A Nonlinear Algorithm for ABR Congestion Control in ATM Networks

Orhan Ç. IMER[*], Tamer BAŞAR[†‡], R. SRIKANT[§]

Abstract

We present a novel ABR congestion control algorithm which is adaptive to changing network conditions and robust to network delays. The algorithm is also easy to implement, as it only requires a single design parameter, and no centralized knowledge about the status of the network. Further, max-min fairness and queue length stability under MCR and PCR constraints are automatically achieved without requiring any additional computation or tuning of the parameters.

1 Introduction

Asynchronous Transfer Mode (ATM) technology is intended to support a variety of services and applications. Performance of an ATM network is fundamentally linked to the ability of the network to provide Quality of Service (QoS) differentiation for different network applications. For that, a set of five different service categories are specified by the ATM Forum [1]:

- Constant Bit Rate (CBR): This service category is used by connections that request a static amount of bandwidth that is continuously available during the connection lifetime. Telephone and television use this service.
- Variable Bit Rate (VBR): The cell rate is variable and is mainly intended for bursty sources. This service can be real-time VBR (video conferencing) or nonreal-time VBR (multimedia e-mail).
- Available Bit Rate (ABR): In this service category, the cell rate depends on the availability of the network. It is basically designed for bursty traffic whose bandwidth range is known roughly. A rate control mechanism is specified by the ATM Forum. Examples of this service category include browsing the Web and e-mail.
- Unspecified Bit Rate (UBR): This category uses the leftover network capacity. UBR service does not specify traffic-related service guarantees. Background file transfer uses this service.

[*]Department of Electrical and Computer Engineering and Coordinated Science Laboratory, University of Illinois, 1308 West Main Street, Urbana, IL 61801-2307, USA. Email: imer@uiuc.edu.

[†]Department of Electrical and Computer Engineering and Coordinated Science Laboratory, University of Illinois, 1308 West Main Street, Urbana, IL 61801-2307, USA. Tel: (217) 333-3607; Fax: (217) 244-1653; Email: tbasar@decision.csl.uiuc.edu.
Research supported by NSF Grant NSF ANI 98-13710, AFOSR MURI Grant AF DC 5-36128.

[‡]Author for correspondence.

[§]Department of General Engineering and Coordinated Science Laboratory, University of Illinois, 1308 West Main Street, Urbana, IL 61801-2307, USA. Tel: (217) 333-2457; Fax: (217) 244-1642; Email: rsrikant@uiuc.edu.
Research supported by an NSF CAREER Award NCR 9701525, NSF Grant ANI 98-13710.

- Guaranteed Frame Rate (GFR): This service category is intended to support non-real-time applications requiring a minimum rate guarantee. It does not require adherence to a rate control protocol. An example application is frame relay interworking.

In principle, any UBR or GFR application can take advantage of the ABR flow control protocol to achieve a low cell loss ratio. Thus, the network traffic generated by various sources can be thought of as composed of CBR, VBR and ABR connections only.

When a virtual circuit (VC) is established between a source and a destination, both the customer (source) and the carrier (switch) must agree on a contract defining the Quality of Service (QoS). The ATM Forum defines three QoS parameters:

- Cell Delay Variation (CDV): The QoS parameter CDV describes the variability in the pattern of cell arrivials with reference to the negotiated peak rate.

- Maximum Cell Transfer Delay (maxCTD): The maxCTD for a given connection is specified in terms of the probability density function of Cell Transfer Delay (CTD). Any cell delivered maxCTD after its transmission is assumed to be late or lost.

- Cell Loss Ratio (CLR): The Cell Loss Ratio is defined for a connection as:

$$\mathrm{CLR} = \frac{\text{Lost Cells}}{\text{Total Transmitted Cells}}$$

All ATM service categories attempt to minimize CLR by reducing network cell losses to an acceptable level. In CBR and VBR services, a traffic contract specifying the CDV and maxCTD guarantees is negotiated during the virtual circuit setup phase and is maintained for the duration of the connection. The ABR service, on the other hand, does not require bounding the delay or the delay variation experienced by a given connection. Therefore, with CBR and VBR traffic, it is generally not possible for the source to decrease its rate, even when an intermediate switch becomes congested, because of the QoS guarantees made at VC setup time. However, ABR sources might adjust their rates to the level of available service at times of congestion. Thus, ABR traffic can be used to control congestion in the network.

About four years ago, the ATM Forum adopted a rate-based congestion control scheme for the ABR service [1]. In this scheme, explicit rate control messages are sent from intermediate nodes to the sources using special cells called Resource Management (RM) cells. The goal of this congestion control mechanism is to fairly share the bandwidth left over from high-priority traffic (CBR and VBR) among the ABR sources while making sure that the links throughout the network are fully utilized.

Although the rate-based congestion control schemes are standardized, developing good explicit rate computation algorithms is still an open issue. As the link speeds continue to rise, the delay-bandwidth product (i.e., the product of the round-trip propagation delay and the link capacity) increases. An issue of importance that arises in this context is how to deal with action delays, which is the time from the moment control information is sent to a source, until an action is taken by it, and until subsequently that action affects the state of the switch that initiated that command.

In designing rate control algorithms, another challenge is posed by the traffic characteristics of the network. The term *connection lifetime* refers to the duration of a switched virtual connection (SVC). ABR sources can only initiate SVC's, as the permanent virtual connections (PVC) are reserved for CBR and VBR services [1]. Some SVC's may have a shorter connection lifetime than the smallest round-trip delay in the network. Such connections are said to be *short-lived*. It is hard, if not impossible, to exercise any control over these type of connections, as by the time the feedback information is received by the sender, the transmission process is already over. This fact manifests itself in the network traffic as frequent bursts, which makes it difficult to predict the exact number of ABR sources at a given instant.

In this paper, we introduce a control-based mathematical model that helps us address these design issues. Our initial modeling assumption is that of a single ATM switch shared by a number of ABR sources. We assume that the ATM switch design uses output queueing, in which the cells destined to the same output line are queued on a finite-size output buffer, and are served on a FIFO (First In First Out) basis at a rate available to the ABR traffic. We propose a robust adaptive congestion control algorithm, which requires only the knowledge of the maximum round-trip network delay and the maximum number of simultaneous connections. We assume that the switch keeps track of the queue length, available service rate, and incoming cell rate at each one of its output lines. We show that if the gain of the proposed controller is picked properly, our algorithm achieves max-min fairness along with queue length stability, despite inaccuracies in the knowledge of actual number of ABR sources and their corresponding round-trip delays. Finally, we present some simulation results to support and illustrate the approach taken.

Several other types of ABR congestion control designs have been considered. We briefly summarize these here to compare and contrast them with our approach. The simplest feedback control mechanism is called *rate matching*. In rate matching, the node measures the average rate available to ABR sources at periodic intervals and simply divides a fraction of this capacity equally among the various users. This is the basic approach used in [2], although several modifications are used in the actual implementation. The main advantage of this scheme is its simplicity, but it is difficult to control queue length optimally to avoid buffer overflows. However, this scheme **is** stable (i.e., the queue length remains bounded in an appropriate stochastic sense [3]). Queue length information is not used in the basic algorithm, although [2] allows one to incorporate queue length information in an ad hoc manner. Alternatively, this problem can be viewed as a feedback control problem where queue length is used for explicit feedback. This approach is used in [4] - [5] to study this problem using classical control techniques or using a state-space approach. As in rate matching, the primary goal is not optimality, but simply queue length stability. In these approaches, the available bandwidth to ABR sources is treated as an unmodeled disturbance. Thus, the algorithms in [4] - [5] ensure stability in the presence of this disturbance, but do not address the issue of performance. In a recently published work [6], a closed-loop proportional-derivative controller is proposed, which achieves max-min fairness plus queue length stability, but the design falls short of addressing the issue of robustness against uncertainty in delays.

This paper is organized as follows. In the next section, we describe the ATM ABR service in detail. The mathematical model of an ATM switch is given in Section 3. In Section 4, we present the robust adaptive algorithm, and the analysis of the algorithm is carried out in Section 5. In Section 6, we present some simulation results, and the paper ends with the concluding remarks of Section 7.

2 Available Bit Rate Service

2.1 Flow Control Model for ABR

In the ABR service, the source adapts its rate to changing network conditions. As mentioned in the introduction, information about the state of the network, such as bandwidth availability, state of congestion, and impending congestion, is conveyed to the source through special control cells called RM cells. ABR flow control occurs between a source and a destination, which are connected via bi-directional links. The forward direction is the direction from the source to the destination, and the backward direction is the direction from the destination to the source.

A source generates forward RM cells every Nrm data cells, where Nrm is generally taken to be 32. These cells travel along the same path as the data cells but are treated specially by the switches along the way (Fig. 1). The switch may:

- Directly insert feedback control information into RM cells by using the Explicit Rate (ER) field of RM cells.
- Provide binary feedback by marking the Congestion Indication (CI) or No Increase (NI) bit in the RM cells.
- Indirectly inform the source about congestion by setting the Explicit Forward Congestion Indication (EFCI) bit in the data cell header, and rely on the destination to convey congestion information back to the source by marking the CI bit in the backward RM cells it generates.
- Spontaneously generate backward RM cells and ship them back to the source.

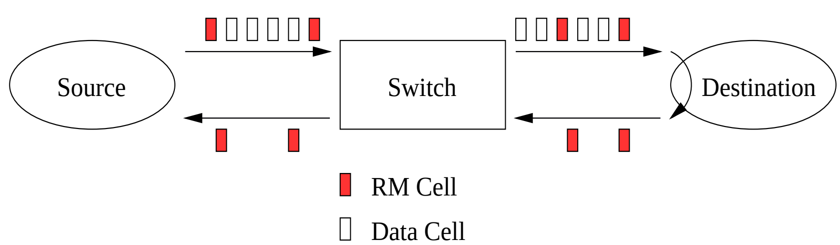

Figure 1: ABR traffic management model.

Note that Fig. 1 is only a generic representation of the control loop, as there may be more than one switch between the source and the destination. On the establishment of an ABR connection, the source specifies to the network both a maximum required bandwidth and a minimum usable bandwidth. These are designated as peak cell rate (PCR), and the minimum cell rate (MCR), respectively. The MCR may be specified as zero. The bandwidth available from the network may vary, but should not become less than MCR. Each

ABR source has a current cell rate, allowed cell rate (ACR), which it must modify upon receiving feedback from the network via RM cells. ACR always falls somewhere between MCR and PCR. When a source sends out a forward RM cell, it sets the ER field of the RM cell to the rate at which it would currently like to transmit. As the RM cell passes through the various switches on the way to the destination and back to the source, those that are congested may reduce the ER. When the source receives the RM cell back, it takes one of the following actions depending on the settings of the ER field and CI and NI bits.

- When there is no congestion (both CI and NI bits are not set), ACR can be increased (but not above PCR) by a quantity RIF*PCR, where RIF is the rate increase factor. However, ACR cannot be increased above the explicit rate specified in the ER field. RIF is negotiated at VC setup phase.
- When a source receives a backward RM cell with CI bit set, it decreases its ACR (but not below MCR) by a quantity RDF*PCR, where RDF is the rate decrease factor. However, again ACR of the source cannot be larger than the explicit rate specified in the ER field. RDF is negotiated in the VC setup phase.
- Finally, when the source receives a backward RM cell with only the NI bit set, it sets its rate to the minimum of ACR and ER.

This set of functions performed by the source can be summarized as follows:

$$\mathrm{ACR} \leftarrow \begin{cases} \min\{\mathrm{PCR}, \min\{\mathrm{ACR} + \mathrm{RIF} \times \mathrm{PCR}, \mathrm{ER}\}\} & \text{if CI} = 0,\ \text{NI} = 0 \\ \max\{\mathrm{MCR}, \min\{\mathrm{ACR} - \mathrm{RDF} \times \mathrm{PCR}, \mathrm{ER}\}\} & \text{if CI} = 1 \\ \max\{\mathrm{MCR}, \min\{\mathrm{ACR}, \mathrm{ER}\}\} & \text{if CI} = 0,\ \text{NI} = 1 \end{cases}$$

In explicit rate control RIF $= 1$, and RDF $= 0$. This simplifies the source action to

$$\mathrm{ACR} \leftarrow \max\{\mathrm{MCR}, \min\{\mathrm{PCR}, \mathrm{ER}\}\}$$

2.2 Performance Criteria

Numerous algorithms have been and are being developed for ATM ABR congestion control. To be able to study and compare these algorithms on equal grounds, we need to set some performance measures independent of the particular algorithm under investigation. The main goal of ABR rate control is to provide fairness among all VCs with a minimal CLR and maximal utilization of network resources. The latter two of these objectives can be achieved by regulating the queue length at intermediate switches around a desirable level. Tracking such a nominal queue length (whose exact value is determined based on QoS requirements) is desirable in order to avoid losses due to overflow and waste of the buffer capacity due to underflow.

