Congestion Control as a Stochastic Control Problem with Action Delays ⋆

Eitan Altman [a] Tamer Başar [b] R. Srikant [c]

[a] *INRIA B.P. 93, 06902 Sophia Antipolis Cedex, France. Email: altman@sophia.inria.fr*

[b] *Corresponding author, Coordinated Science Laboratory and Department of Electrical and Computer Engineering, University of Illinois, 1308 West Main Street, Urbana, IL 61801, USA. Email: tbasar@decision.csl.uiuc.edu*

[c] *Coordinated Science Laboratory and Department of General Engineering, University of Illinois, 1308 West Main Street, Urbana, IL 61801, USA. Email: rsrikant@uiuc.edu*

Stochastic control tools can be used to develop effective algorithms for control of congestion in high-speed communication networks with varying bandwidth.

Abstract

We consider the design of explicit rate-based congestion control for high-speed communication networks and show that this can be formulated as a stochastic control problem where the controls of different users enter the system dynamics with different delays. We discuss the existence, derivation and the structure of the optimal controller, as well as of suboptimal controllers of the certainty-equivalent type—a terminology that is precisely defined in the paper for the specific context of the congestion control problem considered. We consider, in particular, two *certainty-equivalent controllers* which are easy to implement, and show that they are stabilizing, i.e., they lead to bounded infinite-horizon average cost, and stable queue dynamics. Further, these controllers perform well in simulations.

Key words: Communication networks; stochastic control; certainty equivalence; optimal control

⋆ Research supported by NSF Grants NSF ECS 93-12807, NSF NCR 9701525 and NSF ANI 98-13710, and AFOSR MURI Grant AF DC 5-36128.

1 Introduction

Traffic in high-speed communication networks can be broadly classified as guaranteed-service traffic and best-effort service traffic. Guaranteed service refers to a contract between the network service provider and the end user which requires the network to provide a fixed quality-of-service (QoS) to the traffic. The QoS guarantees could be in the form of upper bounds on packet loss probability, delay, etc. In contrast, best-effort traffic is guaranteed very little *a priori*. In the Internet, there are no guarantees, while in the context of Asynchronous Transfer Mode (ATM) networks, the best-effort traffic (in particular the Available Bit Rate (ABR) service) may be guaranteed a minimum data rate and a bound on the loss rate. Instead of guaranteeing fixed QoS parameters, the idea is to *fairly* allocate the network resources to competing users. The guaranteed-service traffic (referred to as either Constant Bit Rate (CBR) or Variable Bit Rate (VBR) traffic in ATM networks) gets a higher scheduling priority compared to the best-effort traffic. In other words, at a given node, when both best-effort traffic and guaranteed-service traffic are backlogged (i.e., have packets that they are waiting to send), the packets from the guaranteed-service traffic are processed first, and the best-effort traffic is served only if there are no packets in the guaranteed-service queue.

From a control point of view, each best-effort source or user may be thought of as an entity that generates data at a rate that is upperbounded by what is specified by the network. The network exercises control over the best-effort traffic by assigning these rates based on the congestion in the network. This is referred to as *congestion control*. In the absence of such a control mechanism, the buffers at each node in the network (which store packets temporarily) may overflow and lead to packet losses. Loosely speaking, the network attempts to operate on the following principle: If the sources obey the control commands issued by the network, then the network attempts to transfer the packets without any loss. Of course, this can be achieved by making the data rates equal to zero for all the sources, which is clearly not desirable. Thus, another goal of the network is to maximize utilization, i.e., the the sum of the data rates of all the sources should be nearly equal to the total capacity available for best-effort sources.

There are two basic approaches to congestion control. In the Internet, the protocol used for this purpose is called TCP (see, (Jacobson, 1988) for the original congestion avoidance scheme in TCP, and (Keshav, 1997), and references within, for subsequent modifications). In the TCP context, each source slowly increases the rate at which it transmits data and upon detection of congestion, the data rate is reduced. Congestion is detected when buffers overflow and packets are lost. It is the responsibility of the destination to inform the source of lost packets and intermediate nodes in the network do not provide any feedback directly to the source.

In contrast, in ATM ABR service, the communication protocol between the source and the network allows for intermediate nodes to inform the source when queues start to build up or if the input rate is in excess of the available capacity for ABR sources at that node. In fact, instead of exactly providing this information back to the sources, the intermediate nodes are capable of restricting the data rate of the ABR sources (Benmohammed and Meerkov, 1993, Mascolo et al., 1996, Kalyanaraman et al., 1997, Kolarov and Ramamurthy, 1997, Altman et al., 1997, Altman et al., 1998, Benmohammed and Wang, 1998). In this paper, we adopt this framework. Thus, from the point of view of the current technology, one can view our work as a congestion control mechanism for ATM ABR sources.

For the rest of the paper, we will adopt the point of view that there is a single bottleneck node that plays a dominant role in determining the performance of a given set of sources. In this case, 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 by Kalyanaraman et al. (1997), although several modifications are used in the actual implementation. The main advantage of this scheme is its simplicity, but it is difficult to optimally control queue length to avoid buffer overflows. However, this scheme is stable, i.e., the queue length remains bounded in an appropriate stochastic sense, provided that the targeted utilization is less than one (Ait-Hellal et al., 1997). Queue length information is not used in the basic algorithm, although Kalyanaraman et al. (1997) allows one to incorporate queue length information in an ad hoc manner.

Alternatively, this problem can also be viewed as a feedback control problem where queue length is used as the explicit feedback. This approach is used in Benmohammed and Meerkov (1993), Mascolo et al. (1996), Kolarov and Ramamurthy (1997), and Benmohammed and Wang (1998) to study this problem using classical control techniques or using a state-space approach. As in rate matching, the primary goal has not been one of optimality, but simply of queue-length stability. In these approaches, the available bandwidth to ABR sources is treated as an unmodelled disturbance. Thus, the algorithms in Benmohammed and Meerkov (1993), Mascolo et al. (1996), Kolarov and Ramamurthy (1997), and Benmohammed and Wang, (1998) ensure stability in the presence of this disturbance.

In contrast to the above approaches, we use both available rate and queue length to compute the data rates for the ABR sources. To allow this in a control-theoretic formulation, we model the available bandwidth as an autoregressive (AR) process driven by a white noise process, as in Altman and Başar

(1995). This allows us to represent the fluctuations of available bandwidth due to variations in instantaneous throughputs of high-priority sessions as well as to variations in the number of connected high-priority sessions. The modeling of real time traffic as an AR process is often used, both for characterization of correlation and traffic parameters, as well as for queueing performance models. Some examples in the literature are Maglaris et al. (1988), Melamed et al. (1992), and Melamed et al. (1994). In particular, a queueing performance for AR traffic is given in Addie and Zukerman (1972). AR processes for traffic have also been widely used in the context of control. Ramamurthy and Sengupta (1993) use an auto-regressive approach for designing a predictive congestion control, in the presence of other real time traffic. The stability of rate based flow control of ABR (Available Bit Rate) traffic in ATM in the presence of exogenous VBR traffic is analyzed in Ait-Hellal et al. (1997) and Altman et al. (1994). In this paper, working with a specific model of the available bandwidth enables us to achieve much better performance than other existing algorithms, as also demonstrated through simulations.

There is a fundamental difficulty in obtaining good congestion control performance, namely the presence of *action delay*, which is defined in terms of two components (see Figure 1). The first component is the *downstream delay*,

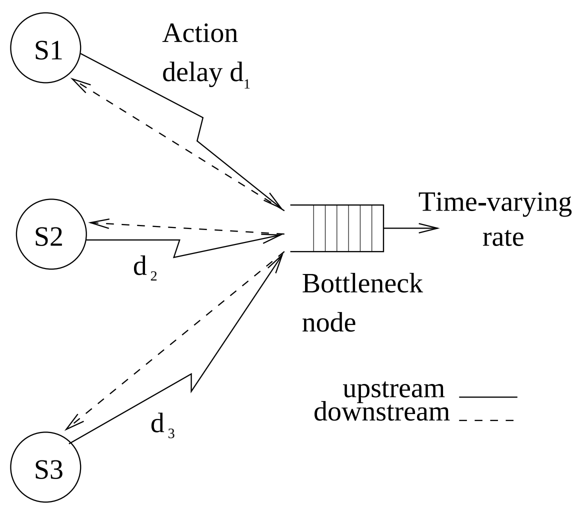

Fig. 1. Bottleneck node model with three ABR sources. The sum of the upstream and downstream delays is called the action delay, and is denoted by d_i for Source i.

i.e., the delay between the time that the bottleneck node issues its command to the time that it takes for a source to receive this command. The second component is the *upstream delay*, i.e., the time that it takes for the data packets generated by the source to reach the bottleneck node. The sum of these delays is the action delay. It is well known in the control literature that the presence of delays in the feedback path generally poses difficulties. Here, this problem is further amplified due to the fact that the action delays are different for different sources. While the simple rate matching algorithms (Ait-Hellal et al., 1997), (Kalyanaraman et al., 1997) do not account for delay, the control-theoretic approaches (Benmohammed and Meerkov, 1993), (Mascolo et al., 1996), (Kolarov and Ramamurthy, 1997), (Benmohammed and Wang, 1998) account for feedback delay in their solutions. In our work, we also explicitly take delay into account, but since we also stochastically model the available capacity, this poses a more interesting challenge due to the fundamental role of information structure in stochastic control problems, as we will see.

We formulate the congestion control problem as an LQG stochastic control problem, where the control actions of all users are actually determined centrally by the node, which however has to take into account the fact that these different actions will affect the queue dynamics at different times due to upstream and downstream delays. Even though this is not directly related to the analysis of this paper, one can actually show that (Altman et al., 1998) the centralized control problem with action delays is equivalent to a decentralized team problem with information delays, where now the decision are made by the users. In the parlance of team theory, this fits into the class of LQG teams with *nested information* (Chu and Ho, 1972, Witsenhausen, 1971), and hence is not intractable as those with nonclassical information (Witsenhausen, 1968, Bansal and Başar, 1987, Srikant and Başar, 1992).

