{\frac{1}{a}}. \tag{39}
$$

Then, from (39) the fixed point $x^{(N)*}$ is $\left(\frac{C}{N}\right)^{\frac{b}{a+b+1}}$,[4] and the eigenvalue is given by $\lambda^{(N)}(x^{(N)*}) = -\frac{b+1}{a}$ and is independent of N. Therefore, the stability of the system does not depend on the number of users in the system. This can also be explained using the price elasticity of demand. Since, given a utility function of

[4]Here we assume that the fixed point is smaller than C.

the form in (29) for some $a > 0$, the price elasticity of the demand is constant for all $x > 0$ from (30), one would expect the stability of the system to be independent of the operating point, *i.e.*, the fixed point, and capacity, but only on the choices of the utility and price functions that determine the responsiveness of the users and resource, respectively.

Clearly, the network designer can rescale the price function by a scalar, *i.e.*,

$$p(y) = \gamma \cdot \left(\frac{y}{C}\right)^b \ , \tag{40}$$

where $\gamma > 0$. When the price function is of the form in (40), the fixed point of the system with N flows is given by $x^* = \gamma^{-\frac{1}{1+a+b}} \left(\frac{C}{N}\right)^{\frac{b}{1+a+b}}$. Furthermore, the value of γ does not change the eigenvalue at the fixed point, *i.e.*, the stability condition does not depend on γ. Hence, if the number of flows traversing the resource is known, then the resource can select an appropriate value of γ so that the fixed point of the system achieves high utilization. However, smaller values of γ reduces the responsiveness of the price function.

5.1.1 Nature of period doubling bifurcation

In this subsubsection we investigate the nature of emerging period doubling bifurcation as different parameters in utility and cost functions are varied. Nonlinear stability analysis of period doubling bifurcation is important from the point of view of ensuring graceful degradation and avoiding a catastrophic collapse in case of loss of stability. To this end we need to evaluate

$$E(F) = \frac{1}{3}F^{(3)}(x) + \frac{1}{2}\left(F''(x)\right)^2 \tag{41}$$

at $x = x^*$ and $a = b+1$, where $F^{(3)}$ is the third-order derivative [6]. In order for the period doubling bifurcation not to be subcritical, $E(F)$ evaluated at $x = x^*$ and $a = b+1$ needs to be non-negative. Since $F'(x^*) = -1$ when $a = b+1$, this is equivalent to showing that the Schwarzian derivative is non-positive, *i.e.*,

$$S(F) = \frac{F^{(3)}}{F'} - \frac{3}{2}\left(\frac{F''}{F'}\right)^2 \leq 0 \ . \tag{42}$$

Computing higher-order derivatives for the map $F(\cdot)$ gives

$$F'' = \frac{b+1}{a}\left(\frac{b+1}{a}+1\right)\frac{C^{b/a}}{x^{\frac{b+1}{a}+2}} \tag{43}$$

$$F^{(3)}(x^*) = = -\frac{b+1}{a}\left(\frac{b+1}{a}+1\right)\left(\frac{b+1}{a}+2\right)\frac{C^{b/a}}{x^{\frac{b+1}{a}+3}} \tag{44}$$

Substituting (43) and (44) in (42) yields

$$\begin{aligned} S(F) &= \frac{(\frac{b+1}{a}+1)(\frac{b+1}{a}+2)}{x^2} - \frac{3}{2}\left(\frac{\frac{b+1}{a}+1}{x}\right)^2 \\ &= \frac{a^2-(b+1)^2}{2a^2x^2} \\ &= 0 \ . \end{aligned}$$

Therefore, one can see that (41) is non-negative.

This implies that under the instability condition, *i.e.*, $\frac{b+1}{a} > 1$, a period doubling bifurcation is guaranteed to be not subcritical when it occurs. Hence, when the system loses stability, the magnitude of oscillation in rate will gradually increase with parameters and the system will not experience a sudden appearance of oscillation with a large magnitude.

5.2 Heterogeneous Users

With the above utility and resource price functions, the underlying discrete time map from (25) is given by

$$\frac{1}{x_{i,n+1}^{a_i}} = x_{i,n}\left(\sum_{l\in r_i}\left(\frac{\sum_{j\in I_l} x_{j,n}}{C_l}\right)^{b_l}\right)$$

$$\Rightarrow x_{i,n+1} = x_{i,n}^{-\frac{1}{a_i}}\left(\sum_{l\in r_i}\left(\frac{\sum_{j\in I_l} x_{j,n}}{C_l}\right)^{b_l}\right)^{-\frac{1}{a_i}} \tag{45}$$

Note that $x_{i,n+1}$ is strictly decreasing in each of $x_{j,n}, j \in \mathcal{I}$.

We define $b_{max}^i = \max_{l\in r_i} b_l$ and $C^i = \min_{l\in r_i} C_l$ for all $i \in \mathcal{I}$, and assume that users are ordered by increasing a_i, *i.e.,* $a_1 \geq a_2 \geq \cdots \geq a_N$. Let x^* be the unique solution of the optimization problem. We assume that $A^T x^* < C$. A sufficient condition for this is that $C_l > |I_l|$. Let $\sigma = -\max_{i\in\mathcal{I}} \frac{b_{max}^i+1}{a_i}$. Suppose that $D_0 = \prod_{i=1}^N D_0^i$, where

$$D_0^i = [\overline{\beta} x_i^* \ , \ \overline{\alpha} x_i^*] \ ,$$

$\overline{\alpha}$ is some finite constant larger one, and $\overline{\beta}$ is a positive constant such that

$$\hat{F}(\overline{\beta}x^*) < \overline{\alpha}x^* \text{ and } \overline{\beta}x^* < \hat{F}(\overline{\alpha}x^*) \ . \tag{46}$$

Lemma 1 *Suppose that $a_i > b_{max}^i + 1$ for all $i \in \mathcal{I}$. Then, any $\overline{\beta}$ such that $\overline{\alpha}^{1/\sigma} < \overline{\beta} < \overline{\alpha}^{\sigma}$ satisfies (46).*

Proof: The proof is given Appendix D ■

Now, for $k = 1, 2, \cdots$, we define

$$D_k = \begin{cases} \prod_{i=1}^N [\overline{\alpha}^{\sigma^k} x_i^* \ , \ \overline{\beta}^{\sigma^k} x_i^*] \ , & k \text{ odd} \\ \prod_{i=1}^N [\overline{\beta}^{\sigma^k} x_i^* \ , \ \overline{\alpha}^{\sigma^k} x_i^*] \ , & k \text{ even} \end{cases} \tag{47}$$

Lemma 2 *Suppose that $a_i > b_{max}^i + 1$ for all $k = 1, 2, \cdots$. Then, $\hat{F}(D_{k-1}) \subset D_k \subset int(D_{k-1})$, where $int(D_{k-1})$ is the interior of D_{k-1}, and $\cap_{k=0}^{\infty} D_k = \{x^*\}$.*

Proof: The proof is given in Appendix E. ■

Theorem 5.1 *Suppose that $a_i > b_{max}^i + 1$ for all $i \in \mathcal{I}$. If the initial functions ϕ lie in $C([0, -1], int(D_0))$, then $x(t)$ converges to x^* asymptotically for all $T > 0$ and $\kappa_i > 0$.*

Proof: The theorem follows from Lemma 2 and Theorem 5. ■

Now note that as $\overline{\alpha}$ increases, $\hat{F}(\overline{\alpha})$ goes to $\underline{0}$. Hence, since the rates of the users are in practice constrained by the link capacities, we can see that starting from any arbitrary rate vector satisfying the capacity constraint, the rates converge to x^* asymptotically from the above results.

