Rate control for communication networks:
shadow prices, proportional fairness and stability *

F.P. Kelly, A.K. Maulloo and D.K.H. Tan
University of Cambridge, UK

This paper analyses the stability and fairness of two classes of rate control algorithm for communication networks. The algorithms provide natural generalizations to large-scale networks of simple additive increase/multiplicative decrease schemes, and are shown to be stable about a system optimum characterized by a proportional fairness criterion. Stability is established by showing that, with an appropriate formulation of the overall optimization problem, the network’s implicit objective function provides a Lyapunov function for the dynamical system defined by the rate control algorithm. The network’s optimization problem may be cast in primal or dual form: this leads naturally to two classes of algorithm, which may be interpreted in terms of either congestion indication feedback signals or explicit rates based on shadow prices. Both classes of algorithm may be generalized to include routing control, and provide natural implementations of proportionally fair pricing.

Keywords: ATM network, congestion indication, elastic traffic, Internet, Lyapunov function, proportionally fair pricing, queues, routing, tatonnement.

Introduction

The design and control of modern communication networks raise several issues well suited to study using techniques of operational research such as optimization, network programming and stochastic modelling. In this paper we illustrate this theme, through the presentation and analysis of a mathematical model that arises in connection with the development and deployment of large-scale broadband networks.

In future communication networks there are expected to be applications that are able to modify their data transfer rates according to the available bandwidth within the network. Traffic from such applications is termed elastic [1]; a typical current example is TCP traffic over the Internet [2], and future examples may

*Correspondence: Frank Kelly, Statistical Laboratory, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, UK. Email: f.p.kelly@statslab.cam.ac.uk.
include the controlled-load service of the Internet Engineering Task Force [3] and the Available Bit Rate transfer capability of ATM (asynchronous transfer mode) networks [4].

The key issue we address in this paper concerns how the available bandwidth within the network should be shared between competing streams of elastic traffic; in particular, we present a tractable mathematical model and use it to analyse the stability and fairness of a class of rate control algorithms. Traditionally stability has been considered an engineering issue, requiring an analysis of randomness and feedback operating on fast time-scales, while fairness has been considered an economic issue, involving static comparisons of utility. In future networks the intelligence embedded in end-systems, acting on behalf of human users, is likely to lessen the distinction between engineering and economic issues and increase the importance of an interdisciplinary view. (This general theme was the subject of the 1996 Blackett Memorial Lecture; further aspects are developed elsewhere [5].)

There is a substantial literature on rate control algorithms, recently reviewed by Hernandez-Valencia et al. [6]. Key early papers of Jacobson [2] and Chiu and Jain [7] identified the advantages of adaptive schemes that either increase flows linearly or decrease flows multiplicatively, depending on the absence or presence of congestion. Important recent papers of Bolot and Shankar [8], Fendick et al. [9] and Bonomi et al. [10] have analysed the stability of networks with a single bottleneck resource, where congestion is signalled by the build-up of a queue at the bottleneck’s buffer, and where propagation delays are significant. (In wide-area networks propagation times may be significant in comparison with queueing times: for a transatlantic link of 600 Megabits per second, ten million bits may be in flight between queues.) The framework we adopt in this paper is simpler than that analysed by these authors in that we directly model only rates and not queue lengths, but more complex in that we model a network with an arbitrary number of bottleneck resources. Theoretical work [11], [12] on queues serving the superposition of a large number of streams indicates circumstances when the busy period preceding a buffer overflow may be relatively short, and several authors have argued the advantages of preventing queue build-up through the bounding of rates (see, for example, Charny et al. [13]).

Any discussion of the performance of a rate control scheme must address the issue of fairness, since there exist situations where a given scheme might maximize network throughput, for example, while denying access to some users. The most commonly discussed fairness criterion is that of max-min fairness: loosely, a set of rates is max-min fair if no rate may be increased without simultaneously decreasing another rate which is already smaller. In a network with a single bottleneck resource max-min fairness implies an equal share of the resource for each flow through it. Mazumdar et al. [14] have pointed out that from a game-theoretic standpoint such an allocation is not special, and have advocated instead the Nash bargaining solution, from cooperative game theory, as capturing natural assumptions as to what constitutes fairness.
The need for networks to operate in a public (and therefore potentially non-cooperative) environment has stimulated work on charging schemes for broadband networks: see Kelly [15] for a scheme based on time and volume measurements for non-elastic traffic, MacKie-Mason and Varian [16] for a description of a ‘smart market’ based on a per-packet charge when the network is congested, and the collection edited by McKnight and Bailey [17] for several further papers and references. Kelly [18] describes a model for elastic traffic in which a user chooses the charge per unit time that the user is willing to pay; thereafter the user’s rate is determined by the network according to a proportional fairness criterion applied to the rate per unit charge. It was shown that a system optimum is achieved when users’ choices of charges and the network’s choice of allocated rates are in equilibrium. There remained the question of how the proportional fairness criterion could be implemented in a large-scale network. In this paper we show that simple rate control algorithms, using additive increase/multiplicative decrease rules or explicit rates based on resource shadow prices, can provide stable convergence to proportional fairness per unit charge, even in the presence of random effects and delays.

Mechanisms by which supply and demand reach equilibrium have, of course, long been a central concern of economists, and there exists a substantial body of theory on the stability of what are termed tatonnement processes [19], [20], [21]. From this viewpoint the rate control algorithms described in this paper are particular embodiments of a ‘Walrasian auctioneer’, searching for market clearing prices. The ‘Walrasian auctioneer’ of tatonnement theory is usually considered a rather implausible construct; we show that the structure of a communication network provides a natural context within which to investigate the consequences for a tatonnement process of stochastic perturbations and time lags.