The problem posed here does in fact admit an optimal solution, which is characterized in terms of the solution of a discrete-time algebraic Riccati equation (DARE) whose dimension is determined by the magnitude of the largest delay and the order of the AR process describing the available capacity (Imer and Başar, 1999b). There are, however, other solutions, easier to implement (they involve the solution of a scalar DARE), which share a common appealing feature of *certainty equivalence* in addition to being also stabilizing. It is the derivation of these suboptimal controllers that this paper is focussed on, along with their stability, and a comparative study of their performance through simulations. We also clarify the issue of certainty equivalence for problems of the type considered here.

The rest of the paper is organized as follows. In Section 2, we present the mathematical formulation of the congestion control problem, discuss the nature of the optimal solution, and introduce a notion of *certainty equivalence*, tailored to the problem at hand. In Section 3, we consider a special finite-horizon version of this problem, derive the optimal as well as two suboptimal certainty-equivalent solutions. This then leads to the presentation, in Section 4, of the two certainty-equivalent controllers for the original infinite-horizon problem. In Section 5, we show that, while these two certainty-equivalent controllers may not be optimal, they are still stabilizing. Simulation results are presented in Section 6, and concluding remarks are included in Section 7.

2 Mathematical model and the notion of certainty equivalence

2.1 Mathematical model

We consider here a discrete-time model, where a time unit corresponds to the length of the minimum measurement interval. Let q_n denote the queue length at a bottleneck link, and μ_n denote the effective service rate available for ABR traffic in that link at the beginning of the nth time slot. Let r_{mn} denote the effective source rate for source m ($m = 1, \ldots, M$) at the input of the bottleneck link during the nth time slot, which is actually the outcome of an action taken by source m several time steps earlier, based on command signal sent by the switch even earlier. We denote the total time it takes for the decision of the switch on the transmission rate of source m to reach that source and subsequently for the effect of this decision to reach the bottleneck node (i.e., the sum of *downstream* and *upstream* delays–using the terminology introduced in the previous section) by d_m,[1] and the command decision of the switch for source m at time n by v_{mn}, which we will sometimes also write as $v_{m,n}$. Hence, we have the relationship:

$$v_{m,n-d_m} = r_{mn} \,. \tag{1}$$

Now, in terms of the notation introduced, the queue length evolves according to

$$q_{n+1} = q_n + \sum_{m=1}^{M} r_{mn} - \mu_n \equiv q_n + \sum_{m=1}^{M} v_{m,n-d_m} - \mu_n \,. \tag{2}$$

The above equation corresponds to a linearized version of the actual queue dynamics, since we have ignored the fact that the queue length cannot be negative. Simulations in Altman et al. (1997) and Altman et al. (1998) show that this linearization is in fact valid when the controllers are successful in maintaining the queue size around a positive target value Q, sufficiently away from *zero*. The service rate μ_n available to the sources may change over time in an unpredictable way, since this is the capacity left over from high priority traffic. We model this available capacity by a p-dimensional stable AR process:

$$\mu_n = \mu + \xi_n \tag{3}$$

[1] Without any loss of generality, we take the d_m's to be ordered in accordance with their indices, that is $d_1 \le d_2 \le \ldots, \le d_M$.

$$\xi_n = \sum_{i=1}^{p} \alpha_i \, \xi_{n-i} + \phi_{n-1} \,, \tag{4}$$

where μ is the known constant nominal service rate, α_i, $i = 1, \ldots, p$, are known parameters, and $\{\phi_n\}_{n\geq 1}$ is a zero-mean *i.i.d.* sequence with finite variance denoted by k^2.

The objective function, to be minimized by the switch, involves the transmission rates of all the sources which use the isolated bottleneck node, as well as the length of the queue at that node, and is given by

$$J = \limsup_{N\to\infty} \frac{1}{N} E \left\{ \sum_{n=1}^{N} \left[(q_n - Q)^2 \; + \sum_{m=1}^{M} \frac{1}{c_m^2} \left(r_{mn} - a_m \, \mu_n \right)^2 \right] \right\} \tag{5}$$

where Q is the target queue length, c_m's are some positive constants, and $\sum_{m=1}^{M} a_m = 1$. The *first additive term* above represents a penalty for deviating from a desirable queue length. The *second additive term* is a measure of the quality with which the input rate for each source tracks a given fraction of the available service rate, where the c_m's are weighting terms that serve to prioritize relative importance of these individual terms (among different sources as well as collectively with respect to the first additive term). For example, if we desire "fair" sharing of the available bandwidth, we would choose

$$a_1 = a_2 = \cdots = a_M = \frac{1}{M},$$

assuming that everything else is also symmetric for the sources.

The information available to the switch at time n is I_n, where

$$I_n = \{\xi_n, \xi_{n-1}, \ldots; q_n, q_{n-1}, \ldots; v_{jk}, j = 1, \ldots, M, \; k \leq n\}$$

Hence,

$$v_{mn} = \gamma_{mn}(I_n), \; n = 1, 2, \ldots, ; \quad m = 1, 2, \ldots, M,$$

where γ_{mn} is some Borel-measurable function, with respect to which J will be minimized.

We now introduce the new (appropriately shifted) variables

$$x_n := q_n - Q \tag{6}$$
$$u_{mn} := v_{mn} - a_m \mu \tag{7}$$

which will serve as the state and the control, respectively. The queue dynamics (2) can be re-written in terms of these quantities as:

$$x_{n+1} = x_n + \sum_{m=1}^{M} u_{m,n-d_m} - \xi_n \tag{8}$$

$$\xi_{n+1} = \sum_{i=1}^{p} \alpha_i \, \xi_{n+1-i} + \phi_n \tag{9}$$

and the cost function J can be re-written as:

$$J = \limsup_{N\to\infty} \frac{1}{N} E \left\{ \sum_{n=1}^{N} \underbrace{\left[(x_n)^2 + \sum_{m=1}^{M} \frac{1}{c_m^2} (u_{m,n-d_m} - a_m \, \xi_n)^2 \right]}_{\ell_n(x_n,\xi_n;u_n)} \right\} \tag{10}$$

where ℓ_n denotes the instantaneous (running) cost.

Now, assume for the moment that all delays are zero, and introduce

$$\tilde{u}_{mn} := u_{mn} - a_m \, \xi_n \, ; \quad \tilde{u}_n := (\tilde{u}_{1n}, \ldots, \tilde{u}_{Mn})' \, .$$

Then, the instantaneous cost takes the pure quadratic form

$$\ell_n(x_n, \xi_n; u_n) = \tilde{\ell}_n(x_n; \tilde{u}_n) = (x_n)^2 + |\tilde{u}_n|_R^2 \, . \tag{11}$$

where $|\cdot|_R$ denotes the Euclidean norm in the M-dimensional vector space, weighted by R, and

$$R := \ \mathrm{diag} \left(\frac{1}{c_1^2}, \cdots, \frac{1}{c_M^2} \right) . \tag{12}$$

Also, in terms of $\tilde{u}$, the dynamics for x become:

$$x_{n+1} = x_n + b' \, \tilde{u}_n \, , \tag{13}$$

where

$$b' := (\, 1 \ 1 \ \cdots \ 1 \,) \, . \tag{14}$$

Hence, if there were no action delays, we would have a standard discrete-time linear-quadratic regulator problem, which admits the unique solution (Anderson and Moore, 1990):

$$\tilde{u}_n = -[R + bb's]^{-1} bs\, x_n \,, \tag{15}$$

where s is the unique positive root of the algebraic equation

$$s = 1 + s[1 - b'(R + bb's)^{-1} bs] \,. \tag{16}$$

This positive root can in fact be explicitly computed:

$$s = \frac{1 + \sqrt{1 + 4c^2}}{2}\,; \quad c^2 := \left(\sum_{m=1}^{M} \frac{1}{c_m^2} \right)^{-1} \,. \tag{17}$$

The corresponding control $\tilde{u}$ renders the closed-loop queue dynamics asymptotically stable:

$$x_{n+1} = \frac{1}{2} \left(\frac{1 + 2c^2 - \sqrt{1 + 4c^2}}{c^2} \right) x_n \,.$$

Note that in this case the queue dynamics are completely decoupled from the AR process (which is stable by our initial hypothesis). Further note that in spite of this decoupling, the original control u_n is still a function of ξ_n, since the inverse transformation from $\tilde{u}_n$ to u_n yields

$$u_n = \tilde{u}_n + \begin{pmatrix} a_1 \\ \vdots \\ a_M \end{pmatrix} \xi_n \,. \tag{18}$$

Expression (15) can be simplified further by applying the following matrix inversion lemma (Appendix D, Sage and White, 1977) to the matrix $(R + bb's)$.

Lemma 1 *Suppose that the square and compatible-dimensional matrices A, C and $A + B'C^{-1}B$ are all invertible. Then*

$$(A + B'C^{-1}B)^{-1} = A^{-1} - A^{-1}B'(BA^{-1}B' + C)^{-1}BA^{-1}.$$

■

Using this lemma, we obtain

$$(R + bb's)^{-1} = R^{-1} - R^{-1}b(b'R^{-1}b + \frac{1}{s})^{-1}b'R^{-1}.$$

Since R is a diagonal matrix, we have

$$R^{-1} = diag(c_1^2, c_2^2, \ldots, c_M^2)\,,$$

and

$$b'R^{-1}b = \sum_{i=1}^{M} c_i^2,$$

and hence the $(i,i)^{\text{th}}$ element of $(R + bb's)^{-1}$ is given by

$$c_i^2 - \left(sc_i^4 / \left(1 + s\sum_{\ell=1}^{M} c_\ell^2\right)\right),$$

and, for $i \neq j$, the $(i,j)^{\text{th}}$ element is given by

$$-sc_i^2c_j^2 / \left(1 + s\sum_{\ell=1}^{M} c_\ell^2\right).$$

Thus, the m^{th} element of $(R + bb's)^{-1}bs$, to be denoted by p_m, is

$$\begin{aligned} p_m &= c_m^2 - \frac{sc_m^4}{1 + s\sum_{\ell=1}^{M} c_\ell^2} - \frac{sc_m^2 \sum_{\ell=1,\ell\neq m}^{M} c_\ell^2}{1 + s\sum_{\ell=1}^{M} c_\ell^2} \\ &= c_m^2 / \left(1 + s\sum_{\ell=1}^{M} c_\ell^2\right). \end{aligned} \tag{19}$$

In view of this, we can now rewrite (18) for the individual components of u_n as:

$$u_{mn} = -p_m\, x_n + a_m\, \xi_n\,, \quad m = 1\ldots, M\,. \tag{20}$$

This is therefore the optimum transmission rate for source m if there were no delay (upstream or downstream) in the network. In the presence of delay, however, these rates could lead to an unstable queue system, and hence there is a need to take into account the upstream and downstream delays on various links.