6 Effects of Non-Responsive Traffic

Some of applications in the Internet, such as real-time streaming, cannot react to congestion as fast as responsive flows and hence adopt a different transport layer protocol that adapts to congestion state much slower. Since their reaction times are much larger than those of responsive flows, for modeling purposes in the time scale of interest they can be modeled as non-responsive flows, whose rates do not vary. In this section, we analyze the effect of the presence of these non-responsive flows. As non-responsive traffic has no dynamics and contributes to aggregate rate presented to the resource, the dynamics of the model given by (34) and (35) gets modified. Here we only consider a single flow case. However, similar results can be shown for multiple heterogeneous user case.

$$
\begin{aligned}
\frac{1}{x_{n+1}^{a}} &= x_n\left(\frac{x_n+q}{C}\right)^{b},\ x_n>0 \\
x_{n+1} &= \left(\frac{C^{b}}{(x_n+q)^{b}x_n}\right)^{\frac{1}{a}}, \qquad (48)
\end{aligned}
$$

where q is the aggregate load from non-responsive flows.

The unique fixed point for this system will be given as the solution of following equation:

$$
\begin{aligned}
& x^{a+1}(x+q)^{b} = C^{b} \\
& \Rightarrow x^{\frac{a+1}{b}}(x+q) = C \\
& \Rightarrow x+q = \frac{C}{x^{\frac{a+1}{b}}} \qquad (49)
\end{aligned}
$$

It is clear from (49) that as the amount of non-responsive traffic q increases, the solution $x^*(q)$ of (49) decreases which is intuitive. Now we will analyze the eigenvalue of (48) to look into the effects of non-responsive traffic on stability. Computing the eigenvalue $\lambda^q(x^*(q))$ gives the following:

$$
\lambda^q(x^*(q)) = -\frac{C^{\frac{b}{a}}}{a}\left[\frac{1}{x^{\frac{a+1}{a}}(x+q)^{\frac{b}{a}}} + \frac{b}{x^{\frac{1}{a}}(x+q)^{\frac{a+b}{a}}}\right]\Bigg|_{x=x^*(q)} \qquad (50)
$$

Substituting (49) for $(x+q)$ in (50) yields

$$
\begin{aligned}
\lambda^q(x^*(q)) &= -\frac{C^{\frac{b}{a}}}{a}\left[\frac{1}{C^{\frac{b}{a}}} + \frac{bx^{\frac{a+b+1}{b}}}{C^{\frac{a+b}{a}}}\right]\Bigg|_{x=x^*(q)} \\
&= -\frac{1}{a}\left[1+\frac{bx^{\frac{a+b+1}{b}}}{C}\right]\Bigg|_{x=x^*(q)} \qquad (51)
\end{aligned}
$$

This expression can be verified by using the expression of fixed point in the absence of non-responsive traffic from (36) which yields the same expression as the eigenvalue in the responsive traffic only case given by (37). From the earlier observation that the fixed point decreases with the load of non-responsive traffic q, it can be easily seen that second term in the eigenvalue will decrease with q, and hence eigenvalue will become smaller in magnitude rendering the system stable. This demonstrates that the stability of the system improves with q. A similar observation has been made in the context of TCP-RED [17].

Although (51) tells us qualitatively that the stability of the system improves with increasing q, it does not give us quantitative answers as to how the trade-off between the stability and responsiveness of the user or resource price is changed. Here we study the improvement in stability by computing the supremum of the values of b that result in a stable system for a fixed value of a or the infimum of the values of a that lead to a stable system for a fixed value of b.

From (48) one can see that if $x_n > x^*(q)$, then $x_{n+1} < x^*(q)$ and similarly $x_n < x^*(q)$ implies that $x_{n+1} > x^*(q)$. Let $x_n = \beta \cdot x^*$ and

$$\tilde{\beta} = \frac{x_n + q}{x^* + q} \ . \tag{52}$$

We denote $x^*(q)$ simply by x^* when there is no confusion. Then, we can rewrite (48)

$$\begin{aligned} x_{n+1} &= \left(\frac{C}{\tilde{\beta}(x^* + q)} \right)^{\frac{b}{a}} (\beta x^*)^{-\frac{1}{a}} \\ &= \left(\frac{C}{x^* + q} \right)^{\frac{b}{a}} \tilde{\beta}^{-\frac{b}{a}} \beta^{-\frac{1}{a}} x^{*-\frac{1}{a}} \\ &= x^* \tilde{\beta}^{-\frac{b}{a}} \beta^{-\frac{1}{a}} \\ &= x^* \left(\tilde{\beta}^b \beta \right)^{-\frac{1}{a}} \end{aligned} \tag{53}$$

For stability it suffices to have

$$\left(\tilde{\beta}^b \beta \right)^{\frac{1}{a}} < \beta \ \text{ if } \beta > 1 \ \text{ and } \ \left(\tilde{\beta}^b \beta \right)^{\frac{1}{a}} > \beta \ \text{ if } \beta < 1$$

Since $\tilde{\beta}^b \beta = 1 = \beta$ when $\beta = 1$, a locally sufficient condition for $|\lambda^q(x^*)| < 1$ is

$$\left. \frac{d\gamma(\beta)}{d\beta} \right|_{\beta=1} < 0 \ , \tag{54}$$

where $\gamma(\beta) = \tilde{\beta}^b \beta^{1-a}$. By substituting (52)

$$\begin{aligned} \frac{d}{d\beta}\gamma(\beta) &= \frac{d}{d\beta} \frac{(\beta x^* + q)^b \beta^{1-a}}{(x^* + q)^b} \\ &= \frac{\beta^{-a} (\beta x^* + q)^{b-1} (b\beta x^* + (1-a)(\beta x^* + q))}{(x^* + q)^b} \end{aligned} \tag{55}$$

Thus, in order to satisfy (54) we need

$$bx^* + (1-a)(x^* + q) < 0$$

or equivalently

$$b + 1 < a + \frac{(a-1)q}{x^*} \ . \tag{56}$$

When $q = 0$, *i.e.*, there is no non-responsive traffic, (56) yields $b + 1 < a$, which is the necessary and sufficient condition for global stability given in subsection 5.1. Since $x^*(q)$ is decreasing in q, the ratio $\frac{q}{x^*(q)}$ increases with q. Hence, it is clear that the value of a that satisfies $b + 1 - a = \frac{(a-1)q}{x^*(q)}$, which is given by $\frac{b+1+q/x^*(q)}{1+q/x^*(q)}$, is decreasing in q. Thus, (56) hints at the trade-off between the responsiveness of the user and resource price and stability similarly as in the case only with elastic traffic discussed in subsection 5.1.