The organization of the paper is as follows. In the next section we describe our basic model of a network, describe two classes of rate control algorithm, and provide an outline of our results. Detailed proofs are provided in the next two sections, following which we illustrate our theoretical results through a discussion of some numerical examples. We then consider user adaptation and routing, and finally conclude with some remarks on open issues.

Outline of results

The basic model

Consider a network with a set $J$ of resources, and let $C_j$ be the finite capacity of resource $j$, for $j \in J$. Let a route $r$ be a non-empty subset of $J$, and write $R$ for the set of possible routes. Set $A_{jr} = 1$ if $j \in r$, so that resource $j$ lies on route $r$, and set $A_{jr} = 0$ otherwise. This defines a $0-1$ matrix $A = (A_{jr}, j \in J, r \in R)$.

Associate a route $r$ with a user, and suppose that if a rate $x_r$ is allocated to
user \( r \) then this has utility \( U_r(x_r) \) to the user. Assume that the utility \( U_r(x_r) \) is an increasing, strictly concave and continuously differentiable function of \( x_r \) over the range \( x_r \geq 0 \) (following Shenker [1], we call traffic that leads to such a utility function elastic traffic). Assume further that utilities are additive, so that the aggregate utility of rates \( x = (x_r, r \in R) \) is \( \sum_{r \in R} U_r(x_r) \). Let \( U = (U_r(\cdot), r \in R) \) and \( C = (C_j, j \in J) \). Under this model the system optimal rates solve the following problem.

\[
SYSTEM(U, A, C):
\begin{align*}
\text{maximize} & \quad \sum_{r \in R} U_r(x_r) \\
\text{subject to} & \quad Ax \leq C \\
& \quad x \geq 0.
\end{align*}
\]

While this optimization problem is mathematically fairly tractable (with a strictly concave objective function and a convex feasible region), it involves utilities \( U \) that are unlikely to be known by the network. We are thus led to consider two simpler problems.

Suppose that user \( r \) may choose an amount to pay per unit time, \( w_r \), and receives in return a flow \( x_r \) proportional to \( w_r \), say \( x_r = w_r / \lambda_r \), where \( \lambda_r \) could be regarded as a charge per unit flow for user \( r \). Then the utility maximization problem for user \( r \) is as follows.

\[
USER_r(U_r; \lambda_r):
\begin{align*}
\text{maximize} & \quad U_r \left( \frac{w_r}{\lambda_r} \right) - w_r \\
& \quad w_r \geq 0.
\end{align*}
\]

Suppose next that the network knows the vector \( w = (w_r, r \in R) \), and attempts to maximize the function \( \sum_r w_r \log x_r \). The network’s optimization problem is then as follows.

\[
NETWORK(A, C; w):
\begin{align*}
\text{maximize} & \quad \sum_{r \in R} w_r \log x_r \\
\text{subject to} & \quad Ax \leq C \\
& \quad x \geq 0.
\end{align*}
\]

It is known [18] that there always exist vectors \( \lambda = (\lambda_r, r \in R) \), \( w = (w_r, r \in R) \) and \( x = (x_r, r \in R) \), satisfying \( w_r = \lambda_r x_r \) for \( r \in R \), such that \( w_r \) solves \( USER_r(U_r; \lambda_r) \) for \( r \in R \) and \( x \) solves \( NETWORK(A, C; w) \); further, the vector \( x \) is then the unique solution to \( SYSTEM(U, A, C) \).
A vector of rates \( x = (x_r, r \in R) \) is *proportionally fair* if it is feasible, that is \( x \geq 0 \) and \( Ax \leq C \), and if for any other feasible vector \( x^* \), the aggregate of proportional changes is zero or negative:

\[
\sum_{r \in R} \frac{x_r^* - x_r}{x_r} \leq 0. \quad (1)
\]

If \( w_r = 1, r \in R \), then a vector of rates \( x \) solves \( \text{NETWORK}(A, C; w) \) if and only if it is proportionally fair. Such a vector is also the Nash bargaining solution (satisfying certain axioms of fairness \([22]\)) and, as such, has been advocated in the context of telecommunications by Mazumdar et al. \([14]\).

A vector \( x \) is such that the *rates per unit charge* are proportionally fair if \( x \) is feasible, and if for any other feasible vector \( x^* \)

\[
\sum_{r \in R} w_r \frac{x_r^* - x_r}{x_r} \leq 0. \quad (2)
\]

The relationship between the conditions (1) and (2) is well illustrated when \( w_r, r \in R \), are all integral. For each \( r \in R \), replace the single user \( r \) by \( w_r \) identical sub-users, construct the proportionally fair allocation over the resulting \( \sum_r w_r \) users, and provide to user \( r \) the aggregate rate allocated to its \( w_r \) sub-users; then the resulting rates *per unit charge* are proportionally fair. This construction also illustrates the need to adapt the notion of fairness to a non-cooperative context, where it is possible for a single user to represent itself as several distinct users. It is straightforward to check \([18]\) that a vector of rates \( x \) solves \( \text{NETWORK}(A, C; w) \) if and only if the rates per unit charge are proportionally fair.

We note in passing that if, for a fixed set of users and arbitrary parameters \( w = (w_r, r \in R) \), the network solves \( \text{NETWORK}(A, C; w) \), then the resulting rates \( x = (x_r, r \in R) \) solve a variant of the problem \( \text{SYSTEM}(U, A, C) \), with a weighted objective function \( \sum_r \alpha_r U_r(x_r) \) where \( \alpha_r = w_r/(x_r U'_r(x_r)) \) for \( r \in R \). Thus a choice of the parameters \( w = (w_r, r \in R) \) by the network (rather than by users) corresponds to an implicit weighting by the network of the relative utilities of different users, with weights related to the users’ various marginal utilities.