2.2 *The optimal solution and certainty equivalence*

The optimal solution to the problem, as formulated by (8), (9), and (10), can be obtained either by dynamic programming (Imer and Başar, 1999a) or by converting it into a standard linear-quadratic optimal control problem in an extended state space (Imer and Başar, 1999b). In this latter approach, one introduces additional state variables (precisely d_M of them) so that the controls u_{mn}'s enter the state equation in this extended space without delay. Furthermore, one has to introduce additional state variables (precisely, $p-1$ of them) so as to convert equation (9) describing the AR process into a first-order difference equation. Finally, one also has to rewrite the second set of additive terms in (10) (by an appropriate expansion, and by making use of the fact that future values of ϕ_n's are independent of the past controls) in such a way that the term corresponding to $u_{m,n-d_m}$ involves only $\xi_j,\ j \leq n-d_m$ – this being so for all m and n. These transformations and extensions bring the problem into a linear-quadratic control problem with perfect state information, and with the d_M+p+1-dimensional state equation driven by a zero-mean white noise sequence. The optimal transmission rates for the sources are then obtained as linear feedback on the d_M+p+1-dimensional state, with the gain determined from the solution of a d_M+p+1-dimensional discrete-time algebraic Riccati equation (DARE). Since the overall system is stabilizable-detectable, this DARE admits a unique solution in the class of nonnegative-definite matrices, and the optimal controller leads to stable queue dynamics (Anderson and Moore, 1990). Our interest in this paper is to explore the possibility of obtaining other (clearly suboptimal, but still queue-stabilizing) controllers which would not require the (numerical) solution of a high-dimensional DARE, which in real implementations will have to be updated periodically as the parameters (α_i's and p) of the AR process and the various delays $(d_1, \ldots, d_M)$ are revised. For this, our starting point will be the "no-delay" optimal solution given by (20), which we now rewrite as follows by noting that in the presence of delay the control u_{mn} should actually be replaced by $u_{m,n-d_m}$:

$$u_{m,n-d_m} = -p_m\, x_n + a_m\, \xi_n\,, \quad m = 1, \ldots, M\,. \tag{21}$$

But this is not implementable, since the state (x_n, ξ_n) is not available to $u_{m,n-d_m}$, but only I_{n-d_m} is. Hence, one approach (that uses (21) as a starting point) would be to replace x_n and ξ_n in (21) by their *predicted values,*[2] given the information I_{n-d_m}, and the controllers $u_{jk_j},\ j<m,\ k_j<n-d_j$. We will call such a controller, which is derived from (21) and is compatible with the available information, a *certainty-equivalent* controller, and write it as:

[2] This terminology is used here in a rather loose sense. It will be made precise in the next two sections.

$$u_{m,n-d_m} = -p_m\, \hat{x}_{n|n-d_m} + a_m\, \xi_{n|n-d_m} \tag{22}$$

Hence, a *certainty-equivalent* controller has the property that if the predicted values of x and ξ were to coincide with the actual values used in the optimal solution that ignores the presence of delays, then it would be an optimal controller, but since this can never happen in the presence of delays, a certainty-equivalent controller will lead to a performance that is worse than that attained under (21). Now, an important point to note here is that (22) above is not the only certainty-equivalent controller, simply because (21) is not the only optimal controller for the delay-free problem. Any *representation* of (21) that shows explicit dependence on the controls of other sources, and which reduces to (21) when all these controllers are set to their optimal choices according to (21), leads to exactly the same performance as that attained by (21). Within a linear structure, we can characterize all these representations of (21) by:

$$\begin{aligned} u_{m,n-d_m} = &-p_m\, x_n + a_m\, \xi_n \\ &+ \sum_{j=m+1}^{M} \sum_{k<n-d_m} \kappa_{j,k}^{m,n} \left(u_{jk} + p_j x_{k+d_j} - a_j \xi_{k+d_j} \right), \end{aligned} \tag{23}$$

where $\kappa_{j,k}^{m,n}$'s for $j > m, m = 1, \ldots, M; k < n - d_m$ are arbitrary parameters, and the dependence on the controls of other sources exhibit a lower-shifted upper triangular structure consistent with the ordering of the delays (that is, $u_{m\ell}$ is allowed to depend on u_{jk} if and only if $d_m \leq d_j$ and $k \leq \ell - d_m$). This, in a sense, constitutes an equivalence class of controllers for the sources, all leading to the same value for the cost (10). This is though not the entire class of such controllers, because we could also have included nonlinear terms in (23), which we have not; it is, however, the entire class of *linear* controllers with a lower-shifted upper triangular dependence on other controllers as described above. Now, a certainty-equivalent controller derived from each element of that equivalence class would be:

$$\begin{aligned} u_{m,n-d_m} = &-p_m \hat{x}_{n|n-d_m} + a_m \hat{\xi}_{n|n-d_m} \\ &+ \sum_{j=m+1}^{M} \sum_{k<n-d_m} \kappa_{j,k}^{m,n} \left(u_{j,k} + p_j \hat{x}_{k+d_j|n-d_m} - a_j \hat{\xi}_{k+d_j|n-d_m} \right), \end{aligned} \tag{24}$$

where $\hat{x}_{k'|n-d_m}$ and $\hat{\xi}_{k'|n-d_m}$ are defined to be $x_{k'}$ and $\xi_{k'}$ for $k' := k + d_j \leq n - d_m$. A point to stress here is that these certainty-equivalent controllers no longer constitute an *iso-cost* equivalence class, and in fact each one leads to a different value for the cost (10). The optimal policy (whose derivation was discussed at the beginning of this subsection) is indeed a certainty-equivalent controller, and corresponds to (24) for some specific values of the κ_{jk}^{mn}'s. A class of certainty-equivalent controllers where in the inner summation in (24) only the term corresponding to $k = n - d_j$ is retained, that is:

$$\begin{aligned} u_{m,n-d_m} = &-p_m\, \hat{x}_{n|n-d_m} + a_m\, \hat{\hat{\xi}}_{n|n-d_m} \\ &+ \sum_{j=m+1}^{M} \kappa^{m,n}_{j,n-d_j}\, (\, u_{j,n-d_j} + p_j \hat{x}_{n|n-d_m} - a_j \hat{\hat{\xi}}_{n|n-d_m}\,)\,, \end{aligned} \tag{25}$$

carries some appealing properties (from the point of view of ease in implementation), and will be studied further in this paper. In particular, we present in Section 4 two such certainty-equivalent controllers, where we also clarify precisely what we mean by "predicted values" of x_n and ξ_n, and subsequently show in Section 5 that each of these controllers leads to stable queue dynamics. But first we discuss in the next section a simpler finite-horizon problem, to illustrate and further clarify the discussion of this subsection on optimality and certainty equivalence.

3 A Three-Stage Example

We consider here a special finite-horizon version of the optimal control problem described by (8), (9), (10), with delayed actions, to demonstrate the existence of multiple certainty-equivalent controllers, and the derivation of the optimal one which is also certainty equivalent. This special case involves two users and three stages, with user 1 acting at all three stages and having no action delay (i.e., $d_1 = 0$), and user 2 acting only once, but having a two-step action delay (i.e., $d_2 = 0$). In accordance with this, the shifted queue dynamics are

$$\begin{aligned} x_4 &= x_3 + u_{13} + u_{21} - \xi_3 \\ x_3 &= x_2 + u_{12} - \xi_2 \\ x_2 &= x_1 + u_{11} - \xi_1 \end{aligned} \tag{26}$$

where u_{in} is the action of user i at stage n, and ξ_n, $n \geq 1$, is generated by the first-order AR process:

$$\xi_{n+1} = \alpha \xi_n + \phi_n, \qquad n = 1, 2$$

where the quadruple $\{x_1, \xi_1, \phi_1, \phi_2\}$ is a set of independent, zero-mean, second-order random variables, and α is a scalar parameter, with $|\alpha| < 1$.

The expected cost to be minimized for this finite-horizon problem is:

$$J(\gamma_{11}, \gamma_{12}, \gamma_{13}; \gamma_{21}) = E\{(x_4)^2 + (x_3)^2 + (x_2)^2 + (u_{13} - a_1\xi_3)^2 \\ +(u_{21} - a_2\xi_3)^2 + (u_{12} - \xi_2)^2 + (u_{11} - \xi_1)^2\} \tag{27}$$

where $0 < a_1 < 1$, $0 < a_2 < 1$, $a_1 + a_2 = 1$, and γ_{in} is the policy of user i at stage n, which is taken to be a general Borel-measurable function according to:

$$\begin{aligned} u_{11} &= \gamma_{11}(\xi_1, x_1) =: \gamma_{11}(I_1) \\ u_{12} &= \gamma_{12}(\xi_2, \xi_1, x_2, x_1; u_{21}, u_{11}) =: \gamma_{12}(I_2) \\ u_{13} &= \gamma_{13}(\xi_3, \xi_2, \xi_1, x_3, x_2, x_1; u_{21}, u_{11}, u_{12}) =: \gamma_{13}(I_3) \\ u_{21} &= \gamma_{21}(\xi_1, x_1) =: \gamma_{21}(I_1). \end{aligned} \tag{28}$$

Note that user 1 has access to perfect measurement of all the current and past variables and the past controls, while user 2 has access to only the pair (ξ_1, x_1).