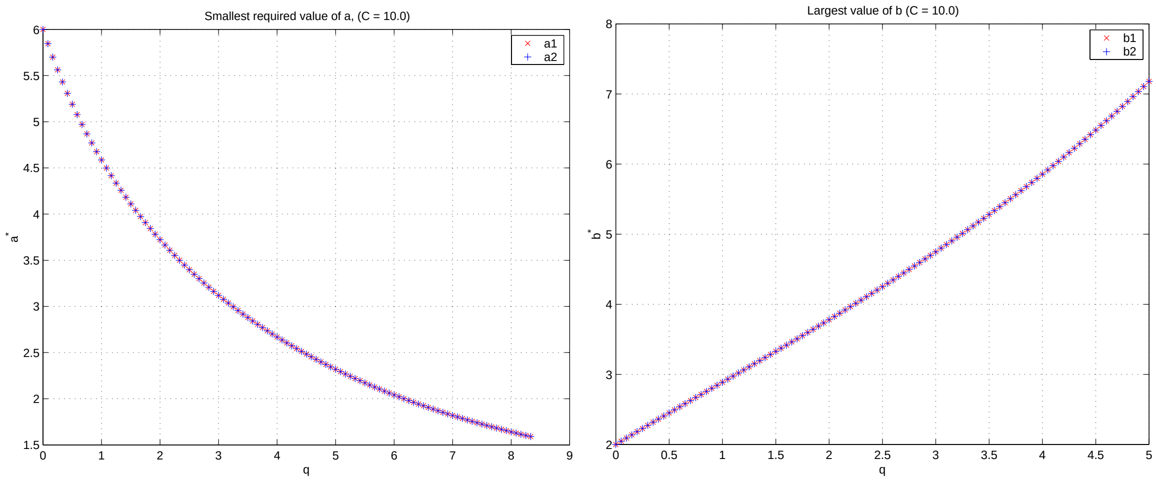


Figure 1: Plot of (a) $\inf\{a > 0 \mid |\lambda(x^*(a,q))| < 1\}$ and (b) $\sup\{b > 0 \mid |\lambda(x^*(b,q))| < 1\}$.

In order to quantify the impact of the presence of non-responsive traffic, in the first case, we first fix the value of b, *i.e.*, the resource price function, and compute the infimum of all values of a that lead to a stable system with increasing value of q. Fig. 1(a) plots $\inf\{a > 0 \mid |\lambda(x^*(a,q))| < 1\}$ with increasing value of q. The value computed from the eigenvalue in (51) is shown as $'\text{x}'$, and the value calculated from (56) is shown as $'+'$. The figure shows that these two values are identical. As expected from (51) the smallest value of a that leads to a stable system decreases with q. This also confirms that adding non-responsive traffic has different effects from simply reducing the system capacity, which in fact does *not* affect the stability condition.

In the second experiment we fix the value of a and study the supremum of all values of b that result in stability. Fig. 1(b) plots these values as a function of q. Again, the values computed from (51) and (56) are shown as $'\text{x}'$ and $'+'$, respectively. The figure demonstrates the with increasing value of q the system becomes more stable and can tolerate more responsive resource price function.

7 Numerical Examples

In this section we present numerical examples to validate our results presented in the previous section.

7.1 Homogeneous Users

Fig. 2 plots $x(t)$ for $C = 5, T = 200, \kappa = 0.2$, $b = 5$, and various values of a. The value of parameter a is set to $2, 6.1$, and 10, respectively. Note that $a = 2$ yields an eigenvalue $-\frac{b+1}{a} = -3$, which violates the stability condition. This is illustrated in Fig. 2. As one can see the system does not converge to the optimal value of 2.73. On the other hand, the value of $a = 6.1$ leads to a stable system and the rate converges to the optimal value of 1.945 as demonstrated in the figure. When we further increase the value of a to 10 one can see that the utilization at the fixed point decreases with a larger value of a. However, the settling time improves with increasing a. Thus, this presents another trade-off between settling time and resource utilization as mentioned in subsection 5.1.

In the second example we take two homogeneous users with $a = 3$, price function with $b = 5$, and link capacity C of 5. It is clear that for these values of a and b rate control algorithm is unstable since $\frac{5+1}{3} = 2 > 1$. The optimal rates for both users in the absence of delay is given by $x^* = \left(\left(\frac{2}{5}\right)^5\right)^{\frac{1}{9}} = 1.6637$.

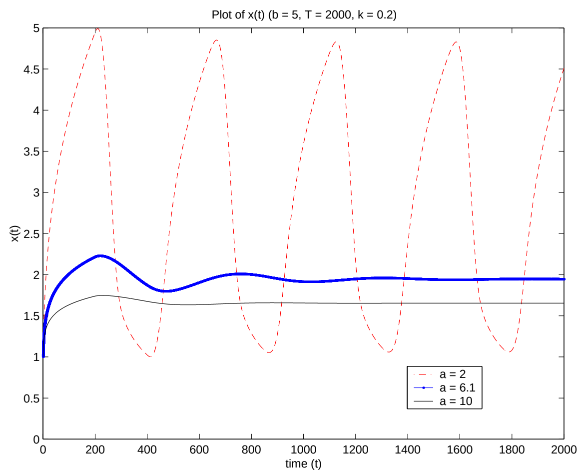


Figure 2: Plot of $x(t)$ for $a = 2$, 6.1, and 10.

Their self imposed upper rate limit will be $C/2 = 2.5$. The lower limit on the solution according to the period two orbit of map F will be given by $F(2/5) = \left(\frac{2}{5}\right)^{\frac{1}{3}} = 0.7368$.

Fig. 3(a) shows the rate waveform for a delay of $T = 1$, which is not sufficient to send the system into the unstable mode, and hence both rates converge to their optimal value of 1.6637 (Fig. 3(b)). However, when delay is increased to $T = 10$, system begins to oscillate as shown in Fig. 3(c). The upper limit of 2.5 and lower limit of 0.7368 can be verified. Finally, in Fig. 3(d), which shows the same waveform for $T = 50$, the waveform is more square-like compared to last figure. In the limit with increasing delay this waveform approaches a square waveform oscillating between the period two orbit of corresponding map.

It is also evident that in both of the oscillating cases the period of the waveform is approximately twice of the delay and the interval between consecutive times when the waveforms cross $y(t) = y^* = 1.6637$ is more than the delay itself. Typically, these oscillating orbits are very difficult to describe as they vary from sinusoidal to square waves with increasing value of delay. This phenomena has been studied earlier in [3].

Clearly, these numerical solutions confirm the upper and lower limits for the trajectories for large enough communication delays. In particular, these periodic orbits remind of a particular periodic solution class devised specifically for delay-differential equations, namely *Slowly Oscillating Periodic (SOP)* orbits [19, 24]. Roughly, an SOP is a periodic orbit with its consecutive zeros (zero corresponds to the fixed point y^* in our case) separated by more than one normalized time unit. The time unit used in our context corresponds to a round-trip time, which arises naturally as a measure for network performance and stability. This also supports the view that round-trip time may be the most useful time scale from the point of view of stability and oscillations [7]. For dynamical (35) we have following conjecture regarding the existence of an SOP:

Conjecture 1 *SOP: For all $0 < \nu < 1/T_0$, where T_0 is given by (18) in linear stability context, (14) has at least one slowly oscillating periodic solution with period $P(\nu) > 2$. Moreover, $T(\nu) \to 2$ as $\nu \to 0$.*

Although proving the existence of an SOP is technically complicated and its asymptotic behavior is even more challenging, we believe that these slowly oscillating periodic orbits are useful for the study of networks and networked control systems to understand the stability and oscillation behavior in the presence of non-negligible delays.