Under the decomposition of the problem \( \text{SYSTEM}(U, A, C) \) into the problems \( \text{NETWORK}(A, C; w) \) and \( \text{USER}_r(U_r; \lambda_r), r \in R \), the utility function \( U_r(x_r) \) is not required by the network, and only appears in the optimization problem faced by user \( r \). The Lagrangian \([23]\) for the problem \( \text{NETWORK}(A, C; w) \) is

\[
L(x, z; \mu) = \sum_{r \in R} w_r \log x_r + \mu^T (C - Ax - z)
\]

where \( z \geq 0 \) is a vector of slack variables and \( \mu \) is a vector of Lagrange multipliers (or shadow prices). Then

\[
\frac{\partial L}{\partial x_r} = \frac{w_r}{x_r} - \sum_{j \in e} \mu_j,
\]
and so the unique optimum to the primal problem is given by
\[ x_r = \frac{w_r}{\sum_{j \in J} \mu_j} \]  
(3)

where \((x_r, r \in R), (\mu_j, j \in J)\) solve
\[ \mu \geq 0, \quad Ax \leq C, \quad \mu^T (C - Ax) = 0 \]  
(4)

and relation (3). Furthermore the associated dual problem quickly reduces, after elision of terms not dependent on the shadow prices \(\mu\), to the following problem.

**DUAL\((A, C; w)\):**

\[
\text{maximize } \sum_{r \in R} w_r \log(\sum_{j \in r} \mu_j) - \sum_{j \in J} \mu_j C_j
\]

over \(\mu \geq 0\).

While the problems **NETWORK\((A, C; w)\)** and **DUAL\((A, C; w)\)** are mathematically tractable, it would be difficult to implement a solution in any centralized manner. A centralized processor, even if it were itself completely reliable and could cope with the complexity of the computational task involved, would have its lines of communication through the network vulnerable to delays and failures. Rather, interest focuses on algorithms which are decentralized and of a simple form: the challenge is to understand how such algorithms can be designed so that the network as a whole reacts intelligently to perturbations. Next we describe two simple classes of decentralized algorithm, designed to implement solutions to relaxations of the problems **NETWORK\((A, C; w)\)** and **DUAL\((A, C; w)\)**.

**A primal algorithm**

Consider the system of differential equations
\[
\frac{d}{dt} x_r(t) = \kappa \left( w_r - x_r(t) \sum_{j \in r} \mu_j(t) \right)
\]  
(5)

where
\[ \mu_j(t) = p_j \left( \sum_{s \in S \setminus j} x_s(t) \right). \]  
(6)

(Here and throughout we assume that, unless otherwise specified, \(r\) ranges over the set \(R\) and \(j\) ranges over the set \(J\).) We may motivate the relations (5)–(6) in several ways. For example, suppose that \(p_j(y)\) is a price charged by resource \(j\), per unit flow through resource \(j\), when the total flow through resource \(j\) is \(y\). Then by adjusting the flow on route \(r\), \(x_r(t)\), in accordance with equations (5)–(6), the network attempts to equalize the aggregate cost of this flow, \(x_r(t) \sum_{j \in r} \mu_j(t)\),
with a target value $w_r$, for every $r \in R$. (For an enlightening description of the technological implementation of such algorithms in an ATM network, see Courcoubetis et al. [24].)

For an alternative motivation, suppose that resource $j$ generates a continuous stream of feedback signals at rate $p_j(y)$ when the total flow through resource $j$ is $y$. Suppose further that when resource $j$ generates a feedback signal, a copy is sent to each user $r$ whose route passes through resource $j$, where it is interpreted as a congestion indicator requiring some reduction in the flow $x_r$. Then equation (5) corresponds to a response by user $r$ that comprises two components: a steady increase at rate proportional to $w_r$, and a multiplicative decrease at rate proportional to the stream of feedback signals received. (For early discussions of algorithms with additive increase and multiplicative decrease see Chiu and Jain [7] and Jacobson [2]; Hernandez-Valencia et al. [6] review several algorithms based on congestion indication feedback.)

Later we establish that under mild regularity conditions on the functions $p_j, j \in J$, the expression

$$U(x) = \sum_{r \in R} w_r \log x_r - \sum_{j \in J} \int_0^{\sum_{s: j \in s} x_s} p_j(y) dy$$

(7)

provides a Lyapunov function for the system of differential equations (5)–(6), and we deduce that the vector $x$ maximizing $U(x)$ is a stable point of the system, to which all trajectories converge.

The functions $p_j, j \in J$, may be chosen so that the maximization of the Lyapunov function $U(x)$ arbitrarily closely approximates the optimization problem $\text{NETWORK}(A, C; w)$, and, in this sense, is a relaxation of the network problem. In our penultimate section we shall see that certain relaxations correspond naturally to a system objective which takes into account loss or delays, as well as flow rates.

The Lyapunov function (7) thus provides an enlightening analysis of the global stability of the system (5)–(6), and of the relationship between this system and the problem $\text{NETWORK}(A, C; w)$. However, the system (5)–(6) has omitted to model two important aspects of decentralized systems, namely stochastic perturbations, and time lags. We analyze these aspects by considering small perturbations to the stable point $x$.