Fairness is an issue that requires more discussion, as it may be hard to visualize what is meant by a *fair* allocation in a large network with multiple nodes. The most widely accepted notion of fairness is the *max-min fairness* criterion [7]. Under this criterion, the fair share of a connection contending for a given link bandwidth should be equal to

$$\text{Fair share} = \mathrm{MCR}_m + \frac{a - \left(u + \overline{\mathrm{MCR}}\right)}{N - M^u} \qquad (1)$$

Here MCR_m is MCR of connection m, $\overline{\mathrm{MCR}}$ is the sum of MCRs of active connections, a is the total available bandwidth for ABR connections on the given link, N is the number of active connections on the link, u is the sum of bandwidth of connections bottlenecked elsewhere (including those limited by their PCR), and M^u is the number of such connections.

3 Basic Model of an ATM Switch

Generally, an ATM switch has several input and output lines. The number of input lines is almost always the same as the number of output lines, because the links are bi-directional. Cells arrive on the input lines asynchronously, but the switching is done synchronously with the help of a master clock. Each input line is connected to a common bus, through which the incoming cells are directed to their corresponding output ports. Most of the commercial ATM switches use output queueing to prevent high cell loss rates. In output queueing each output line has a finite buffer, where the incoming cells are served on a FIFO (first in, first out) basis. Associated with each output line, there is an ER controller that suggests a data rate for each ABR VC.

In what follows we focus on a particular output line of an ATM switch. Other output lines can be treated in a similar manner. The model we adopt here is a discrete-time model, where a time unit corresponds to the interval over which the rate available to ABR sources is determined (that is, the interval over which measurements are made). Further, this measurement interval is assumed to be long enough for the switch to be able to process several cells —a reasonable assumption if the link speeds are high and packet sizes are small. This allows us to ignore the cell-level dynamics and model the ABR traffic as a fluid.

Let $r(n)$ denote the total number of cells that arrive in one of the output buffers of an ATM switch in the interval $[n, n+1)$, and let $a(n)$ denote the number of cells that depart from this buffer in the same time interval. Note that $a(n)$ represents the available bandwidth (unused by higher priority traffic, particularly, CBR and VBR). Denoting the queue length at time n by $q(n)$ we have the evolution

$$q(n+1) = \min\left\{\overline{Q}, \max\left\{0, q(n) + r(n) - a(n)\right\}\right\} \tag{2}$$

where $\overline{Q}$ denotes the size of the output buffer.

Let there be a total number of $N \geq 1$ connections (sources) switched through the output line under study, and the number of cells that arrive from source m during time-slot $[n, n+1)$ be denoted by $r_m(n)$. Clearly,

$$r(n) = \sum_{m=1}^{N} r_m(n).$$

In general, $r(n)$ has two components: 1) The number of cells that arrive from *uncontrolled* sources, i.e., those sources which are bottlenecked elsewhere in the network or are limited by their PCR constraints. The ER controller at this switch has no control over these sources. In other words, the ER field of an uncontrolled source is either overwritten at some other switch, or is replaced by its PCR value at the source. We denote this component of $r(n)$ by $u(n)$. 2) The number of cells that arrive from *controlled* sources, which are bottlenecked at this switch. The ER fields of RM cells of this type of sources are modified to achieve several traffic related service guarantees. In addition, we assume that each source, controlled or uncontrolled, has negotiated a nonnegative MCR with the network, and let MCR_m denote the MCR of the controlled source m.

Let there be a total number of $M \geq 1$ controlled sources, and the number of cells that arrive from controlled source m in the interval $[n, n+1)$ be denoted by $r_m^c(n)$. Then, we have the following relation between $r(n)$, $r^c(n)$, and $u(n)$:

$$r(n) = \sum_{m=1}^{M} r_m^c(n) + u(n).$$

Note that the number of controlled sources, M, may not be known to the switch due to the bursty nature of ABR traffic, but it is reasonable to assume that this number cannot exceed a certain upper bound, $\overline{M}$. To have a feasible problem, we must guarantee that there is enough bandwidth for the controlled sources to transmit at their minimum cell rates, which requires

$$a(n) - u(n) > \sum_{m=1}^{M} \mathrm{MCR}_m = \overline{\mathrm{MCR}}. \tag{3}$$

As mentioned in the introduction, it takes time from the moment the ER decision is made by the switch until an action is taken by a source, and until subsequently that action affects the state of the switch that initiated the action. Thus, the cell rate of source m at time n, $r_m(n)$, is actually an outcome of an action taken d_m time units earlier, where d_m represents the action delay for source m and is taken to be independent of time n. Thus, we have

$$r_m^c(n) = \mathrm{MCR}_m + \mathrm{ER}(n - d_m) \tag{4}$$

where $\mathrm{ER}(n)$ denotes the action of the switch (ER controller) at time n. Without any loss of generality, we assume that the d_m's are ordered such that

$$0 \leq d_1 \leq \cdots \leq d_M \leq \overline{d}$$

where $\overline{d}$ corresponds to the maximum round-trip network delay.

The action delay, d_m, for source m, consists of several components, such as the round-trip propagation delay, the queuing and processing delays, etc. Although the end-to-end round-trip delay may be known to the switch as part of the FRTT (Fixed Round-Trip Time) computation performed at VC setup time, we can still have some error in delay estimates. One source of error is the assumption that the delay is an integer multiple of the time unit (measurement interval), which may not be true. Further, there is variability in the delays due to queues in the virtual circuits. Finally, RM cells are generated only every Nrm data cells and hence, feedback is not instantaneous. Therefore, d_m's cannot be taken to be known to the switch.

As evident from (1), any calculation of the fair share at the switch requires the knowledge of the number of controlled connections, $M = (N - M^u)$. Due to the bursty nature of ABR traffic, an accurate value of M may not be known at any time.

Motivated by these observations, we want to develop an ER control algorithm which is robust to uncertainty in action delays, and at the same time adaptive to the number of controlled connections.

We close our account on this section by rewriting the queue dynamics (2) in a more convenient form:

$$q(n+1) = \min\left\{\overline{Q}, \max\left\{0, q(n) + \overline{\mathrm{MCR}} + \sum_{k=0}^{\overline{d}} m_k \mathrm{ER}(n-k) + u(n) - a(n)\right\}\right\}$$

where m_k denotes the number of controlled sources having k units of action delay.

4 The Algorithm

In order to focus on the derivation of the robust adaptive algorithm, we assume henceforth that the bandwidth available for ABR traffic, a, and the incoming cell rate of uncontrolled connections, u, do not change with time. These assumptions are justified if $a(n)$ and $u(n)$ vary slowly compared to the time constant of the closed-loop system. As the actual number of controlled sources, M, is not known to the switch, we start our analysis by proposing a method to estimate this figure. Let us first consider the case when $u = 0$ (i.e. there are no uncontrolled connections). Our method relies on a simple observation: If the switch sends out the same command, say c, for $(\overline{d}+1)$ time units, at the end of the $(\overline{d}+1)$st step, all of the controlled sources in the system will be transmitting at rate c. Hence, at the end of the $(\overline{d}+1)$st time step, if we divide the incoming cell rate, $r(n)$, by c, we obtain the number of controlled sources on the link. In order to have a running estimate, we construct an estimator, $\hat{M}(n)$, and update it every $(\overline{d}+1)$ time units. Note that this algorithm converges to the exact value of M in only a finite number of steps. Having determined M in this fashion, the ER controller can set the desired rate to $\frac{a}{M}$ and achieve fairness, which in this particular case also corresponds to max-min fairness.

Now if u is not zero, the above scheme fails to converge to M. However, it still does converge, but to a different value. In fact, as we will show shortly, in calculating the share of each controlled connection, if one uses the number to which this algorithm converges, then the resulting bandwidth allocation is max-min fair.

Before proceeding further, we note that the choice of the rate c above is completely arbitrary, because it does not play any role in the estimation process. Thus, we are free to pick any positive rate for the algorithm to work.

Let $\hat{M}(n)$ denote the estimate of M at time n. We propose the following estimator:

$$\hat{M}(n(\overline{d}+1)) = \hat{M}(n(\overline{d}+1)+1) = \cdots = \hat{M}((n+1)(\overline{d}+1)-1) = \frac{r(n(\overline{d}+1))}{\mathrm{ER}((n-1)(\overline{d}+1))} \tag{5}$$

where $\mathrm{ER}(n)$ is kept at the same value for an interval of length $(\overline{d}+1)$. That is,

$$\mathrm{ER}(n(\overline{d}+1)) = \mathrm{ER}(n(\overline{d}+1)+1) = \cdots = \mathrm{ER}((n+1)(\overline{d}+1)-1)$$

For ease of notation, let us introduce the following subsequences:

$$q_s(n) := q(n(\overline{d}+1)),\ \mathrm{ER}_s(n) := \mathrm{ER}(n(\overline{d}+1)),\ r_s(n) := r(n(\overline{d}+1))$$

Now the estimator (5) can be written in a more convenient form using these new definitions as

$$\hat{M}_s(n) = \frac{r_s(n)}{\mathrm{ER}_s(n-1)},\ \hat{M}_s(0) = \overline{M}$$

To complete the design of the robust adaptive controller, we need to specify how $\mathrm{ER}_s(n)$ is updated to achieve the dual goal of max-min fairness and queue length stability. We propose the following design:

$$\mathrm{ER}_s(n) = \max\left\{0, \frac{a}{\hat{M}_s(n)} - \beta(q_s(n) - Q)\right\} \tag{6}$$

where Q is the target queue length, $\beta > 0$ is the controller gain to be selected to ensure stability. Here the max function is introduced to ensure that the switch asks the source to transmit at a positive rate in excess of MCR, as required by the QoS specifications. In (6), the term $-\beta(q_s(n) - Q)$ is introduced to drive the queue length, q_n, to the desired set point by providing negative feedback in the closed-loop system dynamics.

Note that if $q(n)$ converges to Q, $\mathrm{ER}_s(n)$, converges to the solution of the following equation:

$$\mathrm{ER}_s(\infty) = \max\left\{0, \frac{a}{M + \frac{u+\overline{\mathrm{MCR}}}{\mathrm{ER}_s(\infty)}}\right\}, \tag{7}$$

The solution of (7) is given by

$$\mathrm{ER}_s(\infty) = \max\left\{0, \frac{a - (u + \overline{\mathrm{MCR}})}{M}\right\} = \frac{a - (u + \overline{\mathrm{MCR}})}{M}$$

where the second equality follows from the feasibility condition (3). Hence, if we can show that $q(n)$ converges to Q, then by (4), the rate of controlled source m converges to

$$r_m^c(\infty) = \mathrm{MCR}_m + \mathrm{ER}_s(\infty) = \mathrm{MCR}_m + \frac{a - (u + \overline{\mathrm{MCR}})}{M}$$

which is the MCR plus max-min fair share of the available bandwidth.

5 Analysis of the Algorithm

In this section, we show that the robust adaptive algorithm converges to the desired set values, if the gain of the controller, β, is picked appropriately. We start our analysis by rewriting the queue and the estimator dynamics in a more explicit form.

$$\begin{aligned} q(n+1) &= \min\left\{\overline{Q}, \max\left\{0, q(n) + \overline{\mathrm{MCR}} + \sum_{k=0}^{\bar{d}} m_k \mathrm{ER}(n-k) + u - a\right\}\right\} \\ \mathrm{ER}_s(n) &= \max\left\{0, \frac{a}{M + \frac{u+\overline{\mathrm{MCR}}}{c_s(n-1)}} - \beta\left(q_s(n) - Q\right)\right\} \end{aligned}$$

To arrive at a somewhat simpler formulation, we introduce

$$x(n) := q(n) - Q,\ x_0(n) := x_s(n) := q_s(n) - Q,\ x_1(n) := \mathrm{ER}_s(n-1) - \frac{a - (u + \overline{\mathrm{MCR}})}{M},$$

and assume that the saturation nonlinearities of the buffer are not activated for $(\overline{d}+1)$ time units. This can be ensured by picking $\overline{Q}$ sufficiently large. Also let $e := u + \overline{\mathrm{MCR}}$ for ease of notation. The original system can be written in terms of the newly introduced variables as follows:

$$
\begin{aligned}
x_0(n+1) &= \min\left\{\overline{Q}-Q, \max\left\{-Q, x_0(n)+\gamma_1 x_1(n)+\gamma_0 \max\left\{-\frac{a-e}{M}, \frac{e x_1(n)}{M x_1(n)+a}-\beta x_0(n)\right\}\right\}\right\} \quad (8)\\
x_1(n+1) &= \max\left\{-\frac{a-e}{M}, \frac{e x_1(n)}{M x_1(n)+a}-\beta x_0(n)\right\} \quad (9)
\end{aligned}
$$

where

$$
\gamma_0 = \sum_{k=0}^{\overline{d}} (\overline{d}+1-k) m_k, \quad \gamma_1 = \sum_{k=0}^{\overline{d}} k m_k
$$

and γ_0 and γ_1 satisfy $\gamma_0+\gamma_1 = (\overline{d}+1)M$. To obtain these state equations one has to look at the evolution of $x(n)$ for $(\overline{d}+1)$ time units, and express $x((n+1)(\overline{d}+1)) = x_1(n+1)$ in terms of $x(n(\overline{d}+1)) = x_1(n)$ by recursive substitutions. The details of this calculation can be found in [8].