Optimal controller for the delay-free problem

As a benchmark performance, let us first determine the optimal controller for both users under the assumption that user 2 does not have any action delay, or equivalently that the construction of u_{21} is based on the same information as that of u_{13}. Then, this is a standard linear-quadratic stochastic control problem, with a two-dimensional control at stage three, which admits the solution:

$$\begin{aligned} \gamma_{13}^{\mathrm{DF}}(I_3) = \gamma_{21}^{\mathrm{DF}}(I_3) &= -(1/3)x_3 + a_1\xi_3 \\ \gamma_{12}^{\mathrm{DF}}(I_2) &= -(4/7)x_2 + \xi_2\,; \quad \gamma_{11}^{\mathrm{DF}}(I_1) = -(11/18)x_1 + \xi_1 \end{aligned} \tag{29a}$$

with the corresponding unique minimum value for J being:

$$J_{\mathrm{DF}} = \frac{11}{18}\mathrm{var}\ (x_1) \tag{29b}$$

where var $(x_1) := E[x_1^2]$. These correspond to the controllers (21) introduced earlier.

An important point to emphasize here is that, as mentioned earlier, even though the minimum value of J as given above is unique, the solution given is not unique as a control policy, as some other representation of it, such as [3]

$$\tilde{\gamma}_{13}(I_3) = -\frac{1}{3}x_3 + a_1\xi_3 + \kappa_1\left[u_{21} + \frac{1}{3}x_3 - a_2\xi_3\right] \tag{30a}$$

$$\gamma_{21}^{\mathrm{DF}}(I_3) = -\frac{1}{3}x_3 + a_2\xi_3 \tag{30b}$$

$$\gamma_{12}^{\mathrm{DF}}(I_2) = -\frac{4}{7}x_2 + \xi_2 + \kappa_2\left[u_{21} + \frac{1}{3}x_3 - a_2\xi_3\right] \tag{30c}$$

$$\gamma_{11}^{\mathrm{DF}}(I_1) = -\frac{11}{18}x_1 + \xi_1 \tag{30d}$$

would also constitute a solution for every choice of the pair (κ_1, κ_2). This is an important observation because as indicated earlier if one wishes to obtain a certainty-equivalent controller, with the delay-free optimal controller being the starting point, whether one uses representation (29a) or representation (30a)-(30d) with nonzero κ's, will make a difference as we will shortly see.

Optimal controller with action delays

To obtain the solution to the original problem, where user 2 has a 2-step action delay, we can use a dynamic programming approach — though a nonstandard one. We first minimize J, given by (27), with respect to γ_{13}, which leads to

$$\gamma_{13}^{*}(I_3) = -\frac{1}{2}x_3 - \frac{1}{2}u_{21} + \left(a_1 + \frac{1}{2}a_2\right)\xi_3. \tag{31a}$$

Next we substitute this back into (27) and minimize the resulting expression with respect to γ_{12}, leading to

$$\gamma_{12}^{*}(I_2) = -\frac{3}{5}x_2 - \frac{1}{5}u_{21} + \left(1 + \frac{a_2\alpha}{5}\right)\xi_2. \tag{31b}$$

Now, the final step is to minimize J with respect to γ_{11} and γ_{21} (jointly or sequentially), under (31a) and (31b), which leads to the unique solution

$$\gamma_{11}^{*}(I_1) = -\frac{11}{18}x_1 + \xi_1 \tag{31c}$$

$$\gamma_{21}^{*}(I_1) = -\frac{1}{18}x_1 + a_2\alpha^2\xi_1, \tag{31d}$$

[3] This corresponds to the class of controllers (23), with $\kappa_{21}^{13} = \kappa_1, \kappa_{22}^{12} = \kappa_2$, and all other κ's *zero*.

and to the unique optimal cost value:

$$J^* = \frac{11}{18} \text{ var } (x_1) + \frac{3}{2}(a_2)^2 \text{ var } (\phi_2) + \frac{7}{5}\alpha^2(a_2)^2 \text{ var } (\phi_1) \tag{32}$$

where var $(\phi_n) := E[\phi_n^2]$.

Note that the loss in performance in going from the delay-free optimal controller (or the optimal controller with perfect anticipation of the future) to the optimal controller that is compatible with the information structure of the original problem depends only on var (ϕ_1) and var (ϕ_2), and not on var (x_1). This feature could signal some certainty-equivalence property to be associated with this controller – – which is indeed the case provided that an appropriate representation of (29a) is used.

First, note that (31a) is precisely (30a) with $\kappa_1 = -\frac{1}{2}$, and hence γ_{13}^* is indeed part of an optimal solution for the delay-free version of the problem. Also note that $\gamma_{11}^* \equiv \gamma_{11}^{\text{DF}}$, and

$$\gamma_{21}^*(I_1) = E\left[\gamma_{21}^{\text{DF}}(I_3) \mid I_1, \gamma_{21}^{\text{DF}}, \gamma_{11}^{\text{DF}}\right]$$

where the latter holds (with $E[\cdot|\cdot]$ denoting conditional expectation) because

$$E\left[\xi_3 \mid I_1\right] = \alpha^2 \xi_1$$

and

$$E\left[x_3 \mid I_1, u_{12} = \gamma_{12}^{\text{DF}}(I_2), u_{11} = \gamma_{11}^{\text{DF}}(I_1)\right] = \frac{1}{6} x_1.$$

This says that γ_{21}^* is also "certainty equivalent," because it is the conditional mean of γ_{21}^{DF} given the information available for the construction of the controller of user 2 and given the forms of the delay-free optimal controllers of user 1. Now, finally to see the certainty-equivalence property of γ_{12}^*, start with the representation (30c) of γ_{12}^{DF} with $\kappa_2 = -(1/5)$:

$$\tilde{\gamma}_{12}(I_3) = -\frac{4}{7}x_2 + \xi_2 - \frac{1}{5}\left(u_{21} + \frac{1}{3}x_3 - a_2\xi_3\right)$$

and take its conditional mean with respect to I_2:

$$E\left[-\frac{4}{7}x_2 + \xi_2 - \frac{1}{5}\left(u_{21} + \frac{1}{3}x_3 - a_2\xi_3\right) \middle| I_2, \gamma_{12}^{\text{DF}}\right]$$

$$
\begin{aligned}
&= -\frac{4}{7}x_2 + \xi_2 - \frac{1}{5}u_{21} - \frac{1}{35}x_2 + \frac{1}{5}a_2\alpha\xi_2 \\
&= -\frac{3}{5}x_2 - \frac{1}{5}u_{21} + \left(1 + \frac{a_2\alpha}{5}\right)\xi_2
\end{aligned}
\tag{33}
$$

where we have used the property that

$$
E\left[x_3 \mid I_2, \gamma_{12}^{\mathrm{DF}}\right] = \frac{3}{7}x_2.
$$

Note that (33) is identical with (31b), which makes it a certainty-equivalent controller.

Other certainty-equivalent controllers

Even though the optimal controller (31a)-(31d) is a certainty-equivalent controller, it does not have an obvious structure that can be obtained readily as a representation of the optimal controller for the delay-free version of the problem, unless one goes through the actual derivation. There are, of course, other certainty-equivalent controllers, which can be derived directly from (29a) or some other fixed representation of it (such as (30a)-(30d)); these all will, however, necessarily lead to a performance worse than (32). For illustration purposes, and as a prelude to the presentation in the next section, let us present here two such controllers.

The first one (which we will call Controller I) uses the original representation (29a). All we need to do is compute

$$
\gamma_{21}^{I}(I_1) = E\left[\gamma_{21}^{\mathrm{DF}}(I_3) \mid I_1, \gamma_{1n}^{\mathrm{DF}}, n = 1,2,3\right] \;=\; -\frac{1}{18}x_1 + a_2\alpha^2\xi_1,
\tag{34a}
$$

and pick

$$
\gamma_{1n}^{\mathrm{I}} \equiv \gamma_{1n}^{\mathrm{DF}}, \quad n = 1,2,3.
\tag{34b}
$$

Under (34a) and (34b), the cost is

$$
J_{\mathrm{I}} = \frac{11}{18}\ \mathrm{var}\ (x_1) + 2(a_2)^2\left[\ \mathrm{var}\ (\phi_2) + \alpha^2\ \mathrm{var}\ (\phi_1)\right].
\tag{34c}
$$

The second one (which we will call Controller II) uses the representation (30a)-

(30d), with $\kappa_1 = -\frac{1}{2}, \kappa_2 = 0,$[4] and hence the only difference between this one and Controller I is that γ_{13} is now

$$\gamma_{13}^{\text{II}}(I_3) = -\frac{1}{2}x_3 - \frac{1}{2}u_{21} + \left(a_1 + \frac{1}{2}a_2\right)\xi_3, \tag{35a}$$

while

$$\gamma_{12}^{\text{II}} \;\equiv\; \gamma_{12}^{\text{I}}, \gamma_{11}^{\text{II}} \;\equiv\; \gamma_{11}^{\text{I}}, \gamma_{21}^{\text{II}} \;\equiv\; \gamma_{21}^{\text{I}}. \tag{35b}$$

The corresponding value of J is

$$J_{\text{II}} = \frac{11}{18} \text{ var } (x_1) + \frac{3}{2}(a_2)^2[\text{ var } (\phi_2) + \alpha^2 \text{ var } (\phi_1)]. \tag{35c}$$

Note that with var $(\phi_1) > 0$, $\alpha \neq 0$, $a_2 \neq 0$, we have the strict ordering

$$J_{\text{I}} > J_{\text{II}} > J^* > J_{\text{DF}}. \tag{36}$$

4 Two Certainty-Equivalent Controllers

We now return to the original infinite-horizon stochastic optimal control problem, and first present two certainty-equivalent controllers which belong to the class represented by (25), and in a sense constitute two extreme cases in that class; they correspond to the two controllers I and II presented in the previous section. Subsequently, in the next section, we show that both these controllers lead to a stable queue dynamics.