Stochastic perturbations within the network may well arise from a resource’s method of sensing its load. Equation (6) represents the response $y(t)$ of resource $j$ as a continuous function of a load, $y = \sum_{s: j \in s} x_s$, which is assumed known. In practice a resource may assess its load by an error-prone measurement mechanism, and then choose a feedback signal from a small set of possible signals. (See Hernandez-Valencia et al. [6] and Bonomi et al. [10] for more detailed descriptions of binary feedback and congestion indication rate control algorithms.) In the next section we describe how such mechanisms motivate various stochastic models of
the network. One particular model takes the form

\[ dx_r(t) = \kappa \left( w_r dt - x_r(t) \sum_{j \in r} \left( \mu_j(t) dt + \mu_j(t)^{1/2} \varepsilon_j^{1/2} dB_j(t) \right) \right) \]  

(8)

where \( B_j(t) \) is a standard Brownian motion, representing stochastic effects at resource \( j \), and \( \varepsilon_j \) is a scaling parameter for these effects. If the scaling parameters \( \varepsilon_j, j \in J \), are small then the stochastic differential equation (8) has, as solution, a multidimensional Ornstein–Uhlenbeck process, centred on the stable point \( x \) of the differential equations (5)–(6). The stationary distribution for \( (x_r(t), r \in R) \) is a multivariate normal distribution, with covariance matrix that can be explicitly calculated in terms of the parameters of the network.

Similarly we shall describe a model incorporating time lags that generalizes equations (5)–(6), and shall analyse its behaviour close to the stable point \( x \). Our models of stochastic effects and of time-lags provide important insights into the behaviour of the network, and allow us to quantify the various relationships and trade-offs between speed of convergence, the magnitude of fluctuations about the equilibrium point, and the stability of the network.

**A dual algorithm**

The equations (5)–(6) represent a system where rates vary gradually, and shadow prices are given as functions of the rates. Next we consider a system where shadow prices vary gradually, with rates given as functions of the shadow prices. Let

\[ \frac{d}{dt} \mu_j(t) = \kappa \left( x_r(t) - q_j(\mu_j(t)) \right) \]  

(9)

where

\[ x_r(t) = \frac{w_r}{\sum_{k \in r} \mu_k(t)}. \]  

(10)

The relationship between the algorithm (9)–(10) and the problem \( DU AL(A, C; w) \) parallels that between the primal algorithm (5)–(6) and the problem \( NETWORK(A, C; w) \), and, again, we may motivate the algorithm in several ways. For example, suppose that \( q_j(\eta) \) is the flow through resource \( j \) which generates a price at resource \( j \) of \( \eta \). Then an economist would describe the right hand side of equation (9) as the vector of excess demand at prices \( (\mu_j(t), j \in J) \), and would recognise equations (9)–(10) as a tatonnement process by which prices adjust according to supply and demand (Varian [21], Chapter 21).

Later we establish that under mild regularity conditions on the functions \( q_j \), \( j \in J \), the expression

\[ V(\mu) = \sum_{r \in R} w_r \log \left( \sum_{j \in r} \mu_j \right) - \sum_{j \in J} \int_0^{\mu_j} q_j(\eta) d\eta \]  

(11)
provides a Lyapunov function for the system of differential equations (9)–(10), and we deduce that the vector \( \mu \) maximizing \( V(\mu) \) is a stable point of the system, to which all trajectories converge. Further, by appropriate choice of the functions \( q_j, j \in J \), the maximization of the function \( V(\mu) \) can arbitrarily approximate the problem \( DU. AL(A, C; w) \).

We consider stochastic perturbations of system (9)–(10), with a typical example taking the form

\[
d\mu_j(t) = \kappa \left( \sum_{r \in R} \left( x_r(t) dt + x_r(t)^{1/2} \varepsilon_r^{1/2} dB_r(t) \right) - q_j(\mu_j(t)) dt \right)
\]

where \( B_r(t) \) is a standard Brownian motion, representing stochastic effects associated with the flow on route \( r \). If the scaling parameters \( \varepsilon_r, r \in R \), are small then the stationary distribution for \( (\mu_j(t), j \in J) \) is centred on the stable point \( \mu \) of the differential equations (9)–(10), with a covariance matrix that can be explicitly calculated in terms of the parameters of the network. Also it is possible to analyse the stability of the model (9)–(10) when time-lags are introduced.

**User adaptation**

Our analyses of the primal algorithm (5)–(6) and the dual algorithm (9)–(10) assume that the parameters \( (w_r, r \in R) \) chosen by the users are fixed, at least on the time scales concerned in the analyses. With increasing intelligence embedded in end-systems, users may in the future be able to vary the parameters \( (w_r, r \in R) \) even within these short time scales. Both the algorithms may be extended to this situation.

Suppose that user \( r \) is able to monitor its rate \( x_r(t) \) continuously, and to vary smoothly the parameter \( w_r(t) \) so as to track accurately the optimum to \( USER_r(U_r; \lambda_r(t)) \), where \( \lambda_r(t) = w_r(t)/x_r(t) \) is the charge per unit flow to user \( r \) at time \( t \). Then, using revised Lyapunov functions, stability of both the primal and dual algorithms may again be established.

Our next sections provide detailed proofs of the various results outlined above, together with some numerical illustrations. In our penultimate section we shall look again at the system decomposition relating the problems \( SYSTEM(U, A, C) \) and \( NETWORK(A, C; w) \), and extend the discussion to include routing control.

**A primal algorithm**

In this section we establish the global stability of the primal algorithm (5)–(6), determine the rate of convergence, and, by considering perturbations about the stable point, model stochastic effects and time lags.
Global stability

Let the function $U(x)$ be defined by equation (7) where $w_r > 0$, $r \in R$, and suppose that, for $j \in J$, the function $p_j(y)$, $y \geq 0$, is a non-negative, continuous, increasing function of $y$, not identically zero.

**Theorem 1** The strictly concave function $U(x)$ is a Lyapunov function for the system of differential equations (5)–(6). The unique value $x$ maximizing $U(x)$ is a stable point of the system, to which all trajectories converge.