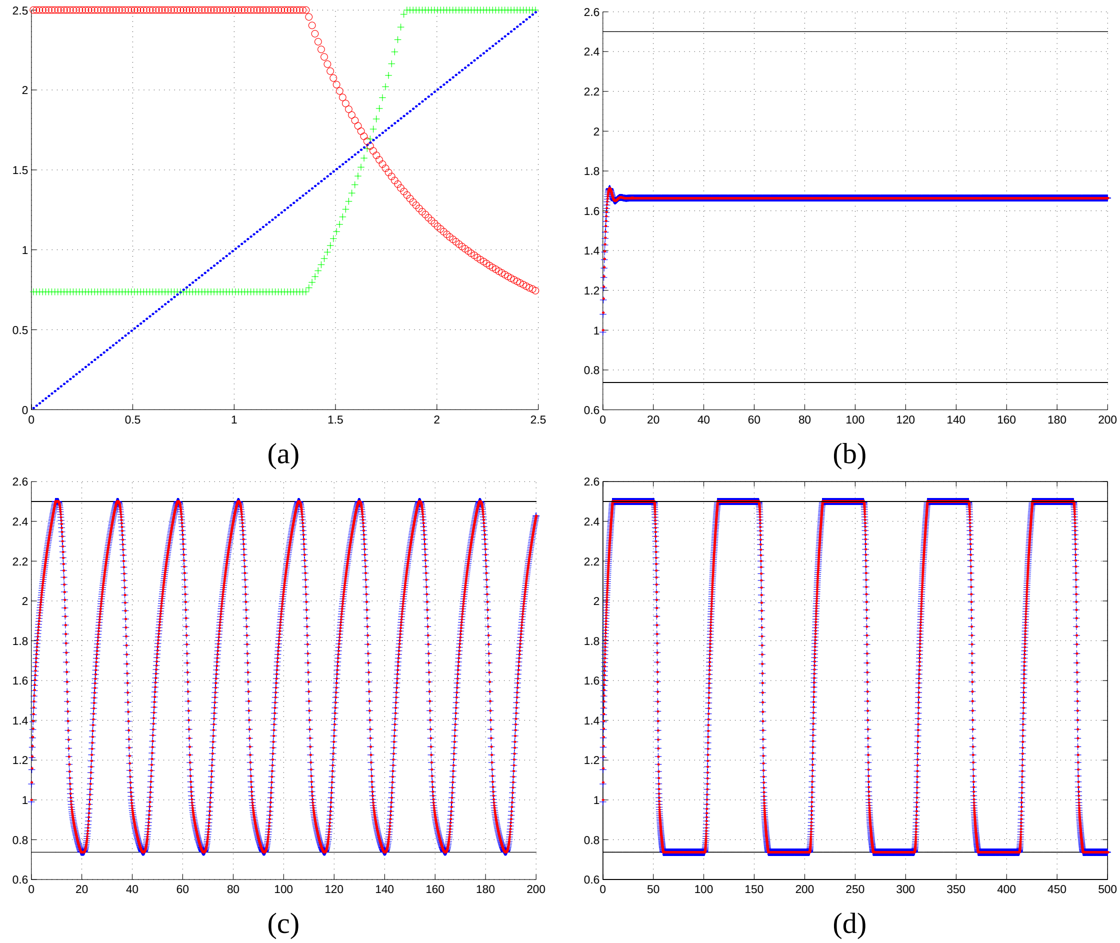


Figure 3: (a) Map given by (35) for above scenario, (b) Rate waveform for the delay of $T = 1$, (c) for the delay of $T = 10$, and (d) for the delay of $T = 50$.

7.2 Heterogeneous Users

In this subsection we present a numerical example with two heterogeneous users. In this example we set the resource price parameter to $b = 2.0$ and users' utility function parameters to $a_1 = 3.1$ and $a_2 = 4.1$ We set the delay T to 100 and 500. Since $a_i > b+1, i = 1,2$, our results state that the system will converge to x^*. Fig. 4 plots $x_i(t), i = 1,2$, for $T = 100$ and 500. As one can see, the system converges to the fixed point of the discrete-time map, which is x^* for both delays. One can also see the synchronization of users' rates in both cases.

7.3 Non-responsive Traffic

The example in this subsection illustrates the effects of non-responsive traffic. There is a single responsive flow with $a = 2.4$ that traverses a link with $C = 10$. The parameter of price function is set to $b = 5$ for the simulation with a delay of $T = 200$ in the system. In the first case, there is no non-responsive flow, while in the second case we introduce non-responsive flow with $q = 2.0$. Fig. 5 plots $x(t)$ in both cases. As one can easily see, the system exhibiting oscillatory behavior without non-responsive traffic, becomes stabilized by the introduction of non-responsive flow. Hence, this demonstrates that the presence of non-responsive flow enhances the stability of system.

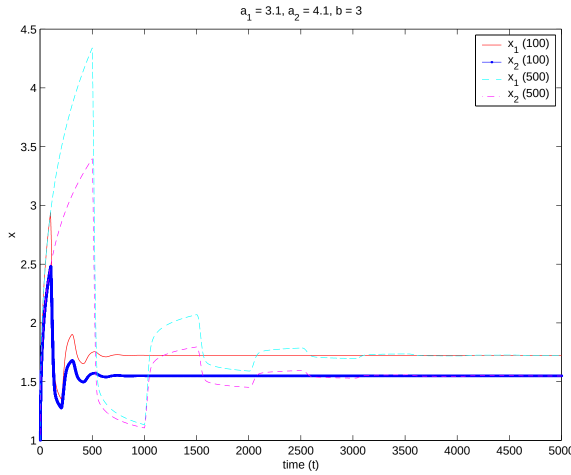


Figure 4: Plot of $\overline{x}(t)$ $(a_1 = 3.1, a_2 = 4.1$, and $b = 2.0)$ with $T = 100$ and 500.

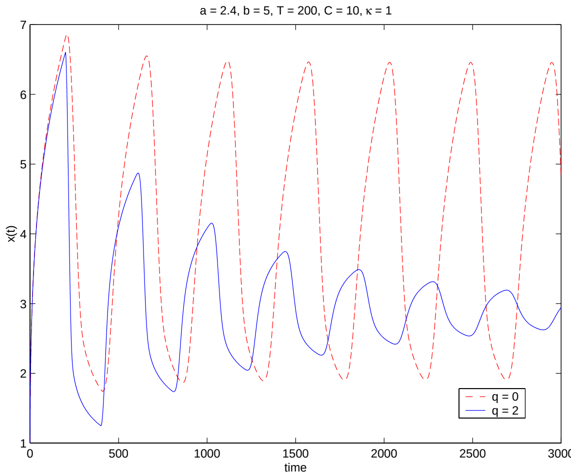
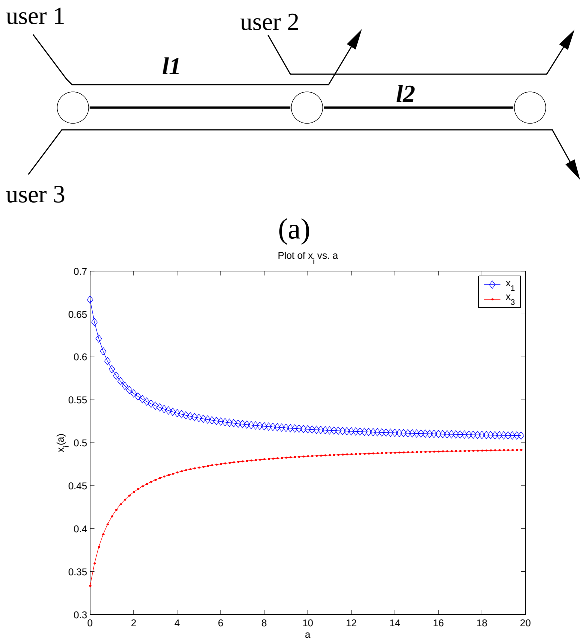


Figure 5: Plot of $\overline{x}(t)$ $(a = 2.4, b = 5$, and $C = 10)$ with $T = 200$.