There are two types of nonlinearities in the state equations above. One is the saturation type nonlinearities, which are caused by min and max functions. These can be analyzed by considering the so-called switching curves, which are the set of points in the $x_1 - x_0$ plane for which the two arguments of the max or min function are equal. The other type of nonlinearity is of a fractional form introduced by the estimation scheme. It can be easily verified that the origin is the only equilibrium point of this system. Thus, what we have here is a two-dimensional nonlinear discrete-time system with an isolated equilibrium point at the origin. We are mainly interested in the stability of this system. In the next subsection, we first conduct a local analysis by linearizing the system around the origin, and conclude that the system is locally asymptotically stable. Then, we carry out a global analysis of the nonlinear system and provide a theorem that estimates the asymptotic region of stability in the $x_1 - x_0$ plane.

5.1 Local Analysis

Around the origin, the nonlinearities of the first type are not active. Thus the state equations take the form

$$
\begin{aligned}
x_0(n+1) &= (1-\beta\gamma_0)x_0(n) + \gamma_1 x_1(n) + \frac{\gamma_0 e x_1(n)}{M x_1(n)+a} \quad (10)\\
x_1(n+1) &= \frac{e x_1(n)}{M x_1(n)+a} \quad (11)
\end{aligned}
$$

Now if we linearize the nonlinear system described by (10)-(11) around the origin, we obtain

$$
\begin{aligned}
x_0^l(n+1) &= (1-\beta\gamma_0)x_0^l(n) + \left(\gamma_1 + \frac{\gamma_0 e}{a}\right) x_1^l(n)\\
x_1^l(n+1) &= -\beta x_0^l(n) + \frac{e}{a} x_1^l(n).
\end{aligned}
$$

The stability of this linearized system depends on the location of the roots of its characteristic equation

$$
\lambda^2 + \left(\beta\gamma_0 - 1 - \frac{e}{a}\right)\lambda + \left(\beta\gamma_1 + \frac{e}{a}\right) = 0, \quad (12)
$$

and by the Schur-Cohn criterion [9], the roots of equation (12) are inside the unit circle if and only if

$$\left|\beta\gamma_1 + \frac{e}{a}\right| < 1, \quad \left|\beta\gamma_0 - 1 - \frac{e}{a}\right| < 1 + \beta\gamma_1 + \frac{e}{a}. \tag{13}$$

Note that $\frac{e}{a} < 1$, as the uncontrolled sources use only a fraction of the total available bandwidth, a. To obtain a bound on β which is independent of network parameters, it is sufficient to consider the case $(\beta\gamma_0 - 1 - \frac{e}{a}) < 0$. This choice of β is equivalent to $\beta < \frac{1+\frac{e}{a}}{\gamma_0}$, as $\gamma_0 > 0$. In addition, the second inequality of (13) implies $\beta > 0$, which, along with the first inequality of (13), further implies $\beta < \frac{1-\frac{e}{a}}{\gamma_1}$. Combining these three inequalities on β, we obtain the stability condition

$$0 < \beta < \min\left\{\frac{1+\frac{e}{a}}{\gamma_0}, \frac{1-\frac{e}{a}}{\gamma_1}\right\}$$

Note that we can bound each term on the right-hand side of this inequality from below by

$$\frac{1+\frac{e}{a}}{\gamma_0+\gamma_1} \le \frac{1+\frac{e}{a}}{\gamma_0}, \quad \frac{1-\frac{e}{a}}{\gamma_0+\gamma_1} < \frac{1-\frac{e}{a}}{\gamma_1}$$

where the first inequality is tight, since γ_1 may achieve the value 0. Thus, also using the fact that $\min(a,b) = \frac{a+b}{2} - \frac{|a-b|}{2}$, we conclude that if β is picked such that

$$0 < \beta < \frac{1}{(\overline{d}+1)M}\left(1 - \frac{e}{a}\right) < \frac{1}{(\overline{d}+1)\overline{M}}\left(1 - \frac{u+\overline{MCR}}{a}\right), \tag{14}$$

then the nonlinear system (8)-(9) is locally asymptotically stable.

5.2 Global Analysis

Having determined local asymptotic stability, we turn our attention to the original system with saturation type nonlinearities. For convenience, we recapitulate here the state equations (8)-(9)

$$\begin{aligned} x_0(n+1) &= \min\left\{\overline{Q} - Q, \max\left\{-Q, x_0(n) + \gamma_1 x_1(n) + \gamma_0 \max\left\{-\frac{a-e}{M}, \frac{e x_1(n)}{M x_1(n) + a} - \beta x_0(n)\right\}\right\}\right\} \\ x_1(n+1) &= \max\left\{-\frac{a-e}{M}, \frac{e x_1(n)}{M x_1(n) + a} - \beta x_0(n)\right\} \end{aligned}$$

A *trajectory* of this system is defined to be the set of points in the $x_1 - x_0$ plane the state equations traverse starting at an initial point $x(0) = (x_1(0), x_0(0))$. And a trajectory of (8)-(9) is said to be stable, if it converges to the origin.

We first observe that due to the max constraint x_1 can be no smaller than $-\frac{a-e}{M}$, and x_0 should lie between $-Q$ and $\overline{Q} - Q$ because of the max and min constraints in (9). Thus, the part of $x_1 - x_0$ plane that is of interest to us is limited to the semi-infinite rectangular region, R, as depicted in Figure 2 below.

Note that any trajectory of (8)-(9) with an initial condition outside R will enter R in only one time step. Thus, from here on we will assume that $x(0) \in R$. In order to further see the effect of the saturation nonlinearities, we invoke a switching curve analysis in which the region R is partitioned into smaller regions by the so-called switching curves, which are formed by the set of points in R for which the two arguments of a min or max function are equal. We first consider the max function in (9). Setting the two arguments of this function equal to each other, we obtain the curve

$$x_0 = \frac{1}{\beta}\left(\frac{e x_1}{M x_1 + a} + \frac{a-e}{M}\right)$$

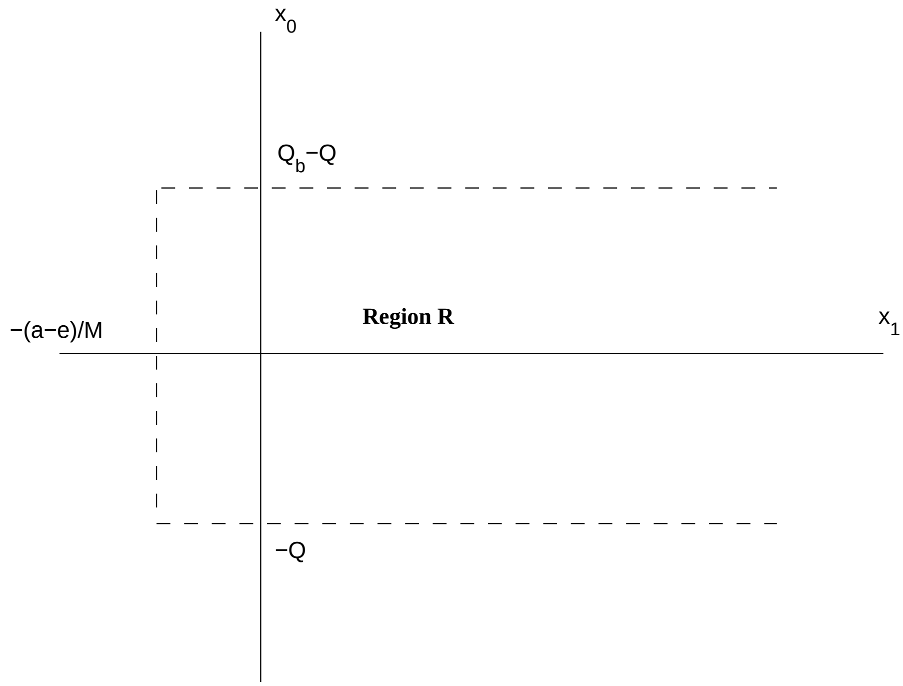

Figure 2: Region R

which we shall call Switching Curve 1, or SC1 for short. SC1 intersects the line $x_1 = -\frac{a-e}{M}$ at $x_0 = 0$, and it intersects the x_0 axis at $\frac{1}{\beta}\left(\frac{a-e}{M}\right)$. Now, if we let

$$\beta < \frac{a-e}{M\left(\overline{Q} - Q\right)}, \tag{15}$$

then SC1 intersects the line $x_0 = \overline{Q} - Q$ at a point $x_1 < 0$; see Figure 3.

Depending on what part of SC1 the states are, the state equations (8)-(9) take two different forms. We now investigate these as two separate cases.

Case 1 $x_0 \geq \frac{1}{\beta}\left(\frac{ex_1}{Mx_1+a} + \frac{a-e}{M}\right)$: In this case, the second state equation (9) reduces to

$$x_1(n+1) = -\frac{a-e}{M}, \tag{16}$$

and the first state equation (8) becomes

$$x_0(n+1) = \min\left\{\overline{Q} - Q, \max\left\{-Q, x_0(n) + \gamma_1 x_1(n) - \gamma_0 \frac{a-e}{M}\right\}\right\} \tag{17}$$

The line

$$x_0 = -\gamma_1 x_1 + \gamma_0 \frac{a-e}{M} - Q,$$

which is obtained by setting the two arguments of the max function equal to each other, lies completely below SC1 for all $x_1 > -\frac{a-e}{M}$ if Q is picked such that

$$Q > \gamma_0 \frac{a-e}{M}.$$

Thus, the max function in (17) always returns its second argument, resulting in the following equation for (9):

$$x_0(n+1) = \min\left\{\overline{Q} - Q, x_0(n) + \gamma_1 x_1(n) - \gamma_0 \frac{a-e}{M}\right\} \tag{18}$$

Now the switching line for the min nonlinearity in (18) is

$$x_0 = -\gamma_1 x_1 + \gamma_0 \frac{a-e}{M} - Q + \overline{Q}$$

This time, the line is completely above SC1 for all $x_1 > -\frac{a-e}{M}$ if Q is chosen as

$$Q > (\gamma_0 + \gamma_1)\left(\frac{a-e}{M}\right) \tag{19}$$

As a result, the min function in (18) always returns its second argument, and the state equation (9) simplifies to

$$x_0(n+1) = x_0(n) + \gamma_1 x_1(n) - \frac{a-e}{M}. \tag{20}$$

We denote the region in which the foregoing analysis is valid as R_1; see Figure 3. Note that the state equations (8)-(9) can be written as (20)-(16) in R_1.

Case 2 $x_0 < \frac{1}{\beta}\left(\frac{ex_1}{Mx_1+a} + \frac{a-e}{M}\right)$: In this case, the state equations (8)-(9) become

$$\begin{aligned} x_0(n+1) &= \min\left\{\overline{Q} - Q, \max\left\{-Q, x_0(n) + \gamma_1 x_1(n) + \gamma_0\left(\frac{ex_1(n)}{Mx_1(n)+a} - \beta x_0(n)\right)\right\}\right\} && (21) \\ x_1(n+1) &= \frac{ex_1(n)}{Mx_1(n)+a} - \beta x_0(n) && (22) \end{aligned}$$

The switching curve for the max nonlinearity in (21) is given by

$$x_0 = -\frac{1}{1-\beta\gamma_0}\left(\frac{\gamma_0 ex_1}{Mx_1+a} + \gamma_1 x_1 + Q\right)$$

Let us call this curve SC2. This curve intersects $x_1 = -\frac{a-e}{M}$ line at $x_0 = \frac{1}{a-\beta\gamma_0}\left((\gamma_0+\gamma_1)\left(\frac{a-e}{M}\right) - Q\right) > -Q$, see Figure 3. Unlike Case 1, however, depending on what side of SC2 the states are, we have two sub-cases to investigate.