**Proof.** The assumptions on $w_r > 0$, $r \in R$, and $p_j, j \in J$, ensure that $U(x)$ is strictly concave on $x \geq 0$ with an interior maximum; the maximizing value of $x$ is thus unique. Observe that

$$
\frac{\partial}{\partial x_r} U(x) = \frac{w_r}{x_r} - \sum_{j \in I_r} p_j \left( \sum_{s \in J} x_s \right);
$$

(13)

setting these derivatives to zero identifies the maximum. Further

$$
\frac{d}{dt} U(x(t)) = \sum_{r \in R} \frac{\partial U}{\partial x_r} \frac{d}{dt} x_r(t)
$$

$$
= \kappa \sum_{r \in R} \frac{1}{x_r(t)} \left( w_r - x_r(t) \sum_{j \in I_r} p_j \left( \sum_{s \in J} x_s(t) \right) \right)^2,
$$

establishing that $U(x(t))$ is strictly increasing with $t$, unless $x(t) = x$, the unique $x$ maximizing $U(x)$. The function $U(x)$ is thus a Lyapunov function for the system (5)–(6), and the theorem follows ([25], Chapter 5). □

Define the continuous functions $p_j(y) = (y - C_j + \varepsilon)/\varepsilon^2$ for $j \in J$. Then, as $\varepsilon \to 0$, the maximization of the Lyapunov function $U(x)$ approximates arbitrarily closely the primal problem $NETWORK(A, C; w)$; in particular the vector $x$ maximizing $U(x)$ approaches the solution $x$ to relations (3) and (4). Note, however, that the derivative $p_j'(y)$ may become arbitrarily large as the approximation improves.

**Rate of convergence**

We have seen, in Theorem 1, that the system (5)–(6) converges to a unique stable point: next we investigate the rate of convergence, by linearization about the stable point.

Let $x$ identify the unique vector maximizing $U(x)$, let $\mu_j = p_j(\sum_{s \in s : j \in s} x_s)$, and suppose $p_j$ is differentiable at this point, with derivative $p_j'$. Let $x_r(t) =$
\[ x_r + x_r^{1/2} y_r(t). \] Then, linearizing the system (5)–(6) about \( x \), we obtain
\[
\frac{d}{dt} y_r(t) = -\kappa(y_r(t)) \sum_{j \in r} \mu_j + x_r^{1/2} \sum_{j \in r} \mu_j \sum_{s, j \in s} x_s^{1/2} y_s(t)) \\
= -\kappa \left( \frac{w_r}{x_r} y_r(t) + x_r^{1/2} \sum_{j} \sum_{s} \mu_j A_{jr} A_{js} x_s^{1/2} y_s(t) \right).
\]

We may write this in matrix form as
\[
\frac{d}{dt} y(t) = -\kappa(WX^{-1} + X^{1/2} A^T P' A X^{1/2}) y(t)
\]
where \( X = \text{diag}(x_r, r \in R) \), \( W = \text{diag}(w_r, r \in R) \) and \( P' = \text{diag}(\mu_j, j \in J) \).

Let
\[
\Gamma^T \Phi \Gamma = WX^{-1} + X^{1/2} A^T P' A X^{1/2}
\]
where \( \Gamma \) is an orthogonal matrix, \( \Gamma^T \Gamma = I \), and \( \Phi = \text{diag}(\phi_r, r \in R) \) is the matrix of eigenvalues, necessarily positive, of the real, symmetric, positive definite matrix (15). Then
\[
\frac{d}{dt} y(t) = -\kappa \Gamma^T \Phi \Gamma y(t),
\]
and thus the rate of convergence to the stable point is determined by the smallest eigenvalue, \( \phi_r \), \( r \in R \), of the matrix (15). Note that speed of convergence increases both with the gain parameter \( \kappa \) and with the magnitude of the derivatives \( P' \); we shall see that this conclusion requires qualification in the presence of either stochastic effects or of time-lags.

Our early assumption that \( p_j(y), j \in J \), are increasing functions is convenient and often natural: it implies that \( U \) is strictly concave with an interior maximum. If the functions \( p_j(y), j \in J \), are not increasing, then \( U(x) \) may not have an interior maximum or it may have multiple stationary points: we describe an example later. Provided \( p_j(y), j \in J \), are differentiable at a stationary point, the local behaviour near the stationary point is described by the equation (16).

**Stochastic analysis**

Next we consider a stochastic perturbation of the linearized equation (16). Let
\[
dy(t) = -\kappa \left( \Gamma^T \Phi \Gamma y(t) dt + F dB(t) \right)
\]
where \( F \) is an arbitrary \([R] \times [I]\) matrix and \( B(t) = (B_i(t), i \in I) \) is a collection of independent standard Brownian motions, extended to \(-\infty < t < \infty\). (Later we describe how the modelling of different sources of randomness may lead to various explicit forms for the matrix \( F \)).
The stationary solution to the system (17) is

$$y(t) = -\kappa \int_{-\infty}^{t} e^{-\kappa(t-\tau)} \Gamma \Phi \Gamma F dB(\tau),$$

(18)

as can be checked by differentiating both sides of equation (18) with respect to $t$. The solution (18) is a linear transformation of the Gaussian process $(B(\tau), \tau < t)$; hence $y(t)$ has a multivariate normal distribution, $y(t) \sim N(0, \Sigma)$, where

$$\Sigma = E[y(t) y(t)^T] = \kappa^2 \int_{-\infty}^{0} e^{\kappa \tau \Gamma \Phi \Gamma F F^T \Gamma^T} \kappa \Gamma \Phi \Gamma F F^T \Gamma^T e^{\kappa \tau \Gamma \Phi \Gamma F F^T \Gamma^T} d\tau,$$

$$= \kappa \Gamma^T \left[ \int_{-\infty}^{0} e^{\tau \Gamma \Phi \Gamma F F^T \Gamma^T} e^{\tau \Phi} d\tau \right] \Gamma.$$