8 Discussion

In this section we discuss the implications of the results presented in the previous sections on the fairness of the rate control algorithms. In practice achieving high utilization at bottlenecks is an important issue. We have suggested that this may be achieved by dynamically adjusting one of control parameters at the routers to control the desired utilization *e.g.*, parameter γ in (40). Kunniyur and Srikant [16] have proposed a dynamic mechanism that utilizes a virtual queue associated with each link. The idea behind this approach is to adapt the virtual queue capacity in order to maintain certain desired utilization at the bottlenecks. This allows the end users using the algorithm in (4) to solve the SYSTEM problem of maximizing the aggregate utility of the users in (1).

We have suggested in subsection 5.1 that such a dynamic mechanism, where the shape of the price function is preserved but the price function is rescaled by a constant, may not change the stability conditions on the users' utility functions. This implies that the stability of the system improves with increasing q of utility functions assuming that the similar results extends to more general multiple bottleneck cases, which we are currently investigating. One consequence of this would be that the system stability improves with the fairness among the users. Before considering the general case, let us consider an example. Consider the



(b)

Figure 6: Example. (a) topology, (b) plot of $x_i^*(a)$.

example shown in Fig. 6(a), where there are three users that share two links in the network. The capacity of both links is assumed to be one. Suppose that all users have the same utility function $U_i(x_i) = \frac{-1}{a \cdot x_i}, a > 0$. The optimal rates $x^*(a)$ that solve the SYSTEM problem in (1) for a given a are plotted in Fig. 6(b). Since $x_1^*(a) = x_2^*(a)$, we only plotted $x_1^*(a)$ and $x_3^*(a)$. As one can see, as a increases, $x_i^*(a), i = 1, 2, 3$, converge to 0.5, which is the max-min fair allocation. This can be easily explained by the closed form solution of the problem. After a little algebra, one can show that $x_1^*(a) = \frac{\beta}{1+\beta}$, where $\beta = 2^{\frac{1}{1+a}}$. Hence, as $a \downarrow 0$ ($a \uparrow \infty$), we have $\lim_{a \downarrow 0} \beta = 2$ ($\lim_{a \uparrow \infty} \beta = 1$) and $\lim_{a \downarrow 0} x_1^*(a) = 0.\overline{6}$ ($\lim_{a \uparrow \infty} x_1^*(a) = 0.5$).[5] One can also attain similar results for general multiple bottleneck cases, *i.e.*, the solutions to the SYSTEM problem converge to the max-min fair allocation as $a \uparrow \infty$. This suggests that both (max-min) fairness among the users and system stability improves as a increases.

9 Conclusions

We showed that dynamical stability of rate control problem for a simple one resource case is determined by the interaction of underlying utility and price functions. In particular, we demonstrated a fundamental trade-off between users' price elasticity of demand and the responsiveness of the resource. We have proved that when the users' utility or resource price function is too responsive in relation to the other, it leads to network instability. We explicitly characterized this for a class of utility functions. Furthermore, we showed another trade-off between the global stability of system and the utilization of the resource. These results offer some guidelines for jointly designing the end users algorithms and AQM mechanisms at the routers in the presence of a communication delay between end users and network elements. We illustrated that

[5] Note that the limit of solutions to the SYSTEM problem with $a \downarrow 0$ is the solution to the NETWORK problem with $w_i = w$ for all $i \in \mathcal{I}$, *i.e.*, proportionally fair allocation [12].

discrete-time framework arises as a natural tool to study the dynamics of delayed rate control schemes. It also hints at the structure of periodic trajectories and their bounds.

Finally, we conjecture that SOP orbits may be relevant to study of the structure of periodic orbits arising in engineering applications. These periodic orbits have been studied extensively in mathematics community and also arise when the delay is state-dependent, which is a useful context in networking [7]. These SOP orbits support the earlier belief that the round-trip time may be the most relevant time scale for network stability studies.

References

[1] R. Agrawal, A. Bedekar, R. J. La, and V. G. Subramanian. C3wpf scheduler. In *Proc. ITC*, 2001.

[2] T. Alpcan and T. Basar. A utility-based congestion control scheme for Internet-style networks with delay. In *Proc. of INFOCOM*, 2003.

[3] S. N. Chow, J. K. Hale, and W. Huang. From sine wave to square wave in delay equations. *Proc. Royal Soc, Edinburgh*, 120A:223–229, 1992.

[4] S. Deb and R. Srikant. Global stability of congestion controllers for the internet. In *Proc. of IEEE CDC*, 2002.

[5] P. M. Fitzpatrick. *Advanced Calculus*. PWS Publishing Company, 1996.

[6] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*. Springer-Verlag, NY, 1983.

[7] C. Hollot, V. Misra, D. Towsley, and W. Gong. A control theoretic analysis of RED. In *Proc. of IEEE INFOCOM*, Anchorage AK, 2001.

[8] A. F. Ivanov, E. Liz, and S. I. Trofimchuk. Global stability of a class of scalar nonlinear delay differential equations. *Accepted for Publication, Preprint from Author*, 2003.

[9] A. F. Ivanov, M. A. Pinto, and S. I. Trofimchuk. Global behavior in nonlinear systems with delayed feedback. *Proc. Conference of Decison and Control, Sydney*, 2000.

[10] A. F. Ivanov and A. N. Sharkovsky. Oscillations in singularly perturbed delay equations. *Dynamics Reported*, 1:165–224, 1991.

[11] R. Johari and D. Tan. End-to-end congestion control for the internet: Delays and stability. *IEEE Transactions on Networking*, 2000.

[12] F. Kelly. Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, 8(1):33–7, January 1997.

[13] F. Kelly. Mathematical modelling of the internet. In *Bjorn Engquist and Wilfried Schmid (Eds.), Mathematics Unlimited – 2001 and Beyond*. Springer-Verlag, Berlin, 2001.

[14] F. Kelly, A. Maulloo, and D. Tan. Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49(3):237–252, March 1998.

[15] S. Kunniyur and R. Srikant. End-to-end congestion control schemes: Utility functions, Random losses and ECN marks. In *Proc. of IEEE INFOCOM*, Tel-Aviv, Israel, 2000.

[16] S. Kunniyur and R. Srikant. A time scale decomposition approach to adaptive ecn marking. In *Proc. of IEEE INFOCOM*, Anchorage, AK, 2001.

[17] R. J. La, P. Ranjan, and E. H. Abed. Nonlinearity of TCP and instability with RED. In *Proc. SPIE ITCom*, 2002.