Case 2a $x_0 \geq -\frac{1}{1-\beta\gamma_0}\left(\frac{\gamma_0 ex_1}{Mx_1+a} + \gamma_1 x_1 + Q\right)$: In this case, the max function in (21) yields its second argument resulting in the following equation for $x_0(n)$:

$$x_0(n+1) = \min\left\{\overline{Q} - Q, x_0(n) + \gamma_1 x_1(n) + \gamma_0\left(\frac{ex_1(n)}{Mx_1(n)+a} - \beta x_0(n)\right)\right\} \tag{23}$$

Now, comparing the arguments of the min function, we obtain the switching curve SC3

$$x_0 = -\frac{1}{1-\beta\gamma_0}\left(\frac{\gamma_0 ex_1}{Mx_1+a} + \gamma_1 x_1 + Q - \overline{Q}\right)$$

It can be seen that if

$$x_0 < -\frac{1}{1-\beta\gamma_0}\left(\frac{\gamma_0 ex_1}{Mx_1+a} + \gamma_1 x_1 + Q - \overline{Q}\right),$$

then (21) can be simplified to

$$x_0(n+1) = x_0(n) + \gamma_1 x_1(n) + \gamma_0 \left(\frac{e x_1(n)}{M x_1(n) + a} - \beta x_0(n) \right) \tag{24}$$

The region corresponding to this set of inequalities is denoted by R_3 in Figure 3. If, on the other hand,

$$x_0 \geq -\frac{1}{1-\beta\gamma_0} \left(\frac{\gamma_0 e x_1}{M x_1 + a} + \gamma_1 x_1 + Q - \overline{Q} \right),$$

(21) reduces to

$$x_1(n+1) = \overline{Q} - Q \tag{25}$$

with the corresponding region R_4 shown in Figure 3.

Case 2b $x_0 < -\frac{1}{1-\beta\gamma_0} \left(\frac{\gamma_0 e x_1}{M x_1 + a} + \gamma_1 x_1 + Q \right)$: In this case, the equation for $x_0(n)$ becomes

$$x_0(n+1) = \min \left\{ \overline{Q} - Q, -Q \right\} = -Q. \tag{26}$$

The region of the $x_1 - x_0$ plane for which (26) is valid is indicated by R_2 in Figure 3.

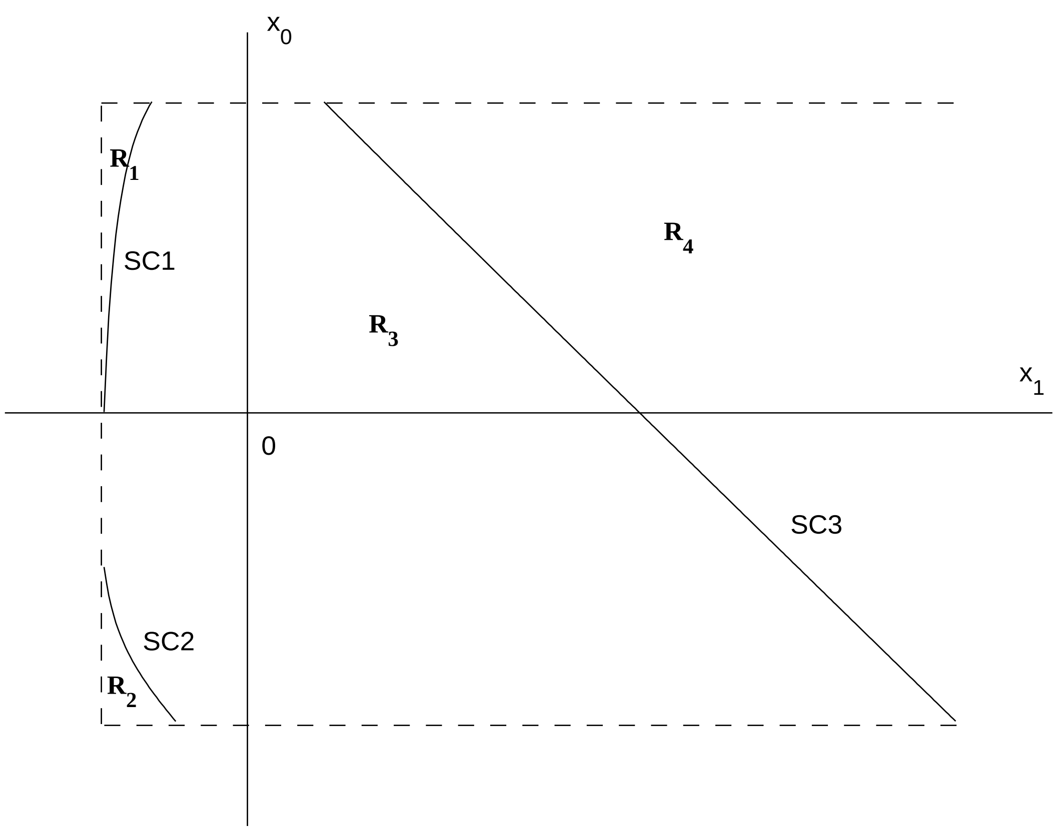

Figure 3: Switching curves and different regions of R

In summary, the semi-infinite rectangular region, R, of the $x_1 - x_0$ plane, can be partitioned into four regions provided that Q and β are picked such that

$$\begin{aligned} (\gamma_0 + \gamma_1)\left(\frac{a-e}{M}\right) < \quad & Q \quad < \overline{Q} & (27) \\ 0 < \quad & \beta \quad < \frac{a-e}{M}\frac{1}{\overline{Q}-Q} & (28) \end{aligned}$$

The region boundaries along with the valid state equations in each region are given below for convenience.

- Region 1 (R1)

  Boundaries:

$$\begin{aligned} x_0 &< \overline{Q} - Q \\ x_0 &\geq \frac{1}{\beta}\left(\frac{ex_1}{Mx_1 + a} + \frac{a-e}{M}\right) \\ x_1 &\geq -\frac{a-e}{M} \end{aligned}$$

  State Equations:

$$\begin{aligned} x_0(n+1) &= x_0(n) + \gamma_1 x_1(n) - \frac{a-e}{M} \\ x_1(n+1) &= -\frac{a-e}{M} \end{aligned}$$

- Region 2 (R2)

  Boundaries:

$$\begin{aligned} x_0 &< \frac{1}{1-\beta\gamma_0}\left(\frac{\gamma_0 e x_1}{Mx_1 + a} + \gamma_1 x_1 + Q\right) \\ x_0 &\geq -Q \\ x_1 &\geq -\frac{a-e}{M} \end{aligned}$$

  State Equations:

$$\begin{aligned} x_0(n+1) &= -Q \\ x_1(n+1) &= \frac{ex_1(n)}{Mx_1(n) + a} - \beta x_0(n) \end{aligned}$$

- Region 3 (R3)

  Boundaries:

$$\begin{aligned} x_0 &< -\frac{1}{1-\beta\gamma_0}\left(\frac{\gamma_0 e x_1}{Mx_1 + a} + \gamma_1 x_1 + Q - \overline{Q}\right) \\ x_0 &< \overline{Q} - Q \\ x_0 &< \frac{1}{\beta}\left(\frac{ex_1}{Mx_1 + a} + \frac{a-e}{M}\right) \\ x_0 &\geq -Q \\ x_0 &\geq \frac{1}{1-\beta\gamma_0}\left(\frac{\gamma_0 e x_1}{Mx_1 + a} + \gamma_1 x_1 + Q\right) \\ x_1 &\geq -\frac{a-e}{M} \end{aligned}$$

State Equations:

$$\begin{aligned} x_0(n+1) &= (1-\beta\gamma_0)\, x_0(n) + \gamma_1 x_1(n) + \frac{\gamma_0 e x_1(n)}{M x_1(n) + a} \\ x_1(n+1) &= \frac{e x_1(n)}{M x_1(n) + a} - \beta x_0(n) \end{aligned}$$

- Region 4 (R4)

  Boundaries:

$$\begin{aligned} x_0 &< \overline{Q} - Q \\ x_0 &\geq -\frac{1}{1-\beta\gamma_0}\left(\frac{\gamma_0 e x_1}{M x_1 + a} + \gamma_1 x_1 + Q - \overline{Q}\right) \end{aligned}$$

  State Equations:

$$\begin{aligned} x_0(n+1) &= \overline{Q} - Q \\ x_1(n+1) &= \frac{e x_1(n)}{M x_1(n) + a} - \beta x_0(n) \end{aligned}$$

We first look at the nonlinear system (8)-(9) in a subset $\mathcal{D}$ of region R_3. Instead of the state equations given earlier for R_3, let us consider the following auxiliary (nominal) system $x(n+1) = f(x(n))$ given by

$$x_0(n+1) = (1-\beta\gamma_0)x_0(n) + \gamma_1 x_1(n) + \frac{\gamma_0 e x_1(n)}{M x_1(n) + a} \tag{29}$$

$$x_1(n+1) = \frac{e x_1(n)}{M x_1(n) + a} \tag{30}$$

where $f : \mathcal{D} \to \mathcal{R}^2$. The original system in $\mathcal{D}$ can be thought of as a perturbation of this nominal system by a function $g(x)$ in the form $g(x) = [0 \ -\beta x_0]^T$, i.e. $x(n+1) = f(x(n)) + g(x(n)) := f_p(x(n))$. We will first show that the nominal system is exponentially stable on $\mathcal{D}$. This fact will enable us to find a Lyapunov function $V(x)$ for this system. Next, we will show that this Lyapunov function can be used for the perturbed system, $f_p(x(n))$, and we will conclude that the original system is exponentially stable in a predictable subset of $\mathcal{D}$ provided that β is sufficiently small. We start our analysis with three Lemmas.

Lemma 1: *Let $\mathcal{D} = \{x \in \mathcal{R}^2 : |x_0| < r_0,\ |x_1| < r_1\} \cap R_3$ for some $0 < r_0 < \infty$, and $0 < r_1 < \frac{a-e}{M}$. Suppose $x(0) \in \mathcal{D}$. Then the nominal system described by (29)-(30) has a unique equilibrium point at the origin.*

Proof: Let $f(x) = [f_0(x) \ f_1(x)]^T$. The function $f_1(x) = \frac{e x_1}{M x_1 + a}$ is a contraction on $|x_1| < r_1$. Therefore, $|x_1| < r_1 \Rightarrow |f_1(x_1)| < |x_1| < r_1$. Thus, if $|x_1(0)| < r_1$, then $|x_1(n)| < r_1$, $\forall n$. We set $x_0(n+1) = x_0(n) = x_0^e$, and $x_1(n+1) = x_1(n) = x_1^e$ to find all equilibrium points of the nonlinear system $x(n+1) = f(x(n))$. This yields

$$\begin{aligned} x_0^e &= (1-\beta\gamma_0)x_0^e + \gamma_1 x_1^e + \frac{\gamma_0 e x_1^e}{M x_1^e + a} \\ x_1^e &= \frac{e x_1^e}{M x_1^e + a} \end{aligned}$$

The second equation has two solutions $x_1^e = 0$ and $x_1^e = \frac{a-e}{M}$. However, since $r_1 < \frac{a-e}{M}$, and $|x_1(n)| < r_1,\ \forall n$, we must have $x_1^e < \frac{a-e}{M}$. Thus, $x_1^e = \frac{a-e}{M}$ cannot be a solution. This leaves us with the unique solution $x_1^e = 0$ for the second equation. Substituting this into the first equation, we obtain

$$x_0^e = (1-\beta\gamma_0)x_0^e \Rightarrow x_0^e = 0$$

Hence, the system (29)-(30) has a unique equilibrium point, $(0,0)$, on $\mathcal{D}$. $\diamond$

Lemma 2: *Let* $\mathcal{D} = \left\{x \in \mathcal{R}^2 : |x_0| < r_0,\ |x_1| < r_1\right\} \cap R_3$ *for some* $0 < r_0 < \infty$, *and* $0 < r_1 < \frac{a-e}{M}$. *Suppose* β *is picked such that*

$$0 < \beta < \min\left\{\frac{a-e}{a}\frac{1}{\gamma_0}, \frac{a-e}{M}\frac{1}{\overline{Q}-Q}\right\} \tag{31}$$

where Q *satisfies (27). Then, the nominal system described by equations (29)-(30) is exponentially stable on* $\mathcal{D}$.