Define the symmetric matrix $[\Gamma F; \Phi]$ by

$$[\Gamma F; \Phi]_{rs} = \left[ \int_{-\infty}^{0} e^{\tau \Phi} \Gamma F F^T \Gamma^T e^{\tau \Phi} d\tau \right]_{rs},$$

$$= \frac{[\Gamma F F^T \Gamma^T]_{rs}}{\phi_r + \phi_s}.$$

Then

$$\Sigma = \kappa \Gamma^T [\Gamma F; \Phi] \Gamma.$$

(19)

Note that the covariance matrix increases linearly with the gain parameter $\kappa$; as $\kappa$ increases, the faster convergence to equilibrium described by relation (16) is at the cost of a greater spread at equilibrium. Varying the derivatives $P^r$ has a more subtle effect, through relation (15) and the construction (19), on the covariance matrix; broadly, as $P^r$ increases, not only is convergence to equilibrium faster, but also the spread at equilibrium decreases. However, we shall see later that, in the presence of time-lags, increasing $P^r$ may compromise stability.

We next illustrate how various sources of randomness may lead to different covariance structures.

**Congestion indication with joint feedback.** Consider the following stochastic version of the system (5)–(6). Let $(N_j(\tau), \tau \geq 0)$, for $j \in J$, be a collection of independent unit rate Poisson processes, and let

$$dx_r(t) = \kappa \left( w_r dt - x_r(t) \sum_{j \in \tau} \varepsilon_j dN_j \left( \varepsilon_j^{-1} \int_{0}^{t} \mu_j(\tau) d\tau \right) \right),$$

(20)

where the functions $\mu_j(\cdot)$, for $j \in J$, are given by equation (6). Equation (20) would describe the following model: resource $j$ generates feedback signals indicating congestion as a time-dependent Poisson process at rate $\varepsilon_j^{-1} \mu_j(t)$; when
resource \( j \) generates a feedback signal, a copy is sent to each user \( r \) whose route passes through resource \( j \); and user \( r \) reacts to such a feedback signal by reducing \( x_r(t) \) by an amount \( \kappa \varepsilon_j x_r(t) \).

Now as \( \varepsilon \to 0 \), the normalized Poisson process \( ((\varepsilon N_j(\tau/\varepsilon) - \tau) \varepsilon^{-1/2}, \tau \geq 0) \) converges in distribution to a standard Brownian motion. This motivates the approximation, valid when \( \varepsilon_j \) are small, of the Poisson driving equation (20) by its Brownian version

\[
dx_r(t) = \kappa \left( w_r dt - x_r(t) \sum_{j \in \mathcal{R}} \left( \mu_j(t) dt + \varepsilon_j^{1/2} \mu_j(t)^{1/2} dB_j(t) \right) \right)
\]

where \( (B_j(t), t \geq 0), \) for \( j \in J \), are a collection of independent standard Brownian motions.

The corresponding Brownian version of the linearized system (16) is just equation (17) where \( B(t) = (B_j(t), j \in J), \) and \( F \) is an \( |R| \times |J| \) matrix with elements

\[
F_{r,j} = \varepsilon_j^{1/2} \mu_j^{1/2} A_{jr} x_r^{1/2}.
\]  

Then

\[
FF^T = X^{1/2} A^T EPAX^{1/2},
\]  

where \( E = \text{diag}(\varepsilon_j, j \in J) \) and \( P = \text{diag}(\mu_j, j \in J) \), and hence the stationary covariance matrix \( \Sigma \) may be calculated from expression (19).

**Congestion indication with individual feedback.** Consider next the Poisson driving equation

\[
dx_r(t) = \kappa \left( w_r dt - \sum_{j \in J} \varepsilon_j dN_{jr} \left( \varepsilon_j^{-1} \int_0^t x_r(\tau) \mu_j(\tau) d\tau \right) \right)
\]

where \( (N_{jr}(\tau), \tau \geq 0), \) for \( j \in J, r \in R, \) are a collection of independent unit rate Poisson processes. This would describe the following model: feedback signals from resource \( j \) to user \( r \) arise at rate \( \varepsilon_j^{-1} x_r(t) \mu_j(t) \); and user \( r \) reacts to such a feedback signal by reducing \( x_r(t) \) by an amount \( \kappa \varepsilon_j \). The Brownian approximation, valid when \( \varepsilon_j \) are small, becomes

\[
dx_r(t) = \kappa \left( w_r dt - x_r(t) \sum_{j \in \mathcal{R}} \left( \mu_j(t) dt + \varepsilon_j^{1/2} x_r(t)^{-1/2} \mu_j(t)^{1/2} dB_{jr}(t) \right) \right),
\]

whose linearization is equation (17) where the \( |R| \times |J| \times |R| \) matrix \( F \) is given by

\[
F_{r,(j,s)} = \varepsilon_j^{1/2} \mu_j^{1/2} A_{jr} I[r = s].
\]  

Thus \( FF^T \) is the matrix \( \text{diag} \left( \sum_{j \in \mathcal{R}} \mu_j \varepsilon_j, r \in R \right) \), and the stationary covariance matrix \( \Sigma \) may be calculated from expression (19). Later we provide a numerical illustration of this calculation, and contrast the results derived from the forms (21) and (24).
Source fluctuations. Consider the Brownian driving equation

\[ dx_r(t) = \kappa \left( w_r dt - x_r(t) \sum_{j \in r} \mu_j(t) dt + \varepsilon_r^{1/2} x_r(t)^{1/2} dB_r(t) \right), \]

which might correspond to fluctuations arising at sources, rather than within the network. For this system the \(|R| \times |R|\) matrix \(F\) is the diagonal matrix \(\text{diag}(\varepsilon_r^{1/2}, r \in R)\).