[18] S. Low and D. Lapsley. Optimization flow control. *IEEE/ACM Transactions on Networking*, 7(6):861–874, June 1997.

[19] J. Mallet-Paret and R. D. Nussbaum. Global continuation and asymptotic behaviour for periodic solutions of a differential-delay equation. *Annali di Matematica Pura ed Applicata*, 145(4):33–128, 1986.

[20] S. Niculescu, E. I. Verriest, L. Dugard, and J. Dion. Stability and robust stability of time delay systems: A guided tour. *Stability and Control of Time-Delay System*, LNCS: 228:16–17, 1997.

[21] P. Ranjan, R. J. La, and E. H. Abed. Trade-offs in rate control with communication delay. Technical report, ISR, University of Maryland at College Park, 2003.

[22] Hassan Sedaghat. *Nonlinear difference equations: theory with applications to social science models*. Kluwer Academic Publishers, Boston, 2003.

[23] A N Sharkovsky, S. F. Kolyada, A. G. Sivak, and V. V. Fedorenko. *Dynamics of one Dimensional maps*. Kluwer Academic, 1997.

[24] A N Sharkovsky, Y. L. Maistrenko, and E. Y. Romaneko. *Difference Equations and their applications: (Trans. from Russian)*. Kluwer Academic, 1995.

[25] Hal R. Varian. *Intermediate Microeconomics*. Norton, New York, 1996.

[26] E. I. Verriest and A. F. Ivanov. Robust stability of systems with delayed feedback. *Circuits, Systems and Signal Processing*, 13(2-3):213–222, 1994.

[27] Z. Wang and F. Paganini. Global stability with time-delay in network congestion control. In *Proc. of IEEE CDC 2002*, Las Vegas, 2002.

[28] J. Wen and M. Arcak. A unifying passivity framework for network flow control. In *Proc. of IEEE INFOCOM*, 2003.

Appendices

A Proof of Theorem 2

Before proving the theorem we will state a lemma which is the key to the proof of Theorem 2.

Lemma 3 *Suppose that an interval J is mapped by F into itself. If none of the endpoints of the interval $F(J)$ is fixed point then for every $\phi \in X_J = C([-1,0],J)$ there exists a finite $t_0 = t_0(\phi,\nu,\kappa)$ such that $y_\phi^\nu(t) \in F(J)$ for all $t \geq t_0$.*

Proof: From Theorem 1 it is clear that $y_\phi^\nu \in J$ for all $t \geq 0$. The claim here is that after certain time t_0 it will be limited by $F(J) \subset J$.

First, assume that $\phi(0) \in \overline{F(J)}$. Then it can be shown that $y_\phi^\nu(t) \in F(J)$ for all $t \geq t_0$ by contradiction. Suppose that this is not true and let t_0 be the first time when $y_\phi^\nu(t)$ leaves the interval $\overline{F(J)}$. In particular assume that it leaves from the right end, *i.e.*, every right-sided neighborhood of t_0 contains a point t_1 such that $y_\phi^\nu(t_1) > \sup \overline{F(J)}$. Then, the same neighborhood also contains a point $t_2, t_0 < t_2 < t_0 + 1$, such that $y_\phi^\nu(t_2) > \sup \overline{F(J)}$ and $\dot{y}_\phi^\nu(t_2) > 0$. As $y_\phi^\nu(t) \in J$ for all $t \in [t_0 - 1, t_0]$, we have $\dot{y}(t_2) = \frac{\kappa(y(t_2))(f(g^{-1}(y(t_2-1))) - y(t_2))}{\nu} < 0$ from (14). This contradicts the earlier assumption that $\dot{y}(t_2) > 0$. The other case where $y_\phi^\nu(t)$ leaves the interval from the left end can be handled similarly.

Now assume that $\phi(0) \notin \overline{F(J)}$. In particular, let $\phi(0) > \sup \overline{F(J)}$. Claim here is that $y_\phi^\nu(t)$ is decreasing for all $t \in [0, t_0]$, where $t_0 \leq \infty$ is the first point with $y_\phi^\nu(t_0) = \sup \overline{F(J)}$. We first argue that $t_0 < \infty$ by contradiction. Suppose $t_0 = \infty$ and, hence, $y_\phi^\nu(t) > \sup \overline{F(J)}$ for all $t \geq 0$. From (14) we have $\dot{y}(t_2) = \frac{\kappa(y(t_2))(f(g^{-1}(y(t_2-1))) - y(t_2))}{\nu} < 0$ because $f(g^{-1}(y(t_2 - 1))) \leq \sup \overline{F(J)}$. Then there exists a limit $\bar{y} = \lim_{t\to\infty} y_\phi^\nu(t) > \sup \overline{F(J)}$ due to Bolzano-Weierstrass theorem [5] which says that every strictly decreasing sequence which is bounded from below has a limit. As $\bar{y}$ is not a fixed point of map F, $\kappa(\bar{y})(\bar{y} - f(g^{-1}(\bar{y}))) := \delta > 0$. This tells us from (14) that $\dot{y}(t) = \frac{\kappa(y(t))(f(g^{-1}(y(t-1))) - y(t))}{\nu} < -\frac{\delta}{2\nu}$ for large enough t. This implies that $y_\phi^\nu(t) \to -\infty$ as $t \to \infty$, which is a contradiction, because this means that $y_\phi^\nu(t)$ crosses $\sup \overline{F(J)}$ for some finite t. Hence, $t_0 < \infty$. Now, we invoke the first part of proof where system is restarted at time t_0 with $y(t_0) = \sup \overline{F(J)}$ and $y(t) \in J\ \forall t \in [t_o - 1\ t_0]$. Using the same argument we can show that $y(t) \in F(J)$ for all $t \geq t_0$. The other case can be handled similarly. ∎

Now we provide the proof for Theorem 2 using this lemma. For any $\phi \in X_{J_0}$, define $m = \inf\{\phi(s),\ s \in [-1\ 0]\}$ and $M = \sup\{\phi(s),\ s \in [-1\ 0]\}$. Clearly, $[m\ M] \subset J_0$. Let J' be the smallest closed invariant interval containing $[m\ M]$ which is a subset of J_0. Then, from the existence of fixed point of the map F, $J' \supset F(J') \supset F(F(J')) \supset \ldots$ and $\cap_{i\geq 0} F^i(J') = x^*$. Using the invariance result and Lemma 3 repeatedly one can find arbitrarily small estimates for the range of trajectories with large enough t. Thus, the theorem follows.

B Proof of Theorem 4

Before we state the invariance theorem we first prove a proposition which establishes orthant invariance for a vector under the multiplication with a positive diagonal matrix.

Proposition 2 *(Orthant-Invariance) For a diagonal positive matrix $K \in \Re^{N\times N}$ and an arbitrary vector $v \in \Re^N$, Kv remains in the same orthant as v.*

Proof: Any vector v can be expressed as a linear combination of basis vectors $\{e_i\}$ or $v = \Sigma_i c_i e_i$. Clearly, due to the diagonal structure of positive matrix K, $Kv = \Sigma_i k_{ii} c_i e_i$ which means that all the coefficients c_i retain their original sign even after the multiplication with K. This ensures the fact that they remain in the same orthant. ∎

Clearly, diagonal structure of positive gain matrix $\kappa(\cdot)$ is useful to determine the directions of right hand side of (24). We note here that diagonal structure gives us more than required in the sense that a non-diagonal positive $\kappa(\cdot)$ can ensure that the right hand side of (24) stays directed towards the interior of a convex domain on any boundary point.