4.1 Controller I

This is obtained from (25) by taking the κ's equal to *zero*, and hence in a sense is the *simplest* certainty-equivalent controller:

$$u_{m,n}^* = -p_m\, \hat{x}_{n+d_m|n} + a_m\, \hat{\xi}_{n+d_m|n}\,, \quad m = 1, \ldots, M\,. \tag{37}$$

Here p_m is as defined by (19), and $\hat{x}_{n+d_m|n}$, $\hat{\xi}_{n+d_m|n}$ are the predicted values of x_{n+d_m} and ξ_{n+d_m}, respectively, based on the information I_n, and given that all other controllers are also in the form (37). These predictors are generated by

[4] The motivation behind this particular choice will become clear in the next section.

$$\hat{x}_{n+j|n} = \hat{x}_{n+j-1|n} + \sum_{i=1}^{M} \hat{u}_{m,n-d_m+j-1|n} - \hat{\xi}_{n+j-1|n} \,, \quad j \geq 1 \,; \tag{38}$$

$$\hat{x}_{n|n} = x_n$$

$$\hat{\xi}_{n+j|n} = \sum_{i=1}^{p} \alpha_i \, \hat{\xi}_{n+j-i|n} \,, \quad j \geq 1 \,; \quad \hat{\xi}_{n-k|n} = \xi_{n-k} \,, \quad k \geq 0 \,, \tag{39}$$

and

$$\hat{u}_{m,n-d_m+j-1|n} := \begin{cases} -p_m \hat{x}_{n+j-1|n} + a_m \hat{\xi}_{n+j-1|n} & \text{if } j \geq d_m + 1 \\ \\ u_{m,n-d_m+j-1} & \text{if } j < d_m + 1 \end{cases} \tag{40}$$

These are the recursive equations generating the predictors for the the queue length and rate information at a future time, where the future time is the current time plus the action delay for the corresponding source. For example, $\hat{\xi}_{n+j|n}$ denotes the predicted value at time n of the value of ξ at some future time $n+j$, based on the information available at time n, which is I_n. A similar interpretation holds for $\hat{x}_{n+j|n}$.

The above algorithm is relatively easy to implement. The estimator algorithms are simple scalar operations and the scalar solution of the Riccati equation s has already been obtained explicitly.

In summary, an easily implementable version of Controller I is given below in the form of a pseudo-code:

Pseudo-code for the node's computation at time n using Controller I

for $j = 1$ to d_M do

for $m = 1$ to M do

if $(n + j - d_m - 1 \geq n)$

$\hat{u}_{m,n+j-d_m-1|n} = -p_m \hat{x}_{n+j-1} + a_m \hat{\xi}_{n+j-1}$

end

$\hat{x}_{n+j} = \hat{x}_{n+j-1} + \sum_{m=1}^{M} \hat{u}_{m,n+j-1+d_m|n} - \hat{\xi}_{n+j-1}$

end

$u_{mn} = -p_m \hat{x}_{n+d_m} + a_m \xi_{n+d_m}$ ◇

4.2 Controller II

The second controller we consider here is again obtained from (25), now by assigning some specific nonzero values to the κ's. These values are picked in such a way that at each stage controls that will experience smaller delay are made to depend in a "robust way" on controls that will experience larger delay. This is accomplished by applying a Cholesky decomposition (Sage and White, 1977) to the positive-definite matrix $(R+bb's)$. In particular, we have

$$(R+bb's)=U'U\,,$$

where U is a non-singular, upper-triangular matrix. In view of this, (18) can be rewritten as

$$U\tilde{u}_n=-\tilde{x}_n, \tag{41}$$

where $\tilde{x}_n := (U')^{-1}bsx_n$. Let U_{ij} represent the $(i,j)^{\text{th}}$ element of U and q_m represent the m^{th} element of $(U')^{-1}bs$. Then, it is easy to see that (41) can be written as

$$\tilde{u}_{mn}=-\frac{1}{U_{mm}}\left(\sum_{k=m+1}^{M}U_{mk}\,\tilde{u}_{kn}+q_m\,x_n\right), \tag{42}$$

or, equivalently, using the relationship between $\tilde{u}$ and u (see (18)):

$$\begin{aligned}u_{mn}=&-\frac{1}{U_{mm}}\left(\sum_{k=m+1}^{M}U_{mk}\,u_{kn}+q_m x_n\right)\\&+\left(a_m+\frac{1}{U_{mm}}\sum_{k=m+1}^{M}U_{mk}a_k\right)\xi_n.\end{aligned} \tag{43}$$

Now, to incorporate the action delays into this expression, we simply replace u_{kn} by $u_{k,n-d_k}$, thus arriving at:

$$\begin{aligned}u_{m,n-d_m}=&-\frac{1}{U_{mm}}\left(\sum_{k=m+1}^{M}U_{mk}\,u_{k,n-d_k}+q_m\,x_n\right)\\&+\left(a_m+\frac{1}{U_{mm}}\sum_{k=m+1}^{M}U_{mk}a_k\right)\xi_n.\end{aligned} \tag{44}$$

The certainty-equivalent Controller II then simply replaces x_n and ξ_n in the expression for $u_{m,n-d_m}$ by their predicted values based on information I_{n-d_m}, while using the new representation (44) for the controllers. Shifting the time forward by d_m units, we arrive at:

$$u^{**}_{mn} = -\frac{1}{U_{mm}}\left(\sum_{j=m+1}^{M} U_{mj} u_{j,n+d_m-d_j} + q_m \hat{x}_{n+d_m|n}\right) + \left(a_m + \frac{1}{U_{mm}}\sum_{k=m+1}^{M} U_{mk}\, a_k\right)\hat{\xi}_{n+d_m|n} \tag{45}$$

where $\hat{x}$, $\hat{\xi}$ are generated by

$$\hat{x}_{n+j|n} = \hat{x}_{n+j-1|n} + \sum_{i=1}^{M} \hat{u}_{m,n-d_m+j-1|n} - \hat{\xi}_{n+j-1|n}\,, \quad j \geq 1\,; \tag{46}$$

$$\hat{x}_{n|n} = x_n$$

$$\hat{\xi}_{n+j|n} = \sum_{i=1}^{p} \alpha_i\, \hat{\xi}_{n+j-i|n}\,, \quad j \geq 1\,; \quad \hat{\xi}_{n-k|n} = \xi_{n-k}\,, \quad k \geq 0\,, \tag{47}$$

$$\tag{48}$$

and

$$\hat{u}_{m,n-d_m+j-1|n} :=$$

$$\begin{cases} -\dfrac{1}{U_{mm}}\left(\displaystyle\sum_{k=m+1}^{M} U_{mk}\, \hat{u}_{k,n-d_k+j-1} + q_m\, \hat{x}_{n+j-1|n}\right) \\ \qquad + \left(a_m + \dfrac{1}{U_{mm}}\displaystyle\sum_{k=m+1}^{M} U_{mk}\, a_k\right)\hat{\xi}_{n+j-1|n} & \text{if } j \geq d_m + 1 \\ \qquad\qquad u_{m,n-d_m+j-1} & \text{if } j < d_m + 1 \end{cases} \tag{49}$$

The *hatted* terms above admit the same interpretations as in the corresponding cases for Controller I, and hence will not be repeated here. We note, however, that even though the implementation of Controller II requires the computation of a Cholesky decomposition of the matrix $(R + bb's)$, this computation can be carried out explicitly due to the special structure of the problem. It is not difficult to see that the expression for the $(i,i)^{th}$ entry of U is:

$$U_{ii} = \sqrt{\left(1 + s\sum_{\ell=1}^{i} c_\ell^2\right) / \left(c_i^2(1 + s\sum_{\ell=1}^{i-1} c_\ell^2)\right)}\,,$$

and the one for the $(i,j)^{th}$ entry, $j > i$, is:

$$U_{ij} = sc_i / \sqrt{(1 + s\sum_{\ell=1}^{i} c_\ell^2)(1 + s\sum_{\ell=1}^{i-1} c_\ell^2)}\,.$$

Since U is upper triangular, $U_{ij} = 0$ for $j < i$. Further, $q_m = U_{mj}$, for any $j > m$.

For convenience, the pseudo-code for Controller II is provided below.

Pseudo-code for the node's computation at time n using Controller II

for $j = 1$ to d_M do

for $m = 1$ to M do

if $(n + j - d_m - 1 \geq n)$ compute $u_{m,n+j-d_m-1|n}$ using (49).

end

$$\hat{x}_{n+j} = \hat{x}_{n+j-1} + \sum_{m=1}^{M} \hat{u}_{m,n+j-1+d_m|n} - \hat{\xi}_{n+j-1}$$

end

$$\begin{aligned} u_{mn} = & -\frac{1}{U_{mm}}\left(\sum_{j=m+1}^{M} U_{mj}u_{jn} + q_m\hat{x}_{n+d_m}\right) \\ & + \left(a_m + \frac{1}{U_{mm}}\sum_{j=m+1}^{M} U_{mj}a_j\right)\hat{\xi}_{n+d_m}. \end{aligned}$$

$\diamond$

The pseudo-codes for the two controllers show that the computations involve only scalar additions and multiplications. This is in contrast to the optimal solution which involves the solution of a $(d_M+p+1 \times d_M+p+1)$-dimensional Riccati equation, in addition to the scalar additions and multiplications. Further, when a new source becomes active or an active source becomes inactive, the Riccati equation has to be solved again to implement the optimal controller. The suboptimal controllers described here do not require this. However, the expressions for the estimators (the *hatted* terms) have to be updated when the number of sources change. But this computation is simple as is evident from the pseudo-codes.