Time lags
Consider next the lagged, discrete time system

\[ x_r[t + 1] = x_r[t] + \kappa \left( w_r - x_r[t] \sum_{j \in r} \mu_j[t - d(j, r)] \right) \quad (25) \]

where

\[ \mu_j[t] = p_j \left( \sum_{s \in s} x_s[t - d(j, s)] \right), \quad (26) \]

and \(d(j, r), j \in J, r \in R,\) are non-negative integers. This might correspond to a model of congestion indication with joint feedback, where there is a delay of \(d(j, r)\) between resource \(j\) generating a feedback signal and user \(r\) receiving it, and another delay of \(d(j, r)\) between user \(r\) changing its rate and the altered flow reaching resource \(j\). Say that a vector \(x\) is an equilibrium point of the system (25)–(26) if \(x_r[t] = x_r, \) for \(t = \ldots, 0, 1, 2, \ldots,\) satisfies these equations.

**Theorem 2** The vector \(x\) maximizing the strictly concave function \(U(x)\) is the unique equilibrium point of the system (25)–(26).

**Proof.** The vector \(x\) is an equilibrium point if and only if it solves

\[ w_r = x_r \sum_{j \in r} p_j \left( \sum_{s \in r} x_s \right). \]

But this is precisely the stationarity condition implied by the partial derivatives (13) of the function \(U(x)\), a strictly concave function with a unique maximum. \(\square\)

For small enough values of \(\kappa\) the equilibrium point will be asymptotically stable, since if we replace \(\kappa\) by \(\kappa \delta\) in equation (25) and let \(x_r(t) = x_r[t/\delta]\), then as \(\delta \to 0\) we may approximate arbitrarily closely a solution to equations (5)–(6). But for small values of \(\kappa\) convergence to the equilibrium point is slow, and so it is of interest to investigate the local stability of the equilibrium point for general values of \(\kappa\).
Let $\mu_j = p_j |\sum_{s \in s} x_s|$, and suppose $p_j$ is differentiable at this point, with derivative $p'_j$. Let $x_r[t] = x_r + x_r^{1/2} y_r[t]$. Then, linearizing the system (25)–(26) about $x$, we obtain

$$y_r[t + 1] = y_r[t] - \kappa \left( y_r[t] \sum_{j \in r} \mu_j + x_r^{1/2} \sum_{j \in r} p'_j \sum_{s \in s} x_s^{1/2} y_s[t - d(j, r) - d(j, s)] \right)$$

$$= y_r[t] - \kappa \left( \frac{p'_r}{x_r} y_r[t] + \sum_j \sum_s p'_j A_{jr} A_{js} x_r^{1/2} x_s^{1/2} y_s[t - d(j, r) - d(j, s)] \right).$$

Define the $|R| \times |R|$ matrices $(L[d], d = 0, 1, \ldots, D)$ by

$$(L[d])_{rs} = \sum_j p'_j A_{jr} A_{js} x_r^{1/2} x_s^{1/2} I[d(j, r) + d(j, s) = d]$$

where $D = \max_{j,r,s}\{d(j, r) + d(j, s)\}$. Thus

$$\sum_{d=0}^{D} L[d] = X^{1/2} A^T P' A X^{1/2},$$

the second term of the key matrix (15). Define the vector $y[t] = (y_r[t], r \in R)$. Then we can rewrite equation (28) in the matrix form

$$\begin{pmatrix}
y[t + 1] \\
y[t] \\
\vdots \\
y[t - D + 1]
\end{pmatrix} = L \begin{pmatrix}
y[t] \\
y[t - 1] \\
\vdots \\
y[t - D]
\end{pmatrix}$$

where

$$L = \begin{pmatrix}
0 & I & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & 0
\end{pmatrix}.$$  \hspace{1cm} (29)

The equilibrium point $x$ of the system (25)–(26) is stable if and only if the spectral radius of the matrix $L$ is less than unity. Recall that in our model of stochastic effects, increasing the derivatives $P'$ had broadly the same reductive effect on the covariance matrix (19) as decreasing the gain parameter $\kappa$; in contrast the destabilising effect on the matrix (29) of increasing $P'$ is broadly the same as increasing $\kappa$.

For simplicity of notation we have used the same gain parameter $\kappa$ for each $r \in R$. If $\kappa$ is replaced by $\kappa_r$ in equation (25), then we again obtain relations (28)–(29), but now with $\kappa$ interpreted as the matrix $\text{diag}(\kappa_r, r \in R)$. An interesting
topic concerns how the time delays within a network affect the choice of gain parameters; we might for example study the problem of choosing \( \text{diag}(\kappa_r, r \in R) \) in order to minimize the spectral radius of the matrix \( L \).

There exist other natural discrete time versions of the equations (5)-(6), and these too may be analysed in a similar manner. For example, consider the method of repeated substitution, \( x_r[t+1] = \frac{w_r}{\sum_{j \in r} \mu_j[t-d(j,r)]} \) or its damped version

\[
x_r[t+1] = (1 - \kappa)x_r[t] + \kappa \frac{w_r}{\sum_{j \in r} \mu_j[t-d(j,r)]}
\]

where \( \mu_j[t] \) is again given by equation (26). Then the linearized relations (28)-(29) are altered in that the top row of the matrix (29) becomes

\[
(I - \kappa(I + XW^{-1}L[0]), -\kappa XW^{-1}L[1], \ldots, -\kappa XW^{-1}L[D]).
\]

A dual algorithm

In this section we investigate the stability of the dual algorithm (9)-(10), including a perturbation analysis of stochastic effects and time lags. Finally we note that the system (9)-(10) is just one example of a dual algorithm, and consider variants that share the Lyapunov function (11).