Proof: Equation (30) can be explicitly solved for $|x_1(0)| < r_1$ to give

$$x_1(n) = \frac{x_1(0)(1-\theta)\theta^n}{\pi x_1(0)(1-\theta^n) + (1-\theta)}$$

where $\theta = \frac{e}{a} < 1$, and $\pi = \frac{M}{a}$. Using this expression, we can bound $|x_1(n)|$ from above as

$$|x_1(n)| \le \begin{cases} |x_1(0)|\theta^n & \text{if } x_1(0) \ge 0 \\ \frac{|x_1(0)|(1-\theta)\theta^n}{1-\theta-\pi r_1} & \text{if } x_1(0) < 0 \end{cases}$$

Thus,

$$|x_1(n)| \le \max\left\{1, \frac{1-\theta}{1-\theta-\pi r_1}\right\}|x_1(0)|\theta^n$$

Since the second term of the max function is always larger than the first one, we have

$$|x_1(n)| \le \frac{1-\theta}{1-\theta-\pi r_1}|x_1(0)|\theta^n,\ \forall x(0) \in \mathcal{D}$$

Now that we have an explicit solution for $x_1(n)$, we can substitute this into equation (29) and obtain the solution for $x_0(n)$. Let $\alpha = 1-\beta\gamma_0$, and pick β such that

$$0 < \beta < \min\left\{\frac{1}{\gamma_0}, \frac{a-e}{M}\frac{1}{\overline{Q}-Q}\right\} \tag{32}$$

Here the min function is introduced to make the choice of β in compliance with our previous bound (28). Also let $s(n)$ denote the sequence

$$s(n) = \gamma_1 x_1(n) + \frac{\gamma_0 e x_1(n)}{M x_1(n) + a},$$

which goes asymptotically to zero as $n \to \infty$, since $x_1(n) \to 0$. With these definitions, the difference equation for $x_0(n)$ becomes

$$x_0(n+1) = \alpha x_0(n) + s(n),$$

which admits the solution

$$x_0(n) = \alpha^n x_0(0) + \sum_{k=0}^{n-1} \alpha^{n-1-k} s(k)$$

Clearly, $|s(k)| \leq \gamma_1|x_1(k)| + \gamma_0|x_1(k+1)|$, since both γ_0 and γ_1 are nonnegative. Now, $x_0(n)$ can be bounded from above as follows:

$$\begin{aligned}
|x_0(n)| &\leq |x_0(0)|\alpha^n + (\gamma_1 + \theta\gamma_0)\frac{1-\theta}{1-\theta-\pi r_1}\alpha^{n-1}\sum_{k=0}^{n-1}\left(\frac{\theta}{\alpha}\right)^k |x_1(0)| \\
&\leq |x_0(0)|\alpha^n + (\gamma_1 + \theta\gamma_0)\frac{1-\theta}{1-\theta-\pi r_1}\frac{\alpha^n - \theta^n}{\alpha - \theta}|x_1(0)|
\end{aligned}$$

where we require $\frac{\theta}{\alpha} < 1$, which is guaranteed with the following choice of β:

$$0 < \beta < \frac{a-e}{a}\frac{1}{\gamma_0}$$

Combining this with (32), and the fact that

$$\min\left\{\frac{1}{\gamma_0}, \frac{a-e}{a}\frac{1}{\gamma_0}\right\} = \frac{a-e}{a}\frac{1}{\gamma_0},$$

we obtain the bound

$$0 < \beta < \min\left\{\frac{a-e}{a}\frac{1}{\gamma_0}, \frac{a-e}{M}\frac{1}{\overline{Q}-Q}\right\}$$

Hence, we conclude that if β is picked according to (31), we have

$$\begin{aligned}
|x_0(n)| &\leq |x_0(0)|\alpha^n + (\gamma_1 + \theta\gamma_0)\frac{1-\theta}{1-\theta-\pi r_1}\frac{1}{\alpha-\theta}|x_1(0)|\alpha^n - (\gamma_1 + \theta\gamma_0)\frac{1-\theta}{1-\theta-\pi r_1}\frac{1}{\alpha-\theta}|x_1(0)|\theta^n \\
|x_1(n)| &\leq \frac{1-\theta}{1-\theta-\pi r_1}|x_1(0)|\theta^n
\end{aligned}$$

for all $x(0) \in \mathcal{D}$. $\diamond$

Lemma 3: *Let* $\mathcal{D} = \left\{x \in \mathcal{R}^2 : |x_0| < r_0,\ |x_1| < r_1\right\} \cap R_3$ *for some* $0 < r_0 < \infty$*, and* $0 < r_1 < \frac{a-e}{M}$. *Then,* $|x_1(n)|$ *can be lower bounded on* $\mathcal{D}$ *so that*

$$|x_1(n)| \geq \frac{1-\theta}{1-\theta+\pi r_1}|x_1(0)|\theta^n$$

Proof: We use the explicit solution of $x_1(n)$ to obtain

$$|x_1(n)| \geq \begin{cases} \frac{|x_1(0)|(1-\theta)\theta^n}{1-\theta+\pi r_1} & \text{if } x_1(0) \geq 0 \\ |x_1(0)|\theta^n & \text{if } x_1(0) < 0 \end{cases}$$

Thus,

$$\begin{aligned}
|x_1(n)| &\geq \min\left\{1, \frac{1-\theta}{1-\theta+\pi r_1}\right\}|x_1(0)|\theta^n \\
\Rightarrow |x_1(n)| &\geq \frac{1-\theta}{1-\theta+\pi r_1}|x_1(0)|\theta^n,\ \forall x(0) \in \mathcal{D}
\end{aligned}$$

and hence the result. $\diamond$

Next, using these three Lemmas we show that there exists a Lyapunov function $V(x)$ on $\mathcal{D}$ for the nominal system (29)-(30) satisfying some norm bounds. We state this as a theorem.

Theorem 1: *Let $\mathcal{D} = \left\{x \in \mathcal{R}^2 : |x_0| < r_0,\ |x_1| < r_1\right\} \cap R_3$ for some $0 < r_0 < \infty$, and $0 < r_1 < \frac{a-e}{M}$. Suppose β is picked as in (31). Then there exists a Lyapunov function $V : \mathcal{D} \to \mathcal{R}$ for the system (29)-(30) that satisfies the inequalities:*

$$C_{11}|x_1| \leq V(x) \leq C_2\|x\|_1 \tag{33}$$

$$\Delta V(x) := V(f(x)) - V(x) \leq -D_3\|x\|_1 \tag{34}$$

for some positive constants C_{11}, C_2, and D_3, where $\|\cdot\|$ denotes the l_1 norm.

Proof: Let $\phi(n, x)$ denote the solution of the system at time n starting at x at time zero; that is, $\phi(0, x) = x$. Let

$$V(x) = \sum_{n=0}^{N-1} \|\phi(n,x)\|_1 = \sum_{n=0}^{N-1} \left[|\phi_0(n,x)| + |\phi_1(n,x)|\right] \tag{35}$$

where N is a positive integer to be chosen later. Due to the exponentially decaying bounds on the trajectories as given in Lemma 2, we have

$$\begin{aligned}
V(x) &= \sum_{n=0}^{N-1} \left[|\phi_0(n,x)| + |\phi_1(n,x)|\right] \\
&\leq \sum_{n=0}^{N-1} \left[|x_0|\alpha^n + (\gamma_1 + \theta\gamma_0)\frac{1-\theta}{1-\theta-\pi r_1}\frac{1}{\alpha-\theta}|x_1|(\alpha^n - \theta^n) + \frac{1-\theta}{1-\theta-\pi r_1}|x_1|\theta^n\right] \\
&= \frac{1-\alpha^N}{1-\alpha}|x_0| + \frac{1-\theta}{1-\theta-\pi r_1}\left[(\gamma_1+\theta\gamma_0)\frac{1}{\alpha-\theta}\left(\frac{1-\alpha^N}{1-\alpha} - \frac{1-\theta^N}{1-\theta}\right) + \frac{1-\theta^N}{1-\theta}\right]|x_1| \\
&= C_{20}|x_0| + C_{21}|x_1| \\
&\leq \max\left\{C_{20}, C_{21}\right\}(|x_0| + |x_1|) \\
&= C_2\|x\|_1
\end{aligned}$$

where

$$\begin{aligned}
C_{20} &= \frac{1-\alpha^N}{1-\alpha} \\
C_{21} &= \frac{1-\theta}{1-\theta-\pi r_1}\left[(\gamma_1+\theta\gamma_0)\frac{1}{\alpha-\theta}\left(\frac{1-\alpha^N}{1-\alpha} - \frac{1-\theta^N}{1-\theta}\right) + \frac{1-\theta^N}{1-\theta}\right]
\end{aligned}$$

and

$$C_2 = \max\left\{C_{20}, C_{21}\right\} \tag{36}$$

Similarly, due to the exponentially decaying lower bound on the trajectory of x_1 by Lemma 3, we have

$$\begin{aligned} V(x) &= \sum_{n=0}^{N-1} [|\phi_0(n,x)| + |\phi_1(n,x)|] \\ &\geq \sum_{n=0}^{N-1} |\phi_1(n,x)| \\ &\geq \sum_{n=0}^{N-1} \frac{1-\theta}{1-\theta+\pi r_1} |x_1| \theta^n \\ &= \frac{1-\theta}{1-\theta+\pi r_1} \frac{1-\theta^N}{1-\theta} |x_1| \\ &= C_{11}|x_1| \end{aligned}$$

where

$$C_{11} = \frac{1-\theta}{1-\theta+\pi r_1} \frac{1-\theta^N}{1-\theta} \tag{37}$$

Next, we look at $\Delta V(x) = V(f(x)) - V(x)$:

$$\begin{aligned} \Delta V(x) &= \sum_{n=0}^{N-1} \|\phi(n, f(x))\|_1 - \sum_{n=0}^{N-1} \|\phi(n,x)\|_1 \\ &= \sum_{n=0}^{N-1} \|\phi(n+1, x)\|_1 - \sum_{n=0}^{N-1} \|\phi(n,x)\|_1 \\ &= \|\phi(N,x)\|_1 - \|\phi(0,x)\|_1 \\ &= \|\phi(N,x)\|_1 - \|x\|_1 \end{aligned}$$

Using Lemma 2 we obtain

$$\begin{aligned} \Delta V(x) &\leq |x_0|\alpha^N + (\gamma_1 + \theta\gamma_0)\frac{1-\theta}{1-\theta-\pi r_1}\frac{1}{\alpha-\theta}|x_1|\left(\alpha^N - \theta^N\right) + \frac{1-\theta}{1-\theta-\pi r_1}|x_1|\theta^N - |x_0| - |x_1| \\ &\leq -\left[1-\alpha^N\right]|x_0| - \left[1 - \frac{1-\theta}{1-\theta-\pi r_1}\left[(\gamma_1+\theta\gamma_0)\frac{1}{\alpha-\theta}\left(\alpha^N-\theta^N\right)+\theta^N\right]\right]|x_1| \\ &= -D_{30}|x_0| - D_{31}|x_1| \\ &\leq -\min\{D_{30}, D_{31}\}(|x_0| + |x_1|) \\ &= -D_3\|x\|_1 \end{aligned}$$

where

$$\begin{aligned} D_{30} &= 1-\alpha^N \\ D_{31} &= 1 - \frac{1-\theta}{1-\theta-\pi r_1}\left[(\gamma_1+\theta\gamma_0)\frac{1}{\alpha-\theta}\left(\alpha^N-\theta^N\right)+\theta^N\right] \end{aligned}$$

and

$$D_3 = \min\{D_{30}, D_{31}\} \tag{38}$$

Note that, both C_{11} and C_2 are positive $\forall N \geq 1$. On the other hand, even though $D_{30} > 0$ for all $N \geq 1$, D_{31} may not be positive for N small. However, for N sufficiently large, D_{31} becomes positive as well, yielding

$D_3 > 0$. Hence, we pick N large enough to make D_{31} positive. In other words, we require

$$\begin{aligned}
\frac{1-\theta}{1-\theta-\pi r_1}\left[(\gamma_1+\theta\gamma_0)\frac{1}{\alpha-\theta}\left(\alpha^N-\theta^N\right)+\theta^N\right] &< 1 \\
\Leftrightarrow (\gamma_1+\theta\gamma_0)\frac{1}{\alpha-\theta}\left(\alpha^N-\theta^N\right)+\theta^N &< \frac{1-\theta-\pi r_1}{1-\theta} \\
\Leftrightarrow (\gamma_1+\theta\gamma_0)\frac{1}{\alpha-\theta}\alpha^N+\left[1-(\gamma_1+\theta\gamma_0)\frac{1}{\alpha-\theta}\right]\theta^N &< \frac{1-\theta-\pi r_1}{1-\theta}
\end{aligned}$$

Let $A=\frac{\gamma_1+\theta\gamma_0}{\alpha-\theta}$, and $B=\frac{1-\theta-\pi r_1}{1-\theta}$. Note that $B<1$. Then, we have

$$A\alpha^N+(1-A)\theta^N<B$$

On the other hand,

$$A\alpha^N+(1-A)\theta^N\le A\left(\max\{\alpha,\theta\}\right)^N+(1-A)\left(\max\{\alpha,\theta\}\right)^N=A\alpha^N+(1-A)\alpha^N=\alpha^N$$

since $\frac{\theta}{\alpha}<1$. Thus, if $\alpha^N<B$, then $A\alpha^N+(1-A)\theta^N<B$, which implies $D_{31}>0$. Therefore, we choose N such that

$$\alpha^N<B\Rightarrow\frac{1}{B}<\left(\frac{1}{\alpha}\right)^N\Rightarrow N>\frac{\ln B}{\ln\alpha}$$

In terms of the original variables this inequality becomes

$$N>\frac{\ln\left(\frac{1-\theta-\pi r_1}{1-\theta}\right)}{\ln\left(1-\beta\gamma_0\right)} \tag{39}$$

Hence, the function defined by

$$V(x)=\sum_{n=0}^{N-1}\|\phi(n,x)\|_1$$

with N satisfying (39) is indeed a Lyapunov function for the nonlinear system (29)-(30), and it satisfies (33)-(34) with positive constants C_{11}, C_2, and D_3 given by (37), (36), and (38), respectively. $\diamond$

Now, consider the perturbed system $x(n+1)=f(x(n))+g(x(n))$, where $g(x)=[0\ \ -\beta x_0]^T$. We want to use the Lyapunov function for the nominal system as a Lyapunov function candidate for the perturbed system, and show that for sufficiently small perturbations the system remains asymptotically stable in some predictable region around the origin. We start our analysis by noting that $\|g(x)\|_1=\beta|x_0|$. Thus, we are dealing with a linear perturbation vanishing at the origin. The next Lemma shows that it is possible to use the Lyapunov function we constructed for the nominal system as a Lyapunov function for the perturbed system provided that β is sufficiently small.