5 Stability via the Certainty-Equivalent Controllers

In this section, we show that the linearized queue dynamics are stable under both Controllers I and II, i.e., both controllers result in a bounded cost. Consider the original system with the following modified N-stage cost function, along with a cost at the terminal state:

$$J_N = s(x_{N+1})^2 + \sum_{n=1}^{N} \left[(x_n)^2 + \sum_{m=1}^{M} \frac{1}{c_m^2} (u_{m,n-d_m} - a_m\, \xi_n)^2 \right]. \tag{50}$$

Let

$$\zeta_{mn} := u_{m,n-d_m} - a_m\, \xi_n\,; \quad \zeta_n := (\zeta_{1n}, \ldots, \zeta_{Mn})'\,.$$

Then, using a "completion-of-squares" argument, J_N can be rewritten as

$$J_N = \sum_{n=1}^{N} |\zeta_n + s(R + bb's)^{-1} b x_n|^2_{(sbb'+R)} + sx_1^2. \tag{51}$$

Defining

$$L := sbb' + R; \qquad k := s(sbb' + R)^{-1} b\,,$$

J_N can be expressed as

$$J_N = sx_1^2 + \sum_{n=1}^{N} \left\{ \sum_{m=1}^{M} (\zeta_{mn} - a_m \xi_n - k_m x_n)^2 L_{mm} \right.$$
$$\left. + 2 \sum_{m=1}^{M} \sum_{j=1}^{m-1} L_{mj} (\zeta_{mn} - a_m \xi_n + k_m x_n)(\zeta_{jn} - a_j \xi_n + k_j x_n) \right\} \tag{52}$$

where L_{mj} is the $(m,j)^{\text{th}}$ element of the matrix L and k_m is the m^{th} element of the vector k.

Substituting the expression for Controller I in (52), we get

$$J_N = sx_1^2 + \sum_{n=1}^{N} \left\{ \sum_{m=1}^{M} (a_m\hat{\xi}_n^{(m)} + k_m\hat{x}_n^{(m)} - a_m\xi_n - k_m x_n)^2 L_{mm} \right.$$
$$+2\sum_{m=1}^{M}\sum_{j=1}^{m-1} L_{mj}(a_m\hat{\xi}_n^{(m)} + k_m\hat{x}_n^{(m)} - a_m\xi_n + k_m x_n)$$
$$\left. \times (a_j\hat{\xi}_n^{(j)} + k_m\hat{x}_n^{(j)} - a_j\xi_n + k_j x_n) \right\} \tag{53}$$

where $\hat{\xi}_n^{(m)} := \hat{\xi}_{n|n-d_m}$, $m = 1, \ldots, M$. and $\hat{x}_n^{(m)} := \hat{x}_{n|n-d_m}$. We recall that

$$\hat{\xi}_n^{(m)} = E(\xi_n | I_{n-d_m}); \qquad \hat{x}_n^{(m)} = E(x_n | I_{n-d_m}),$$

where the latter uses the given structure of the controllers. Since

$$\xi_n = \sum_{i=1}^{p} \alpha_i \, \xi_{n-i} + \phi_{n-1},$$

it follows that

$$\begin{aligned} \xi_n &= f_{\xi,m}(\xi_{n-d_m}, \xi_{n-d_m-1}, \ldots, \xi_{n-d_m-p+1}) \\ &\quad + f_{\phi,m}(\phi_{n-1}, \phi_{n-2}, \ldots, \phi_{n-d_m-1}), \end{aligned} \tag{54}$$

and

$$\hat{\xi}_n^{(m)} = f_{\xi,m}(\xi_{n-d_m}, \xi_{n-d_m-1}, \ldots, \xi_{n-d_m-p+1}), \tag{55}$$

where $f_{\xi,m}$ and $f_{\phi,m}$ are linear functions. Thus,

$$\xi_n - \hat{\xi}_n^{(m)} = f_{\phi,m}(\phi_{n-1}, \phi_{n-2}, \ldots, \phi_{n-d_m-1}). \tag{56}$$

In other words, $(\xi_n - \hat{\xi}_n^{(m)})$ is a linear function of a finite number of primitive random variables. Using a similar argument, it is also possible to show that, $\forall l$ such that $n - d_m < l \leq n$,

$$\hat{\xi}_{l|n-d_m} = f_{\phi,m}^{(l)}(\phi_{n-1}, \phi_{n-2}, \ldots, \phi_{n-d_m-1}),$$

where $f_{\phi,m}^{(l)}$ is some linear function.

Next, we show that $(x_n - \hat{x}_n^{(m)})$ is also a linear function of a finite number of primitive random variables. For any $m \in \{1, 2, \ldots, M\}$, $n - d_m + 1 \leq l \leq n$, we can write x_ℓ as

$$x_l = g_{u,m,l}(\{u_{1k}, u_{2k}, \ldots, u_{m-1,k};\ n - d_m \leq k \leq l\})$$

$$+h_{u,m,l}(\{u_{mk}, u_{m+1,k}, \ldots, u_{Mk}\}, n - d_m \leq k \leq l\}) \tag{57}$$

$$+f_{\xi,m,l}(\xi_{n-d_m}, \xi_{n-d_m-1}, \ldots, \xi_{n-d_m-p}) + \tilde{f}_{\phi,m,l}(\phi_l, \phi_{l-1}, \ldots, \phi_{n-d_m}),$$

where f, g, h and $\tilde{f}$ are all linear functions. Thus, for $m = 1$,

$$\hat{x}_{l|l-d_1} = h_{u,1,l}(\{u_{1k}, u_{2,k}, \ldots, u_{Mk}\}, n - d_1 \leq k \leq l\})$$

$$+\tilde{f}_{\xi,1,l}(\xi_{n-d_1}, \xi_{n-d_1-1}, \ldots, \xi_{n-d_1-p})$$

and

$$x_n - \hat{x}_n^{(1)} = \tilde{f}_{\phi,1,n}(\phi_n, \phi_{n-1}, \ldots, \phi_{n-d_1}).$$

In fact, it is easy to see that $\tilde{f}_{\phi,1,n}$ is independent of n, which we will denote as $\tilde{f}_{\phi,1}$. By an induction on m, one can then show that

$$x_n - \hat{x}_n^{(m)} = \tilde{f}_{\phi,m}(\phi_n, \phi_{n-1}, \ldots, \phi_{n-d_m}), \tag{58}$$

for some linear functions $\tilde{f}_{\phi,m}$. Thus, from (56) and (58), it follows that, there exists a $C > 0$ such that

$$E\left((a_m\hat{\xi}_n^{(m)} + K_m\hat{x}_n^{(m)} - a_m\xi_n - K_m x_n)^2\right) \leq C, \tag{59}$$

independent of m, n. Define $\vartheta_n := a_m\xi_n + K_m x_n$ and $\hat{\vartheta}_m^{(m)} := E[a_m\xi_n + K_m x_n | I_{n-d_m}]$. For $j < m$,

$$\begin{aligned} E[(\hat{\vartheta}_n^{(m)} - \vartheta_n)(\hat{\vartheta}_n^{(j)} - \vartheta_n)] &= E[(\hat{\vartheta}_n^{(m)}\hat{\vartheta}_n^{(j)} - \vartheta_n\hat{\vartheta}_n^{(j)} - \vartheta_n\hat{\vartheta}_n^{(m)} + \vartheta_n^2)] \\ &= E[(\hat{\vartheta}_n^{(m)})^2] - E[(\hat{\vartheta}_n^{(j)})^2] - E[(\hat{\vartheta}_n^{(m)})^2] + E[(\vartheta_n^2)] \\ &= E[(\vartheta_n - \hat{\vartheta}_n^{(j)})^2]\,. \end{aligned}$$

Thus, from (53),

$$J_N \leq C \sum_{n=1}^{N} \left(\sum_{m=1}^{M} L_{mm} + 2 \sum_{m=1}^{M} \sum_{j=1}^{m-1} L_{mj} \right) = N\tilde{C},$$

where

$$\tilde{C} := C\left(\sum_{m=1}^{M} L_{mm} + 2\sum_{m=1}^{M}\sum_{j=1}^{m-1} L_{mj}\right).$$

Therefore,

$$J \leq \lim_{N\to\infty} \frac{J_N}{N} = \tilde{C},$$

which proves that Controller I is stabilizing.

The proof for any other certainty-equivalent controller of the form given in (24) is similar. Instead of working with the form of J_N given by (51), we start by first rewriting of $(R + sbb')$ using a Hermitian similarity transformation of the form $V'PV$, where V is an upper-triangular matrix. In other words, rewrite J_N as

$$J_N = \sum_{n=1}^{N}(\zeta_n + s(R + bb's)^{-1}bx_n)'V'PV(\zeta_n + s(R + bb's)^{-1}bx_n) + sx_1^2.$$

Controller II is a special case of such a transformation with $P = I$. With minor modifications, the rest of the stability proof for all certainty-equivalent controllers follows as for Controller I. In addition to the type of terms in the cost for Controller I, the cost for a general certainty-controller has additional terms of the form

$$E(x_n - \hat{x}_{n|n-d_m})(x_k - x_{k|n-d_m}),$$

where $n < k < n - d_1 + d_M$. Since k is upper-bounded, it is easy to see that the above term is also upper-bounded and stability is assured for all certainty-equivalent controllers.