Let $t_0 \geq 0$ be the first time at which $\bar{y}^\nu(t_0) \in D$ and the solution $\bar{y}^\nu(t)$ leaves the domain D for $t \geq t_0$. This implies that $\nu\dot{\bar{y}}^\nu(t_0)$ is directed towards the outside of the domain. Now, since t_0 is the first such point at which the trajectory leaves the domain D, $F(\bar{y}^\nu(t_0-1))$ lies in the domain D. Therefore, the vectors $F(\bar{y}^\nu(t_0-1)) - \bar{y}^\nu(t_0)$ and $\kappa(\bar{y}^\nu(t_0))\,(F(\bar{y}^\nu(t_0-1)) - \bar{y}^\nu(t_0))$ will both be directed towards the inside of the domain D because $\kappa(\bar{y}^\nu(t_0)) > 0$ and D is convex. This holds because of the orthant-invariance property of vectors under the multiplication with positive diagonal matrix as shown in Proposition 2 and the assumption that D is a product space. However, $\nu\dot{\bar{y}}^\nu(t_0) = \kappa(\bar{y}^\nu(t_0))\;(F(\bar{y}^\nu(t_0-1))\; - \bar{y}^\nu(t_0))$ from (24), which is a contradiction.

C Proof of Theorem 5

We will first prove a lemma which will be used to prove the theorem.

Lemma 4 *Let V be any open product space containing $D_1 \supset F(D_0)$ and contained in D_0 and arbitrary initial functions $\phi \in Y_{D_0}$. (i) If $\phi(0)$ is in closure of the set V, $cl(V)$, then $\bar{y}^\nu_\phi$ is in the closure of V for all $t \geq 0$. (ii) If $\phi(0)$ is not in the closure of V then there exists a finite time $t_0 = t_0(\phi, D_0, \kappa(\cdot))$ such that $\bar{y}^\nu_\phi \in \partial V$ and $\bar{y}^\nu_\phi \in cl(V)$ for all $t \geq t_0$ with ∂V denoting the boundary of V.*

Proof: First, assume $\phi(0) \in cl(V)$. Then, one can show that $\bar{y}^\nu_\phi(t) \in \partial V$ for all $t \geq 0$ following a similar argument in the proof of Theorem 4.

Now suppose that $\phi(0) \notin cl(V)$, and let t_0 the first time such that $\bar{y}^\nu_\phi(t_0) \in \partial V$. Then, one can show that $\bar{y}^\nu_\phi \in cl(V)$ for all $t \geq t_0$, again, following a similar argument in the proof of Theorem 4.

Suppose that we begin with $\phi(0) \notin cl(V)$ and $\bar{y}^\nu_\phi \notin cl(V)$ for all time $t \geq 0$. Also, let M be the maximal open product space containing V such that $M \cap cl(\{\bar{y}^\nu_\phi(t), t \geq 0\}) = \emptyset$. Note that it is possible for M to coincide with V. Since M is maximal such set there exists a sequence $t_i, i = 1, 2, \cdots$, such that $\bar{y}^\nu_\phi(t_i) \to y_0 \in \partial M$. Consider a sequence of finite vectors $\bar{r}_i = \frac{\kappa(\bar{y}^\nu_\phi(t_i))}{\nu}\left[F(\bar{y}^\nu_\phi(t_i-1)) - \bar{y}^\nu_\phi(t_i)\right]$, which form the right hand side of (24) at these time instants. Because $M \subset D_0$ and $F(M) \subset D_1$, these vectors r_i are bounded away from zero, *i.e.,* there exists some $\delta > 0$ such that $|r_i| \geq \delta$, and are directed strictly towards the inside of domain M. Since the origins of the vectors $\bar{r}_i$ converge to the point $y_0 \in \partial M$ and $\bar{r}_i$'s point to the inside of M, there must exist some $t^{'}$ such that $\bar{y}^\nu_\phi(t^{'}) \in M$, which is a contradiction. ■

Now we provide an easy proof for Theorem 5. Since it is possible to construct a sequence of nested open product spaces $U_{i+1} \subset U_i$ such that $F(U_i) \subset U_{i+1}$ and $\cap_{i\geq 0} U_i = \{y^*\}$, by repeated application of Lemma 4 there exists a sequence $\{t_i\} \to \infty$ such that $\bar{y}^\nu_\phi(t_i) \in cl(U_i)$ and $\bar{y}^\nu_\phi(t)$ remains inside due to the invariance theorem for $t \geq t_i$. As $\cap_{i\geq 0} U_i = \{y^*\}$, asymptotic convergence and implied stability for y^* follows.

D Proof of Lemma 1

From (45) we have

$$\hat{F}_i(\bar{\alpha}x^*) = (\bar{\alpha}x_i^*)^{-\frac{1}{a_i}} \left(\sum_{l \in r_i} \left(\bar{\alpha} \frac{\sum_{j \in I_l} x_j^*}{C_l} \right)^{b_l} \right)^{-\frac{1}{a_i}}$$

$$= (\bar{\alpha}x_i^*)^{-\frac{1}{a_i}} \left(\sum_{l \in r_i} \bar{\alpha}^{b_l} \left(\frac{\sum_{j \in I_l} x_j^*}{C_l} \right)^{b_l} \right)^{-\frac{1}{a_i}} .$$

Note that since $\bar{\alpha} \geq \alpha > 1$,

$$\sum_{l \in r_i} \bar{\alpha}^{b_l} \left(\frac{\sum_{j \in I_l} x_j^*}{C_l} \right)^{b_l} \leq \sum_{l \in r_i} \bar{\alpha}^{b^i_{max}} \left(\frac{\sum_{j \in I_l} x_j^*}{C_l} \right)^{b_l} = \bar{\alpha}^{b^i_{max}} \sum_{l \in r_i} \left(\frac{\sum_{j \in I_l} x_j^*}{C_l} \right)^{b_l} \quad (57)$$

Therefore,

$$\begin{aligned}
\hat{F}_i(\overline{\alpha}x^*) &= (\overline{\alpha}x_i^*)^{-\frac{1}{a_i}}\left(\sum_{l\in r_i}\overline{\alpha}^{b_l}\left(\frac{\sum_{j\in I_l}x_j^*}{C_l}\right)^{b_l}\right)^{-\frac{1}{a_i}}\\
&\geq (\overline{\alpha}x_i^*)^{-\frac{1}{a_i}}\left(\overline{\alpha}^{b^i_{max}}\sum_{l\in r_i}\left(\frac{\sum_{j\in I_l}x_j^*}{C_l}\right)^{b_l}\right)^{-\frac{1}{a_i}}\\
&= \overline{\alpha}^{-\frac{1}{a_i}}\overline{\alpha}^{-\frac{b^i_{max}}{a_i}}x_i^{*-\frac{1}{a_i}}\left(\sum_{l\in r_i}\left(\frac{\sum_{j\in I_l}x_j^*}{C_l}\right)^{b_l}\right)^{-\frac{1}{a_i}}\\
&= \overline{\alpha}^{-\frac{b^i_{max}+1}{a_i}}x_i^*\\
&\geq \overline{\alpha}^{\sigma}x_i^*
\end{aligned}$$

where the second inequality follows from $\sigma \leq -\frac{b^i_{max}+1}{a_i}$ for all $i \in \mathcal{I}$. Therefore, if $\overline{\beta} < \overline{\alpha}^{\sigma}$, then $\overline{\beta}$ satisfies the second condition in (46).