Lemma 4: *Let $\mathcal{D}=\left\{x\in\mathcal{R}^2:|x_0|<r_0,\ |x_1|<r_1\right\}\cap R_3$ for some $0<r_0<\infty$, and $0<r_1<\frac{a-e}{M}$. Suppose β is picked such that*

$$0<\beta<\min\left\{\frac{a-e}{a}\frac{1}{\gamma_0},\frac{a-e}{M}\frac{1}{\overline{Q}-Q},\frac{1-\alpha^N}{\left(\frac{1-\theta}{1-\theta-\pi r_1}\right)^2\left[\frac{1-\theta^N}{1-\theta}+\frac{1}{\alpha-\theta}\left(\gamma_1+\frac{\gamma_0 ea}{(a-Mr_1)^2}\right)\left(\frac{1-\alpha^N}{1-\alpha}-\frac{1-\theta^N}{1-\theta}\right)\right]}\right\} \tag{40}$$

where N is as in (39), and Q is as in (27). Then, there exists a Lyapunov function $V : \mathcal{D} \to \mathcal{R}$ for the perturbed system $x(n+1) = f_p(x(n))$ that satisfies the inequalities:

$$C_{11}|x_1| \le V(x) \le C_2\|x\|_1 \tag{41}$$

$$\Delta V_p(x) = V(f_p(x)) - V(x) \le -C_3\|x\|_1 \tag{42}$$

for some positive constants $C_{11}, C_2,$ and C_3.

Proof: Let $V(x)$ be given as in (35). In Theorem 1, we have established that

$$C_{11}|x_1| \le V(x) \le C_2\|x\|_1$$

where C_{11} and C_2 are given by (37) and (36), respectively. Now, let us consider $\Delta V_p(x) = V(f_p(x)) - V(x)$. We have

$$\begin{aligned}
\Delta V_p(x) = V(f(x)+g(x)) - V(x) &= \sum_{n=0}^{N-1} \|\phi(n,(f(x)+g(x))\|_1 - \sum_{n=0}^{N-1} \|\phi(n,x)\|_1 \\
&= \sum_{n=0}^{N-1} \left\| \phi(n,f(x)) + \left.\frac{\partial \phi(n,x)}{\partial x}\right|_{x=z} g(x) \right\|_1 - \sum_{n=0}^{N-1} \|\phi(n,x)\|_1 \\
&\le \sum_{n=0}^{N-1} \|\phi(n,f(x))\|_1 - \sum_{n=0}^{N-1} \|\phi(n,x)\|_1 + \sum_{n=0}^{N-1} \left\| \left.\frac{\partial \phi(n,x)}{\partial x}\right|_{x=z} g(x) \right\|_1 \\
&\le \Delta V(x) + \sum_{n=0}^{N-1} \left\| \left.\frac{\partial \phi(n,x)}{\partial x}\right|_{x=z} \right\|_1 \|g(x)\|_1 \\
&= \Delta V(x) + \beta \sum_{n=0}^{N-1} \left\| \left.\frac{\partial \phi(n,x)}{\partial x}\right|_{x=z} \right\|_1 |x_0|
\end{aligned} \tag{43}$$

for some $z \in \mathcal{D}$. The second line follows from the intermediate value theorem, since $\phi(n,x)$ is differentiable on the open set $\mathcal{D} \subset \mathcal{R}^2$. We proceed with the calculation of $\left\|\frac{\partial \phi(n,x)}{\partial x}\right\|_1$, and show that

$$\left\|\frac{\partial \phi(n,x)}{\partial x}\right\|_1 \le L, \ \forall x \in \mathcal{D}$$

for some $L > 0$. We have

$$\begin{aligned}
\phi_0(x_0,x_1,n) &= \alpha^n x_0 + \sum_{k=0}^{n-1} \alpha^{n-1-k} s(k) \\
\phi_1(x_0,x_1,n) &= \frac{x_1(1-\theta)\theta^n}{\pi x_1(1-\theta^n) + (1-\theta)}
\end{aligned}$$

where

$$s(k) = \gamma_1 y(k) + \frac{\gamma_0 e y(k)}{M y(k) + a}$$

and

$$y(k) = \frac{x_1(1-\theta)\theta^k}{\pi x_1(1-\theta^k) + (1-\theta)}$$

Using these expressions, we calculate the partials of ϕ_0 and ϕ_1 on $\mathcal{D}$ as

$$
\begin{aligned}
\frac{\partial \phi_0}{\partial x_0} &= \alpha^n \\
\frac{\partial \phi_0}{\partial x_1} &= \sum_{k=0}^{n-1} \alpha^{n-1-k}\theta^k \left(\frac{1-\theta}{\pi x_1(1-\theta^k)+(1-\theta)}\right)^2 \left[\gamma_1 + \frac{\gamma_0 ea}{\left(\frac{M x_1(1-\theta)\theta^k}{\pi x_1(1-\theta^k)+(1-\theta)}+a\right)^2}\right] \\
\frac{\partial \phi_1}{\partial x_0} &= 0 \\
\frac{\partial \phi_1}{\partial x_1} &= \left(\frac{1-\theta}{\pi x_1(1-\theta^n)+(1-\theta)}\right)^2 \theta^n
\end{aligned}
$$

On the other hand, we have

$$
\begin{aligned}
\left\|\frac{\partial \phi(n,x)}{\partial x}\right\|_1 &= \max\left\{\left|\frac{\partial \phi_0}{\partial x_0}\right| + \left|\frac{\partial \phi_1}{\partial x_0}\right|, \left|\frac{\partial \phi_0}{\partial x_1}\right| + \left|\frac{\partial \phi_1}{\partial x_1}\right|\right\} \\
&\leq \max\left\{\alpha^n, \left(\frac{1-\theta}{1-\theta-\pi r_1}\right)^2 \left[\theta^n + \left(\gamma_1 + \frac{\gamma_0 ea}{(a-Mr_1)^2}\right)\frac{\alpha^n-\theta^n}{\alpha-\theta}\right]\right\} \\
&= \left(\frac{1-\theta}{1-\theta-\pi r_1}\right)^2 \left[\theta^n + \left(\gamma_1 + \frac{\gamma_0 ea}{(a-Mr_1)^2}\right)\frac{\alpha^n-\theta^n}{\alpha-\theta}\right] := L
\end{aligned}
$$

since

$$
\left(\frac{1-\theta}{1-\theta-\pi r_1}\right)^2 \left(\gamma_1 + \frac{\gamma_0 ea}{(a-Mr_1)^2}\right) > 1
$$

and

$$
\frac{\alpha^n-\theta^n}{\alpha-\theta} = \sum_{k=0}^{n-1} \alpha^k \theta^{n-1-k} \geq \alpha^n
$$

Using this bound on $\left\|\frac{\partial \phi(n,x)}{\partial x}\right\|_1$ in (43) we obtain

$$
\begin{aligned}
\Delta V_p(x) &\leq \Delta V(x) + \beta \sum_{n=0}^{N-1} \left(\frac{1-\theta}{1-\theta-\pi r_1}\right)^2 \left[\theta^n + \left(\gamma_1 + \frac{\gamma_0 ea}{(a-Mr_1)^2}\right)\frac{\alpha^n-\theta^n}{\alpha-\theta}\right] |x_0| \\
&= \Delta V(x) + \beta \left(\frac{1-\theta}{1-\theta-\pi r_1}\right)^2 \left[\frac{1-\theta^N}{1-\theta} + \frac{\gamma_1 + \frac{\gamma_0 ea}{(a-Mr_1)^2}}{\alpha-\theta}\left(\frac{1-\alpha^N}{1-\alpha} - \frac{1-\theta^N}{1-\theta}\right)\right] |x_0| \\
&\leq -\left[1-\alpha^N - \beta\left(\frac{1-\theta}{1-\theta-\pi r_1}\right)^2 \left[\frac{1-\theta^N}{1-\theta} + \frac{\gamma_1 + \frac{\gamma_0 ea}{(a-Mr_1)^2}}{\alpha-\theta}\left(\frac{1-\alpha^N}{1-\alpha} - \frac{1-\theta^N}{1-\theta}\right)\right]\right] |x_0| \\
&\quad - \left[1 - \frac{1-\theta}{1-\theta-\pi r_1}\left[(\gamma_1+\theta\gamma_0)\frac{1}{\alpha-\theta}\left(\alpha^N-\theta^N\right)+\theta^N\right]\right] |x_1| \\
&= -C_{30}|x_0| - C_{31}|x_1| \\
&\leq -\min\{C_{30}, C_{31}\}(|x_0|+|x_1|) \\
&= -C_3\|x\|_1
\end{aligned}
$$

where

$$
\begin{aligned}
C_{30} &= 1-\alpha^N - \beta\left(\frac{1-\theta}{1-\theta-\pi r_1}\right)^2 \left[\frac{1-\theta^N}{1-\theta} + \frac{\gamma_1 + \frac{\gamma_0 ea}{(a-Mr_1)^2}}{\alpha-\theta}\left(\frac{1-\alpha^N}{1-\alpha} - \frac{1-\theta^N}{1-\theta}\right)\right] \\
C_{31} &= 1 - \frac{1-\theta}{1-\theta-\pi r_1}\left[(\gamma_1+\theta\gamma_0)\frac{1}{\alpha-\theta}\left(\alpha^N-\theta^N\right)+\theta^N\right]
\end{aligned}
$$

and

$$C_3 = \min\{C_{30}, C_{31}\}$$

We note that C_{11} and C_2 are positive as in Theorem 1. Now, $C_{31} = D_{31}$, and hence the choice of N as in (39) guarantees that $C_{31} > 0$. Finally, if β is selected according to (40), C_{30} becomes positive as well, resulting in $C_3 > 0$. Hence, we conclude that the function

$$V(x) = \sum_{n=0}^{N-1} \|\phi(n,x)\|_1$$

is a Lyapunov function for the perturbed system satisfying the inequalities of (41)-(42). $\diamond$

The fact that we can find a Lyapunov function satisfying (41)-(42) enables us to conclude that the perturbed system $x(n+1) = f_p(x(n))$ is asymptotically stable in some neighborhood of the origin. The next theorem estimates this region of asymptotic convergence.

Theorem 2: *Let* $\mathcal{D} = \left\{x \in \mathcal{R}^2 : |x_0| < r_0, |x_1| < r_1\right\} \cap R_3$, *and* $\mathcal{D}_0 = \left\{x \in \mathcal{D} : \frac{C_{20}}{C_{21}}|x_0| + |x_1| \leq \frac{C_{11}}{C_{21}}\rho\right\}$, *where* $0 < r_0 < \infty$, $0 < r_1 < \frac{a-e}{M}$, *and* $\rho < r_1$. *Suppose* β *is picked as in (40). Then, the perturbed system*

$$x_0(n+1) = (1-\beta\gamma_0)x_0(n) + \gamma_1 x_1(n) + \frac{\gamma_0 e x_1(n)}{M x_1(n) + a} \tag{44}$$

$$x_1(n+1) = \frac{e x_1(n)}{M x_1(n) + a} - \beta x_0(n) \tag{45}$$

is asymptotically stable on $\mathcal{D}_0 \subset \mathcal{D}$. *Furthermore, trajectories of the system starting at* $x(0) \in \mathcal{D}_0$ *stay in* $\mathcal{D}$.