6 Simulation Results

To evaluate the performance of Controllers I and II, we consider a single congested node accessed by three sources as in Figure 1 in Section 1. Source 1 is subject to no action delay, Source 2 experiences an action delay of 5 time units and Source 3 experiences an action delay of 10 time units. The fairness indices are taken to be $a_1 = a_2 = a_3 = 1/3$. The AR process is assumed to be of order 2, and the parameters of the process are $\alpha_1 = \alpha_2 = 0.4$. The driving

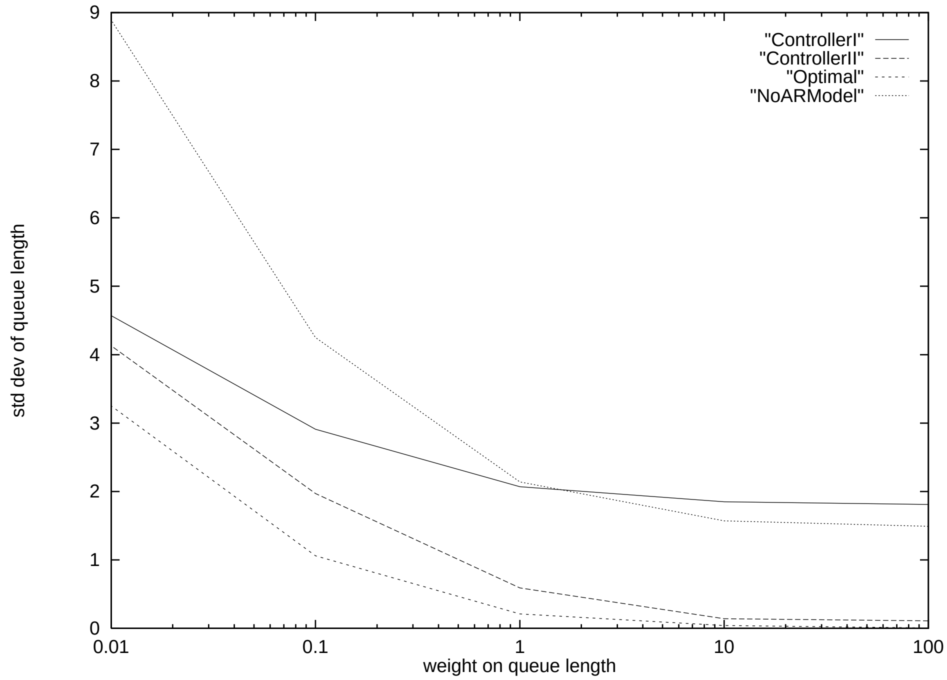

Fig. 2. Standard deviation of the queue length, $\sqrt{E(x^2)}$ as a function of the weight c^2.

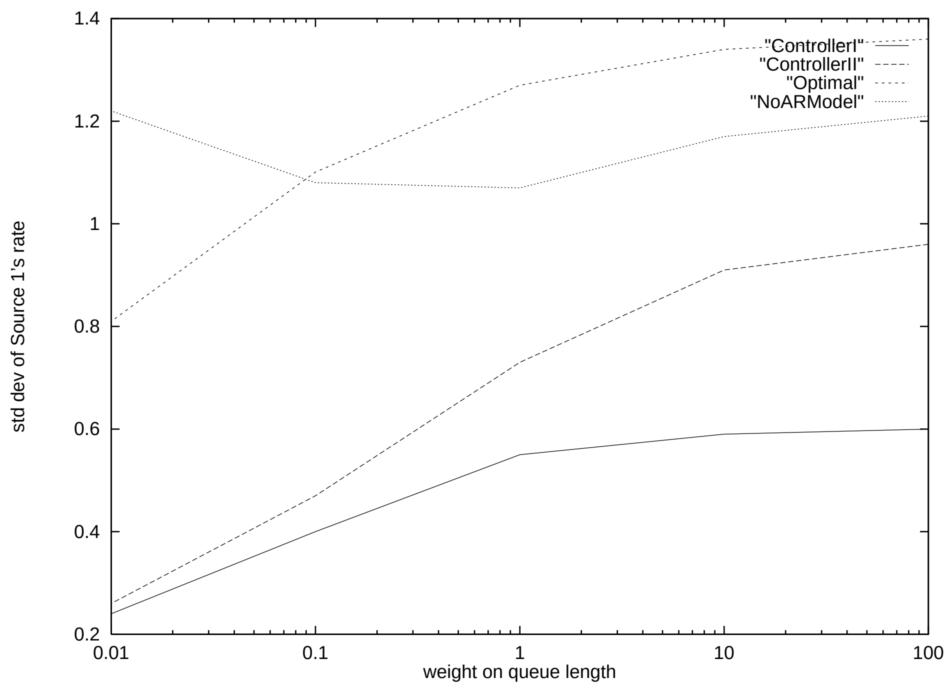

Fig. 3. Standard deviation of Source 1's rate, $\sqrt{E(\tilde{u}_1^2)}$ as a function of the weight c^2.

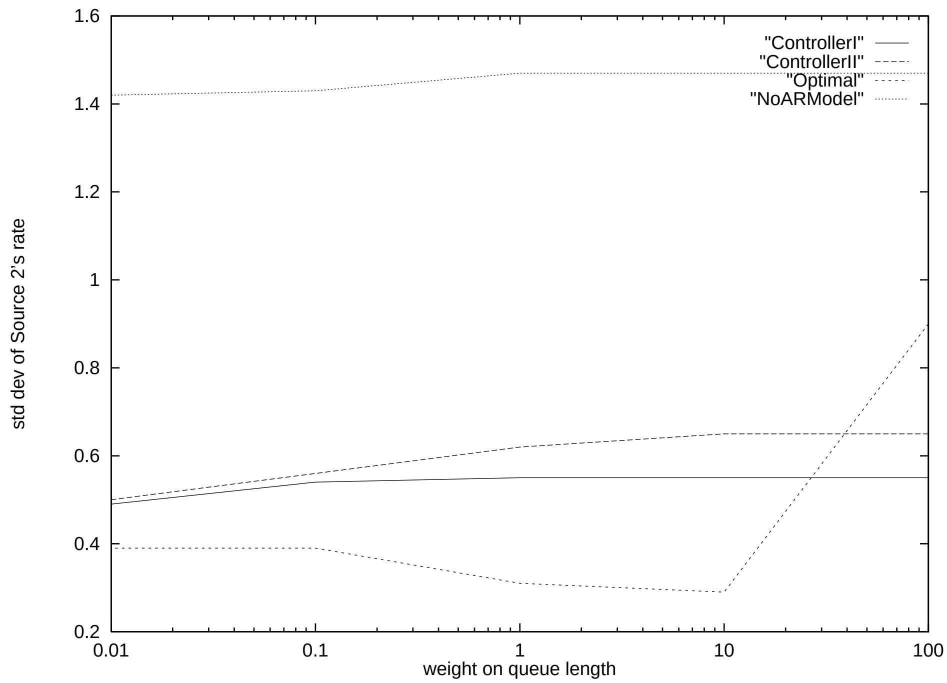

Fig. 4. Standard deviation of Source 2's rate, $\sqrt{E(\tilde{u}_2^2)}$ as a function of the weight c^2.

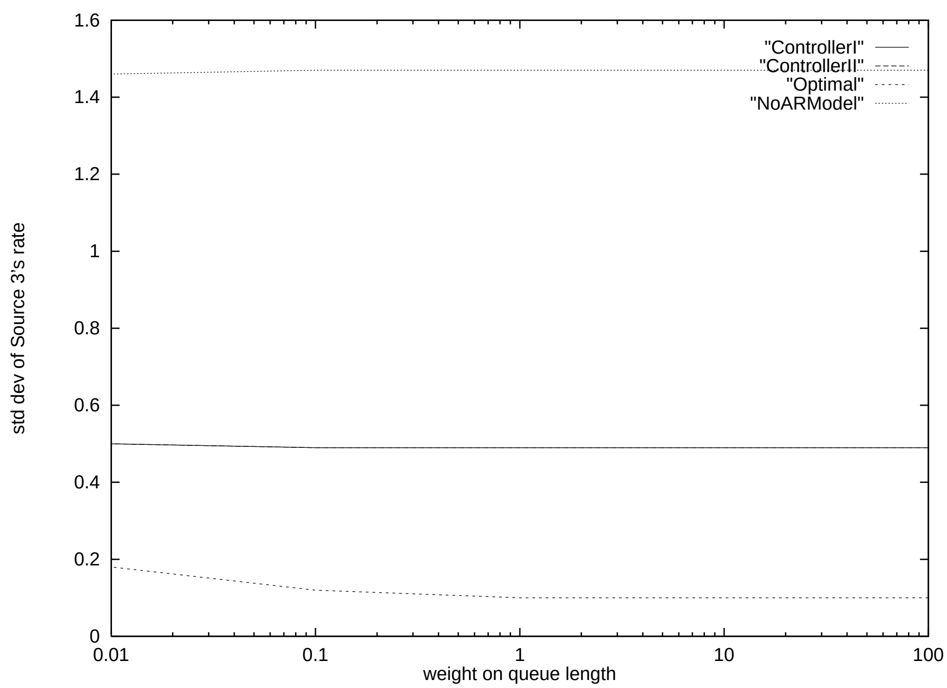

Fig. 5. Standard deviation of Source 3's rate, $\sqrt{E(\tilde{u}_3^2)}$ as a function of the weight c^2.

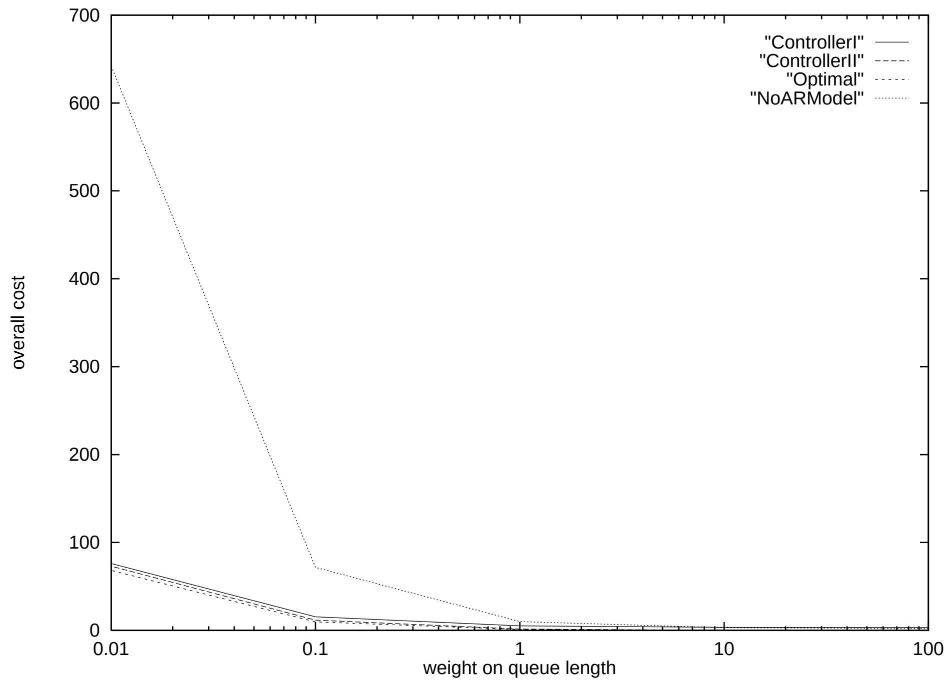

Fig. 6. Overall cost J as a function of the weight c^2.

zero-mean Gaussian white noise process has a variance equal to 1. Figures 2 and 6 depict the performances of the two controllers for various values of the parameter c^2, where $c^2 = c_m^2$, $m = 1, 2, 3$.