Similarly,

$$\begin{aligned}
\hat{F}_i(\overline{\beta}x^*) &= \left(\overline{\beta}x_i^*\right)^{-\frac{1}{a_i}}\left(\sum_{l\in r_i}\overline{\beta}^{b_l}\left(\frac{\sum_{j\in I_l}x_j^*}{C_l}\right)^{b_l}\right)^{-\frac{1}{a_i}}\\
&\leq \left(\overline{\beta}x_i^*\right)^{-\frac{1}{a_i}}\left(\overline{\beta}^{b^i_{max}}\sum_{l\in r_i}\left(\frac{\sum_{j\in I_l}x_j^*}{C_l}\right)^{b_l}\right)^{-\frac{1}{a_i}}\\
&= \overline{\beta}^{-\frac{1}{a_i}}\overline{\beta}^{-\frac{b^i_{max}}{a_i}}x_i^{*-\frac{1}{a_i}}\left(\sum_{l\in r_i}\left(\frac{\sum_{j\in I_l}x_j^*}{C_l}\right)^{b_l}\right)^{-\frac{1}{a_i}}\\
&= \overline{\beta}^{-\frac{b^i_{max}+1}{a_i}}x_i^*\\
&\leq \overline{\beta}^{\sigma}x_i^*
\end{aligned}$$

where the first inequality follows from that $\overline{\beta} < 1$. Hence, if $\overline{\beta}^{\sigma} < \overline{\alpha}$ or $\overline{\beta} > \overline{\alpha}^{1/\sigma}$, then $\hat{F}(\overline{\beta}x^*) < \overline{\alpha}x^*$ and satisfies the first condition in (46).

E Proof of Lemma 2

We first show that $D_{k+1} \subset int(D_k)$, for all $k = 0, 1, \cdots$. This can be easily shown as follows. We know from the proof of Lemma 1 that $\overline{\beta} < \overline{\alpha}^{\sigma}$ and $\overline{\beta}^{\sigma} < \overline{\alpha}$. Hence, $D_1 \subset int(D_0)$. Now similarly as before, since $\overline{\beta} < \overline{\alpha}^{\sigma}$ we have $\overline{\alpha}^{\sigma^2} < \overline{\beta}^{\sigma}$, and because $\overline{\beta}^{\sigma} < \overline{\alpha}$, $\overline{\alpha}^{\sigma} < \overline{\beta}^{\sigma^2}$, which follows from that $-1 < \sigma < 0$. By repeating this we get $\overline{\alpha}^{\sigma^{k+1}} < \overline{\beta}^{\sigma^k}$ and $\overline{\alpha}^{\sigma^k} < \overline{\beta}^{\sigma^{k+1}}$ for odd k and $\overline{\beta}^{\sigma^{k+1}} < \overline{\alpha}^{\sigma^k}$ and $\overline{\beta}^{\sigma^k} < \overline{\alpha}^{\sigma^{k+1}}$ for even k, proving that $D_{k+1} \subset int(D_k)$. The fact that $\cap_{k=0}^{\infty}D_k = \{x^*\}$ follows trivially from that $\lim_{k\to\infty}\sigma^k = 0$ because $|\sigma| < 1$. Hence, $\lim_{k\to\infty}\overline{\alpha}^{\sigma^k} = 1 = \lim_{k\to\infty}\overline{\beta}^{\sigma^k}$.

Now we prove that $\hat{F}(D_k) \subset D_{k+1}, k = 0, 1, \cdots$. Here we prove this only for the case where k is even. The other case can be proved similarly. Following the same approach in the proof of Lemma 1 we have

$$\hat{F}_i(\overline{\alpha}^{\sigma^k}x^*) = \left(\overline{\alpha}^{\sigma^k}x_i^*\right)^{-\frac{1}{a_i}}\left(\sum_{l\in r_i}\left(\overline{\alpha}^{\sigma^k}\frac{\sum_{j\in I_l}x_j^*}{C_l}\right)^{b_l}\right)^{-\frac{1}{a_i}}$$

$$= \left(\overline{\alpha}^{\sigma^k} x_i^*\right)^{-\frac{1}{a_i}} \left(\sum_{l \in r_i} \overline{\alpha}^{\sigma^k b_l} \left(\frac{\sum_{j \in I_l} x_j^*}{C_l}\right)^{b_l}\right)^{-\frac{1}{a_i}} .$$

Since $\overline{\alpha}^{\sigma^k} > 1$ (because k is even),

$$\sum_{l \in r_i} \left(\overline{\alpha}^{\sigma^k}\right)^{b_l} \left(\frac{\sum_{j \in I_l} x_j^*}{C_l}\right)^{b_l} \leq \sum_{l \in r_i} \left(\overline{\alpha}^{\sigma^k}\right)^{b^i_{max}} \left(\frac{\sum_{j \in I_l} x_j^*}{C_l}\right)^{b_l} = \left(\overline{\alpha}^{\sigma^k}\right)^{b^i_{max}} \sum_{l \in r_i} \left(\frac{\sum_{j \in I_l} x_j^*}{C_l}\right)^{b_l} \tag{58}$$

Therefore,

$$\begin{aligned}
\hat{F}_i(\overline{\alpha}^{\sigma^k} x^*) &= \left(\overline{\alpha}^{\sigma^k} x_i^*\right)^{-\frac{1}{a_i}} \left(\sum_{l \in r_i} \overline{\alpha}^{\sigma^k b_l} \left(\frac{\sum_{j \in I_l} x_j^*}{C_l}\right)^{b_l}\right)^{-\frac{1}{a_i}} \\
&\geq \left(\overline{\alpha}^{\sigma^k} x_i^*\right)^{-\frac{1}{a_i}} \left(\left(\overline{\alpha}^{\sigma^k}\right)^{b^i_{max}} \sum_{l \in r_i} \left(\frac{\sum_{j \in I_l} x_j^*}{C_l}\right)^{b_l}\right)^{-\frac{1}{a_i}} \\
&= \left(\overline{\alpha}^{\sigma^k}\right)^{-\frac{1}{a_i}} \left(\overline{\alpha}^{\sigma^k}\right)^{-\frac{b^i_{max}}{a_i}} x_i^{*-\frac{1}{a_i}} \left(\sum_{l \in r_i} \left(\frac{\sum_{j \in I_l} x_j^*}{C_l}\right)^{b_l}\right)^{-\frac{1}{a_i}} \\
&= \left(\overline{\alpha}^{\sigma^k}\right)^{-\frac{b^i_{max}+1}{a_i}} x_i^* \\
&\geq \left(\overline{\alpha}^{\sigma^k}\right)^{\sigma} x_i^* \\
&= \overline{\alpha}^{\sigma^{k+1}} x_i^* .
\end{aligned}$$

Similarly one can show that $\hat{F}_i(\overline{\beta}^{\sigma^k} x^*) \leq \overline{\beta}^{\sigma^{k+1}} x_i^*$. This completes the proof of the lemma.