Proof: Let Ω_ρ be the set

$$\Omega_\rho = \{x \in \mathcal{D} : V(x) \leq C_{11}\rho\}$$

where $V(x)$ is as in Lemma 4. The set Ω_ρ contains $\mathcal{D}_0$ since if $x \in \mathcal{D}_0$, then

$$\begin{aligned} \frac{C_{20}}{C_{21}}|x_0| + |x_1| \leq \frac{C_{11}}{C_{21}}\rho \quad &\Rightarrow \quad C_{20}|x_0| + C_{21}|x_1| \leq C_{11}\rho \\ &\Rightarrow \quad V(x) \leq C_{11}\rho \\ &\Rightarrow \quad x \in \Omega_\rho \end{aligned}$$

On the other hand, Ω_ρ is a subset of $\mathcal{D}$ by definition. Thus

$$\mathcal{D}_0 = \left\{x \in \mathcal{D} : \frac{C_{20}}{C_{21}}|x_0| + |x_1| \leq \frac{C_{11}}{C_{21}}\rho\right\} \subset \Omega_\rho \subset \left\{x \in \mathcal{R}^2 : |x_0| < r_0, |x_1| < r_1\right\} = \mathcal{D}$$

For any $x(0) \in \Omega_\rho$, the solution starting at $x(0)$ at time zero stays in Ω_ρ. This follows from the fact that $\Delta V(x)$ is negative on $\mathcal{D} - \{0\}$; hence $V(x)$ is decreasing on the trajectory. Now, if we assume $x(0) \in \mathcal{D}_0 \subset \Omega_\rho$, then the solution still stays in Ω_ρ. Since Ω_ρ is a subset of $\mathcal{D}$, the solution clearly stays in $\mathcal{D}$. For the rest of the proof we will assume that $x(0) \in \mathcal{D}_0$. By (42) we have

$$\Delta V_p(x) = V(f_p(x)) - V(x) \leq -C_3\|x\|_1$$

Also by using inequality (41) we get

$$V(x) \leq C_2\|x\|_1 \Rightarrow \|x\|_1 \geq \frac{V(x)}{C_2}$$

Thus

$$\begin{aligned} \Delta V_p(x) = V(f_p(x)) - V(x) &\leq -C_3\|x\|_1 \\ &\leq -\frac{C_3}{C_2}V(x) \end{aligned}$$

Since $V(f_p(x)) = V(x(n+1))$, we obtain

$$V(x(n+1)) \leq \left(1 - \frac{C_3}{C_2}\right) V(x(n))$$

Let $y(n)$ be generated by the following difference equation:

$$y(n+1) = \left(1 - \frac{C_3}{C_2}\right) y(n), \quad y(0) = V(x(0)) \tag{46}$$

Clearly, $V(x(n)) \leq y(n),\ \forall n \geq 0$. The solution of (46) is given by

$$y(n) = y(0)\left(1 - \frac{C_3}{C_2}\right)^n = V(x(0))\left(1 - \frac{C_3}{C_2}\right)^n$$

Thus

$$\begin{aligned} V(x(n)) &\leq V(x(0))\left(1 - \frac{C_3}{C_2}\right)^n \\ \Rightarrow C_{11}|x_1(n)| &\leq [C_{20}|x_0(0)| + C_{21}|x_1(0)|]\left(1 - \frac{C_3}{C_2}\right)^n \\ &\leq C_{11}\rho\left(1 - \frac{C_3}{C_2}\right)^n \\ \Rightarrow |x_1(n)| &\leq \rho\left(1 - \frac{C_3}{C_2}\right)^n \end{aligned} \tag{47}$$

Now, for the convergence of the right-hand side of (47), we require $C_3 \leq C_2$, which is indeed the case. As a result, $x_1(n) \to 0$, as $n \to \infty$, but if $x_1(n) = 0$, then we are left with the system

$$x_0(n+1) = (1 - \beta\gamma_0)x_0(n)$$

which is stable with the choice of β as in (40). Hence, we conclude that the perturbed system (44)-(45) is asymptotically stable on $\mathcal{D}_0$, and if $x(0) \in \mathcal{D}_0$, then $x(n) \in \mathcal{D}$, $\forall n \geq 0$. $\diamond$

Remark 1: Theorem 2 basically helps us find a subregion of R_3 on which the nonlinear system (8)-(9) is stable. If $x(0) \in \mathcal{D}_0$, Theorem 2 guarantees that the trajectory of the system cannot escape $\mathcal{D}$. The region $\mathcal{D} = \left\{x \in \mathcal{R}^2 : |x_0| < r_0, |x_1| < r_1\right\} \cap R_3$ lies in R_3, and its size can be adjusted by varying r_0 and r_1. Figure 4 illustrates one choice of the regions $\mathcal{D}$ and $\mathcal{D}_0$ in R_3.

Now that we have determined that there exists a subset of R_3 on which (8)-(9) is stable, we ask what happens if we start outside this stable region. We have already established that trajectories starting outside region R (see Figure 2) converge to R in only one step. Similarly, it can be shown that any trajectory starting in R_4 enters either R_1 or R_3 only after a finite number of time units. To see this, recall that the following version of the state equations is valid in R_4:

$$\begin{aligned} x_0(n+1) &= \overline{Q} - Q \\ x_1(n+1) &= \frac{ex_1(n)}{Mx_1(n) + a} - \beta x_0(n) \end{aligned}$$

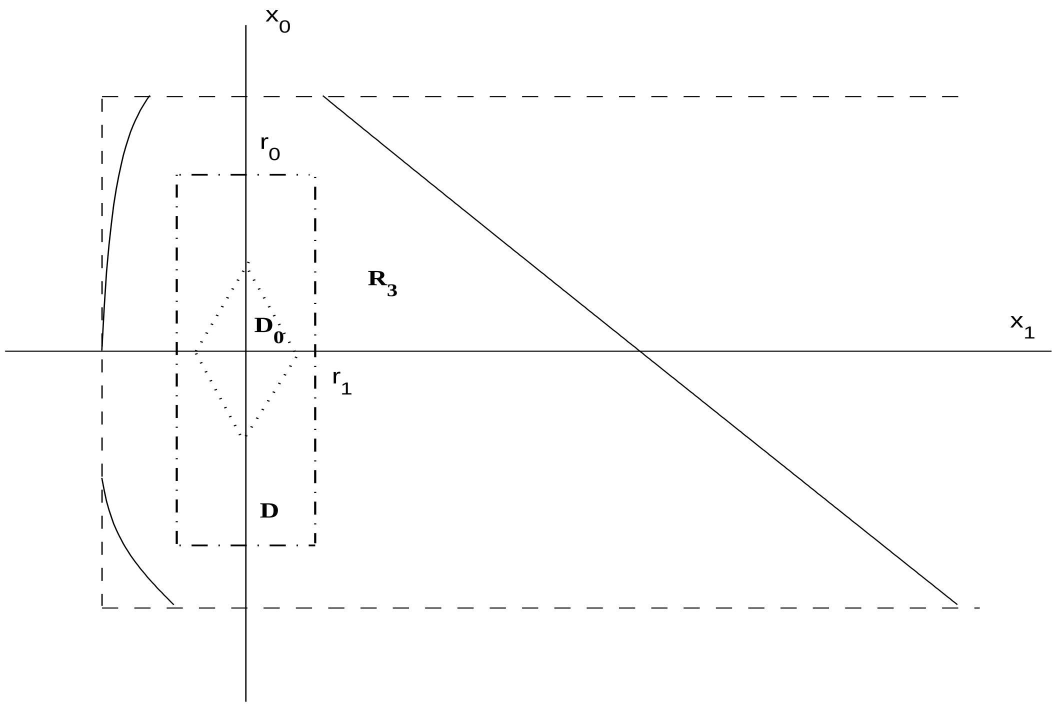

Figure 4: Regions $\mathcal{D}$ and $\mathcal{D}_0$ in R_3

Thus, if $x(0) \in R_4$, then $x_0(1) = \overline{Q} - Q > 0$. Now since the function $f(x) = \frac{ex}{Mx+a}$ is a contraction in x, $x_1(n)$ will decrease until it hits region R_3 or R_1 depending on its initial value. Following a similar argument, it can be shown that any trajectory starting in R_1 enters R_2 or R_3, after only a finite number of steps. And finally, trajectories starting in R_2 reach R_3 in a finite time. Thus, wherever we start in $x_1 - x_0$ plane, we end up in region R_3. Now the question is if we ever leave R_3 again. The foregoing Lyapunov analysis falls short of answering this question, however it establishes that there exists a subset of R_3 which is invariant. Extending this region of asymptotic stability to cover the entire region of R_3 is not immediate, and requires further research. As we show in the next section, simulations conducted with the algorithm strongly suggest that this is indeed the case.

Remark 2: In the foregoing analysis we have established a number of results regarding the convergence of the nonlinear system (8)-(9). These can be immediately linked to the original algorithm for ER congestion control, as the original variables are just shifted versions of x_0 and x_1. Thus, the convergence of x_0 and x_1 to the origin immediately implies convergence of the queue length, $q(n)$, to the desired set point Q, as well as the convergence of the source rates to their max-min fair share of the available bandwidth.

Remark 3: Both local and global analyses suggest that the controller gain β should be picked *small* to guarantee this convergence. How small β needs to be is precisely given by (14) and (40) in local and global cases, respectively. From simulations we have found that a rough guideline for choosing β is to make it inversely proportional to the maximum network delay, $\overline{d}$, maximum number of controlled connections, $\overline{M}$, and ratio of the available bandwidth to the bandwidth available to the controlled connections, i.e. $\frac{a}{a-e}$.

6 Simulations

6.1 Simulation of the Nonlinear Algorithm

In this section, we illustrate the global convergence of the nonlinear algorithm. Figure 5 shows trajectories of (8)-(9) starting in different regions of the state space. For the simulation example we took

$$a = 20,\ u = 1,\ \mathrm{MCR} = 1,\ M = 3,\ \overline{d} = 3,\ Q = 100,\ \overline{Q} = 200$$

and picked $\beta = 0.05$ consistent with (14) and (40). We run the algorithm with the following initial conditions each in different regions:

$$\begin{aligned} &\text{Region 1:} \quad x_1(0) = -6,\ x_0(0) = 50 \\ &\text{Region 2:} \quad x_1(0) = -6,\ x_0(0) = -95 \\ &\text{Region 3:} \quad x_1(0) = 10,\ x_0(0) = -50 \\ &\text{Region 4:} \quad x_1(0) = 25,\ x_0(0) = 50 \end{aligned}$$

The simulation run length is 100 time units. Note that the algorithm converges to the origin irrespective of the initial conditions.

6.2 Network Simulation: Single Link Case

We have simulated the robust adaptive ER congestion control algorithm on a hypothetical link, with a constant available bandwidth a, a constant rate of uncontrolled sources, u, and a slowly varying number of controlled connections, $M(n)$. For the simulation example, we assume that all controlled connections have the same MCR. The following parameter values are used in the simulation:

$$\mathrm{MCR}_m = 100,\ \overline{d} = 5,\ \overline{M} = 5,\ u = 400,\ Q = 1500,\ \overline{Q} = 3000,\ a = 1500,$$

All rates are in cells/time unit. The controller gain is picked $\beta = 0.015$ in accordance with the theory developed. Note that the rate of the uncontrolled connections is unknown to the switch.

As mentioned before, the exact number of controlled connections may not be known due to the bursty nature of the ABR traffic. We model this uncertainty as a random walk with boundaries at 1 and $\overline{M}$ and a step size of 1. We initially start with $M(0) = 3$ connections, and update this figure every $T(\lambda)$ time units, where $T(\lambda)$ is an exponential random variable with mean $\frac{1}{\lambda}$. We take $\lambda = 0.005$. A typical sample-path of $M(n)$ is depicted in Figure 6. Each user comes with a network delay, $d_m \leq \overline{d}$, which we assume to be unknown to the switch. We would like to demonstrate that our robust adaptive algorithm provides a max-min fair bandwidth allocation with a stable queue length despite the changing network conditions.