For the purposes of comparison, we have also plotted the results using the optimal controller derived in (Imer and Başar, 1999b). In addition, we have also compared the results from a controller which ignores the AR model and simply stabilizes the system using standard LQ theory. For each value of c^2 and for each controller, the simulation was run for $100,000$ time units starting with an initial queue length of 10. The figures show that Controllers I and II perform significantly better than the simple stabilizing controller, especially in regulating the queue length, and their performances are close to the performance of the optimal controller. Evidently, modelling the available bandwidth results in significant improvement in the performance. In the results shown here, as well as many other simulation experiments that we have performed, Controller II often performs better than Controller I. We do not believe, however, that this result will generally hold, since neither controller is optimal. The figures also show that one can control the queue length fairly tightly around a nominal level Q, by increasing c_i's. Also, as expected, the mean values of x, u_1, u_2 and u_3 were all observed to be close to zero. This shows that the controller is fair in the long-term. The fact that the variances of u_i for the three sources are of approximately the same magnitude also shows that the fluctuation from fair allocation is about the same for all sources, despite the different action delays.

7 Discussion and Conclusions

We have presented two algorithms for rate control that are derived using tools from stochastic control theory. They are easy to implement, carry an appealing certainty-equivalence property, and are stabilizing. They take into account the presence of different delays for different users, and are actually implemented at a centralized entity, namely, the node.

There are several avenues of further research that we are currently investigating:

- Measurement interval: We considered a discrete-time problem in which the basic time unit was some minimum interval over which each node in the network would compute the available rate. Very small measurement intervals would lead to poor estimates of the available rate, whereas large measurement intervals would lead to poor utilization. Further, since our model is a discrete-time model, if the measurement interval is large, queue length variations within the measurement interval may become important. Thus, the impact of this discretization is a potential topic for further research.
- RM (rate management) cells: Here we have assumed that each source receives feedback from the controller at the node at each time instant. Under the current ATM standards, the feedback is available only when the source generates an RM cell. The RM cell travels through the entire route of a virtual circuit (i.e., the route taken by the ABR source from source to destination) and collects feedback information along the path. More specifically, the RM cell is generated by the source and travels to the destination, where it is turned back to the source. In the reverse path, each node on the path sets the maximum allowable rate for the source. The minimum of these rates is used by the ABR source. An RM cell is generated every Nrm data cells, where Nrm is typically 32. The only exception is when the source's data rate becomes very low, in which case, other special mechanisms come into play. We believe that the controllers presented here can stabilize queue length even when the feedback mechanism is implemented using RM cells. In fact, since our discrete time unit is the measurement interval which will be typically longer than the time between two RM cells, we always expect that there will be at least one RM cell per time unit.
- Unknown number of active users: In the formulation here, we assumed that the number of active users is known. If this is unknown, then one can use some estimation scheme, such as the one in Kalyanaraman et al. (1997), to estimate the number of active users and use that as the parameter M in our formulation.
- Peak rate constraints and bursty sources: If a source cannot send at the rate at which the node asks it to, either due to peak rate constraints or simply because it is not currently generating traffic at that rate, then one can use

the ER field in the forward RM cell to signal to the node the maximum rate at which the source can currently transmit. Then, this rate could be subtracted from the available mean rate at the node, the number of active sources (M) can be reduced by one, and the algorithm can proceed as usual.

- AR process: We have assumed here that the parameters of the AR process are known. In practice, the node has to estimate this. Preliminary results indicate that the use of standard estimators such as recursive least-squares, in conjunction with our controller, performs well.
- Multi-node network model: In Compans (1998), many of the above modifications have been incorporated into the basic algorithm derived here. Simulation results on many network configurations indicate that, even when the assumption of linearity is violated, either due to the queue length hitting zero or due to the presence of multiple bottleneck nodes, the algorithm still leads to stable queue dynamics without much performance degradation (Compans, 1998).
- Unknown and random delay: ATM's signaling protocol allows one to estimate action delay at call set-up time. However, in addition to this, there could be additional variable delay introduced due to queueing which cannot be measured. The effect of this should be best studied in the context of a real network model, which we plan to do in a subsequent publication.

References

Addie, R. G., & Zukerman, M. (1993). An approximation for performance evaluation of stationary single server queues. In *Proceedings of the Conference on Computer Communications (IEEE INFOCOM)*, *2*, San Francisco, CA, pp. 835–842.

Ait-Hellal, O., Altman, E., & Başar, T. (1997). Rate-based flow control with bandwidth information. *European Transactions on Telecommunications*, *8*(1):55–65.

Altman, E., Başar, T., & Srikant, R. (1997). Multi-user rate-based flow control with action delays: A team-theoretic approach. In *Proc. 36th IEEE Conf. Decision and Control*, San Diego, CA.

Altman, E. , Başar, T., & Srikant, R. (1998). Robust rate control for ABR sources. In *Proc. IEEE INFOCOM*, San Francisco, CA.

Altman, E., Baccelli, F., & Bolot, J. C. (1994). Discrete-time analysis of adaptive rate control mechanisms. *High Speed Networks and their performance*, H.G. Perros and Y. Viniotis Eds., North Holland, pp. 121–140.

Altman, E., & Başar, T. (1995). Optimal rate control for high speed telecommunication networks. In *Proc. 34th IEEE Conf. Decision and Control*, pp. 1389–1394, New Orleans, LA.

Anderson, B.D.O., & Moore, J.B. (1990). *Optimal Control: Linear Quadratic Methods*, Prentice-Hall, N.J.

Bansal, R., & Başar, T. (1987). Solutions to a class of linear-quadratic-Gaussian (LQG) stochastic team problems with nonclassical information. *Systems and Control Letters*, pp. 125–130.

Benmohamed, L. & Meerkov, S. M. (1993). Feedback control of congestion in packet switching networks: The case of a single congested node. *IEEE/ACM Transactions on Networking*, *1*(6):693–707.

Benmohamed, L., & Wang, Y. T. (1998). A control-theoretic ABR explicit rate algorithm for ATM switches with per-VC queueing. In *Proceedings of IEEE INFOCOM*, San Francisco, CA.

Compans, S. (1998). *Flow Control for ABR Service in ATM Networks.* M.S. Thesis, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign.

Ho, Y.-C., & Chu, K.-C. (1972). Team decision theory and information structures in optimal control problems: Part I. *IEEE Transactions on Automatic Control*, *17*(1):15–21.

Imer, O.Ç., & Başar, T. (1999a). Constructive solutions to a decentralized team problem with application in flow control in networks. In D.E. Miller and L. Qiu, editors, *Topics in Control and Its Applications*, Springer Verlag (to appear).

Imer, O.Ç., & Başar, T. (1999b). Optimal solution to a team problem with information delays: An application in flow control for communication networks. *38th IEEE Conference on Decision and Control*, Phoenix, AZ.

Jain, R. (1996). Congestion control and traffic management in ATM networks: recent advances and a survey, *Computer Networks and ISDN Networks.*

Jacobson, V. (1988). Congestion avoidance and control. *ACM Computer Communication Review*, *18*:314–329.

Kalyanaraman, S., Jain, R., Fahmy, S., Goyal, R., & Vandalore, B. (1997). The ERICA switch algorithm for ABR traffic management in ATM networks. http://www.cis.ohio-state.edu/ jain/papers.

Keshav, S. (1997). *An Engineering Approach to Computer Networks.* Addison-Wesley, Reading, MA.

Kolarov A., & Ramamurthy, G. (1997). A control theoretic approach to the design of closed loop rate based flow control for high speed ATM networks. In *Proceedings of IEEE INFOCOM.*

Maglaris, B., Anastasiou, D., Sen, P., Karlsson, G., & Robbins, J. (1988). Performance models of statistical multiplexing in packet video communications. *IEEE Trans. on Communications*, *36*:834–844.

Mascolo, S., Cavendish, D., & Gerla, M. (1996). ATM rate based congestion control using a Smith predictor: an EPRCA implementation. In *Proceedings of IEEE INFOCOM*, San Francisco, CA.

Melamed, B., Raychaudhuri, D., Sengupta, B., & Zdepski, J. (1992). TES-based traffic modeling for performance evaluation of integrated networks. In *Proceedings of the Conference on Computer Communications (IEEE INFOCOM)*, *1*:75–84. Florence, Italy (1C.1).

Melamed, B., Raychaudhuri, D., Sengupta, B., & Zdepski, J. (1994). TES-Based video source modeling for performance evaluation of integrated networks. *IEEE Transactions on Communications*, *42*.

Ramamurthy, G., & Sengupta, B. (1993). A predictive hop-by-hop congestion control policy for high-speed networks. In *Proceedings of the Conference on Computer Communications (IEEE INFOCOM)*, *3*:1033–1041. San Francisco.

Sage, A.P., & White, C.C., (1977). Optimum Systems Control. *Prentice Hall*, Englewood Cliffs, NJ.

Srikant, R., & Başar, T. (1992). Asymptotic solutions to weakly coupled stochastic teams with nonclassical information. *IEEE Transactions on Automatic Control*, *37*(2):163–173.

Witsenhausen, H.S. (1968). A counterexample in stochastic optimal control. *SIAM Journal on Control*, *59*:131–147.

Witsenhausen, H.S. (1971). Separation of estimation and control for discrete-time systems. *Proceedings of the IEEE*, *59*:1557–1566.