As can be seen from Figure 7, the link queue length is regulated around the desired value of $Q = 1500$ cells. The overshoots in Figure 7 occur when the value of $M(n)$ changes. They are inevitable when the amount of change in the number of controlled connections is comparable to the maximum number of connections,

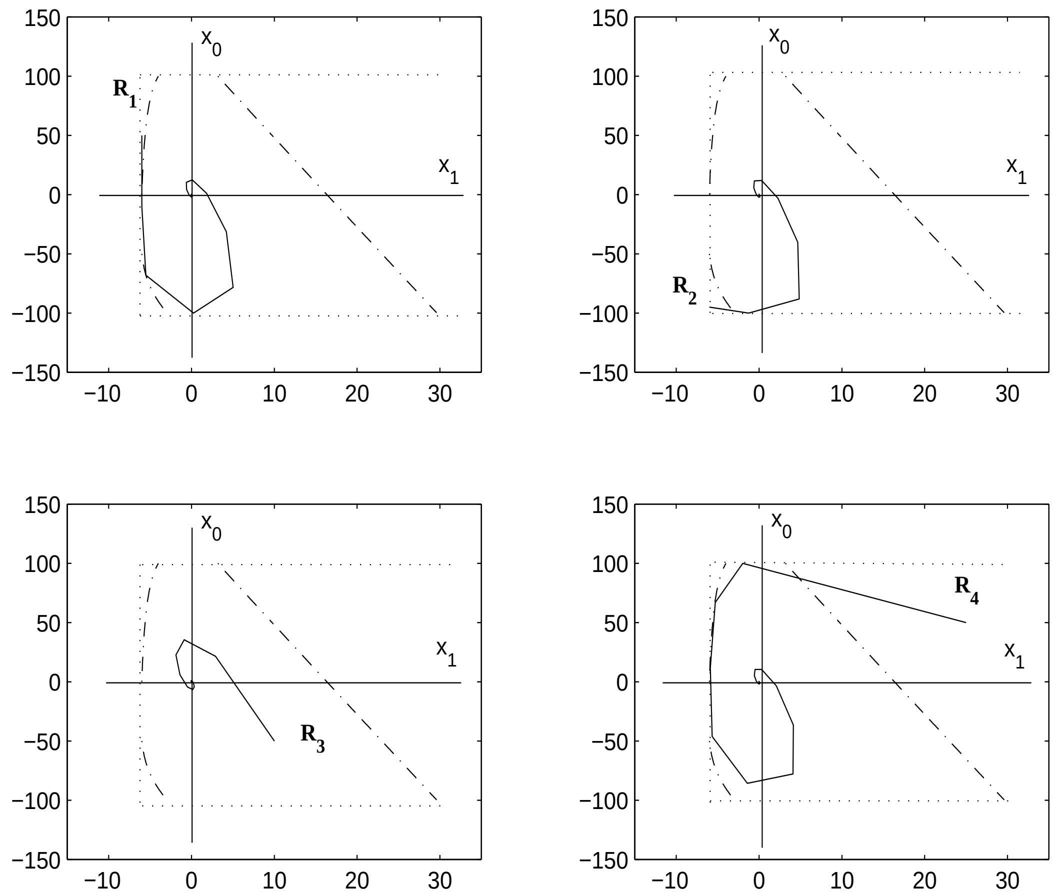

Figure 5: Trajectories of the system starting in Regions R_1, R_2, R_3 and R_4

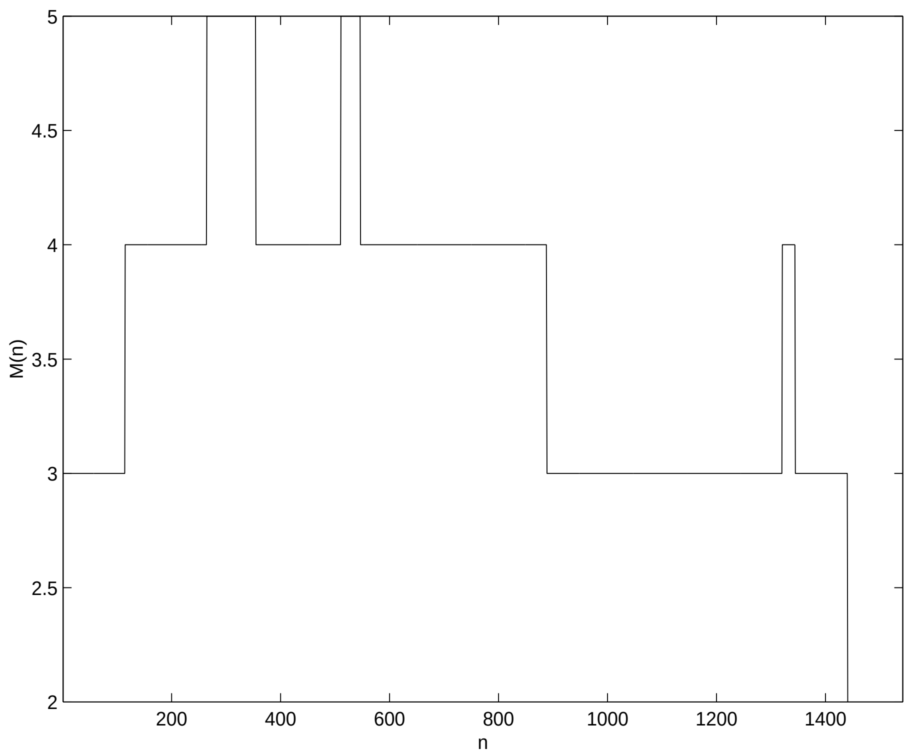

Figure 6: A typical sample-path behavior of the number of controlled connections

which is the case in this simulation example. As the number of connections increase, the transitions in the queue dynamics become smoother. In order to see that our design achieves max-min fairness as well, we first calculate the max-min fair share of the available bandwidth according to (1). As the number of controlled connections varies over time, this share varies, too. We want the output of the ER controller, ER(n), to be equal to this fair share at all time. In Figure 8 we plot ER(n). The flat portions of this graph corresponds to the fair share of the available bandwidth, as can be verified by direct calculation.

Finally, we note that the design parameter β can be chosen to tradeoff between the rate of convergence and the magnitude of overshoots. A smaller value of β results in a smaller overshoot, but a larger settling time.

7 Conclusions

In this paper, we have presented a robust adaptive congestion control algorithm for ABR service in ATM networks, and have shown that it performs well under various criteria, such as max-min fairness, PCR and MCR constraints, while at the same time achieving a high utilization of the available bandwidth. Despite unknown network delays and varying number of ABR connections, our design achieves a max-min fair bandwidth allocation among the ABR sources bottlenecked on a given link. In addition, the algorithm stabilizes the queue length around a desired set point, Q, resulting in an efficient use of network resources. The algorithm we propose here does not suffer from computational complexity, because there is a single

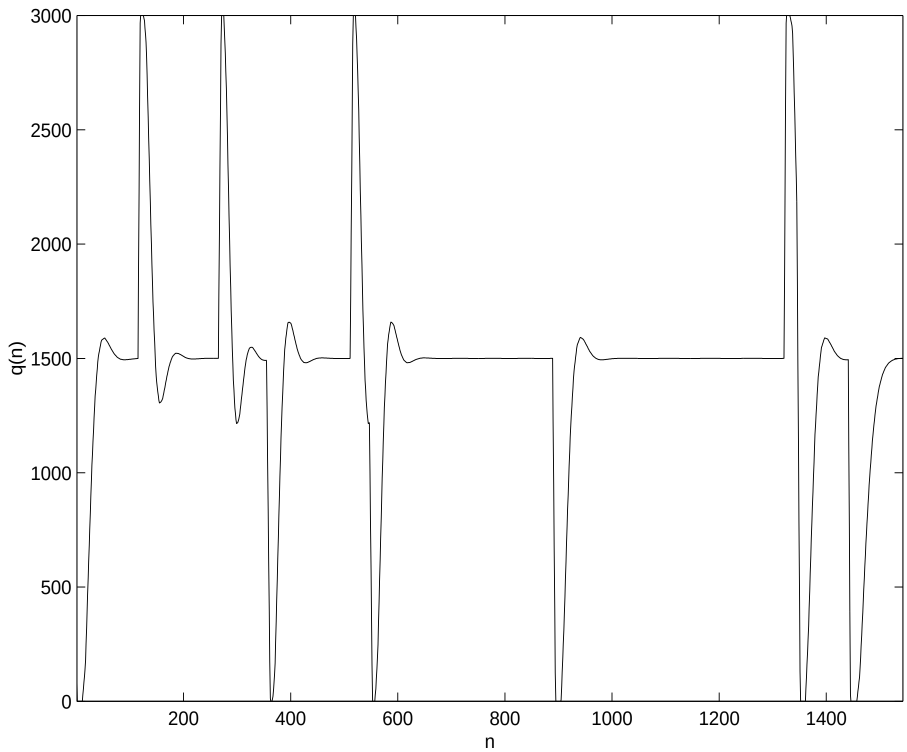

Figure 7: Typical sample-path behavior of the link queue length

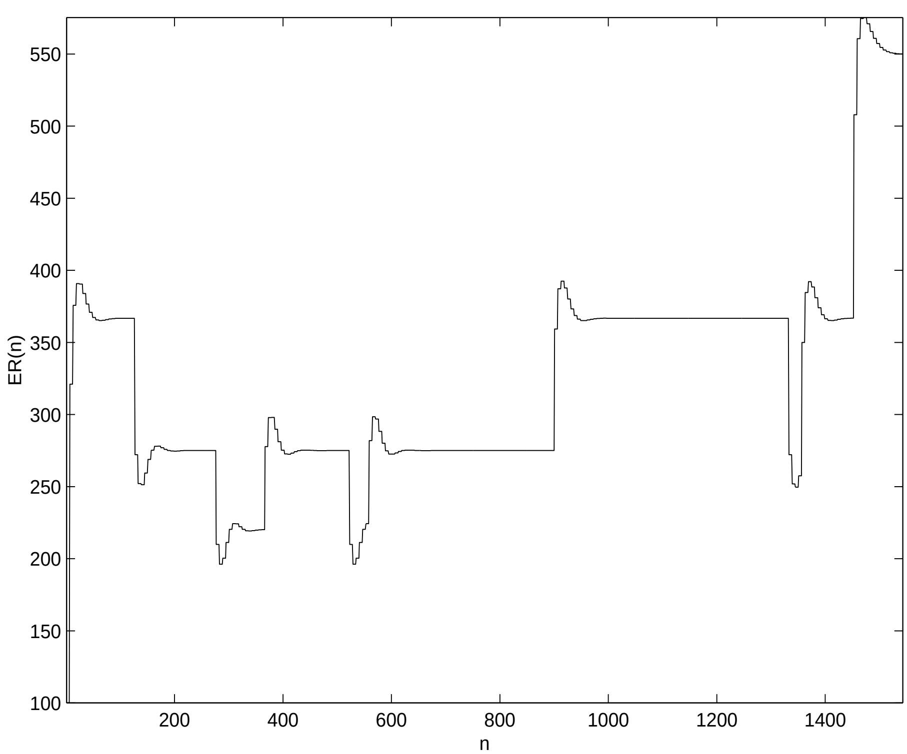

Figure 8: Typical sample-path behavior of the ER controller

design parameter, namely β, to be tuned, and to determine the explicit rate, ER, the switch has to perform only two divisions, one multiplication, and two additions per output line, every $(\overline{d}+1)$ time units. Moreover, the information needed to perform these calculations, $\{r_s(n), q_s(n), a(n)\}$, is locally available to the switch. This feature makes it possible to implement the algorithm in a decentralized fashion.

Further study needs to be done in simulating the robust adaptive algorithm in a real network environment to see how well it reacts to network latencies, as well as to evaluate its performance under various scenarios.

References

[1] The ATM Forum, Technical Committee, "Traffic management, Version 4.1," pp. 43-55, af-tm-0121.000, March 1999.

[2] S. Kalyanaraman, R. Jain, S. Fahmy, R. Goyal, and B. Vandalore, "The ERICA switch algorithm for ABR traffic management in ATM networks," *IEEE/ACM Trans. on Networking*, vol. 8, no. 1, pp. 87-98, February 2000.

[3] O. Ait-Hellal, E. Altman, and T. Başar, "Rate-based flow control with bandwidth information," *European Transactions on Telecommunications*, vol. 8, no. 1, pp. 55-65, 1997.

[4] L. Benmohamed and S. M. Meerkov, "Feedback control of congestion in packet switching networks: The case of a single congested node," *IEEE/ACM Transactions on Networking*, vol. 1, no. 6, pp. 693-707, 1993.

[5] L. Benmohamed and Y. T. Wang, "A control-theoretic ABR explicit rate algorithm for ATM switches with per-VC queueing," in *Proceedings of IEEE INFOCOM*, San Francisco, CA, March 1998.

[6] A. Kolarov and G. Ramamurthy, "A control-theoretic approach to the design of an explicit rate controller for ABR service," *IEEE Transactions on Networking*, vol. 7, no. 5, pp. 741-753, October 1999.

[7] D.P. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[8] O.Ç. Imer, T. Başar, and R. Srikant, "A robust adaptive control algorithm for ABR service in ATM networks," in *Proceedings of IEEE ICCCN*, Las Vegas, AZ, October 2000.

[9] P. J. Antsaklis and A. N. Michel, *Linear Systems*. New York, NY: McGraw-Hill, 1997, pp. 489-505.