Global stability

Let the function \( \mathcal{V}(\mu) \) be defined by equation (11), where \( w_r > 0, r \in R \), and suppose that, for \( j \in J, q_j(0) = 0 \) and \( q_j(\eta), \eta \geq 0 \), is a continuous, strictly increasing function of \( \eta \).

**Theorem 3** The strictly concave function \( \mathcal{V}(\mu) \) is a Lyapunov function for the system of differential equations (9)-(10). The unique value \( \mu \) maximizing \( \mathcal{V}(\mu) \) is a stable point of the system, to which all trajectories converge.

**Proof.** The assumptions on \( w_r > 0, r \in R \), and on \( q_j, j \in J \), ensure that \( \mathcal{V}(\mu) \) is strictly concave on \( \mu \geq 0 \) with an interior maximum; the maximizing value of \( \mu \) is thus unique, and is determined by setting the derivatives

\[
\frac{\partial}{\partial \mu_j} \mathcal{V}(\mu) = \sum_{r,j \in r} \frac{w_r}{\sum_{k \in r} \mu_k} - q_j(\mu_j)
\]

(30)
to zero. Also,

\[
\frac{d}{dt} \mathcal{V}(\mu(t)) = \sum_{j \in J} \frac{\partial \mathcal{V}}{\partial \mu_j} \cdot \frac{d}{dt} \mu_j(t)
\]

\[
= \kappa \sum_{j \in J} \left( \sum_{r,j \in r} \frac{w_r}{\sum_{k \in r} \mu_k(t)} - q_j(\mu_j(t)) \right)^2,
\]

16
establishing that $V(\mu(t))$ is strictly increasing with $t$, unless $\mu(t) = \mu$, the unique value $\mu$ maximizing $V(\mu)$. The function $V(\mu)$ is thus a Lyapunov function for the system (9)–(10), and the theorem follows [25]. □

The maximization of the Lyapunov function $V(\mu)$ becomes the dual problem if, for $j \in J$, $\eta > 0$, $q_j(\eta) = C_j$. These functions violate the assumption that $q_j(\eta)$ is continuous at $\eta = 0$, but they may be arbitrarily closely approximated, for example by the functions $q_j(\eta) = C_j \eta/(\eta + \varepsilon)$ for small positive $\varepsilon$. Note, however, that the derivative $q_j'(\eta)$ may become arbitrarily large as the approximation improves.

Rate of convergence

Let $\mu$ identify the unique vector maximizing $V(\mu)$, let $x_r = w_r/\sum_{k \in R} \mu_k$, and suppose $q_j(y)$ differentiable at the point $y = \mu_j$, with derivative $q_j'$. Let $\mu_j(t) = \mu_j + \xi_j(t)$. Then, linearizing the system (9)–(10) about $\mu$, we obtain, after some reduction,

$$\frac{d}{dt} \xi(t) = -\kappa(AXW^{-1}XAT + Q')\xi(t)$$

where $W = \text{diag}(w_r, r \in R)$ and $Q' = \text{diag}(q_j', j \in J)$. Let

$$\Theta^T \Psi \Theta = AXW^{-1}XAT + Q'$$

(31)

where $\Theta$ is an orthogonal matrix, $\Theta^T \Theta = I$, and $\Psi = \text{diag}(\psi_j, j \in J)$ is the matrix of eigenvalues, necessarily non-negative, of the real, symmetric, positive semi-definite matrix (31). Then

$$\frac{d}{dt} \xi(t) = -\kappa \Theta^T \Psi \Theta \xi(t),$$

(32)

and thus the rate of convergence to the stable point is determined by the smallest eigenvalue of the matrix (31). Note that speed of convergence increases both with the gain parameter $\kappa$ and with the magnitude of the derivatives $Q'$.

Stochastic analysis

Next consider a stochastic perturbation of the linearized equation (32). Let

$$d\xi(t) = -\kappa \left( \Theta^T \Psi \Theta \xi(t) dt - GdB(t) \right)$$

(33)

where $B(t) = (B_i(t), i \in I)$ is a collection of independent standard Brownian motions, and $G$ is a $|J| \times |I|$ matrix.

A similar analysis to that of the last section determines the stationary covariance matrix $\Sigma$ of $\xi(t)$. Define the symmetric matrix $[\Theta G; \Psi]$ by

$$[\Theta G; \Psi]_{jk} = \frac{[\Theta GG^T \Theta^T]_{jk}}{\psi_j + \psi_k}.$$
Then
\[ \Sigma = \kappa \Theta^T [\Theta G; \Psi] \Theta. \]  
(34)

Note that the covariance matrix increases linearly with the gain parameter \( \kappa \); as \( \kappa \) increases, the faster convergence to equilibrium described by relation (32) is at the cost of a greater spread at equilibrium.

Next we describe an example illustrating how a model of the form (33) might arise.

**Shadow prices inferred from fluctuating flow rates.** Consider the Poisson driving equation

\[ d\mu_j(t) = \kappa \left( \sum_{r,j \in R} \varepsilon_r dN_r \left( \varepsilon_r^{-1} \int_0^t x_r(\tau) d\tau \right) - q_j(\mu_j(t)) dt \right) \]

where \((N_r(\tau), \tau \geq 0)\), for \(r \in R\), are a collection of independent unit rate Poisson processes. This would describe a model where, on a very fine time-scale, the flow on route \(r\) takes the form of a time-dependent Poisson process of rate \(x_r(t)/\varepsilon_r\), with each point of the process containing a workload of size \(\varepsilon_r\). The Brownian approximation, valid when \(\varepsilon_r\) are small, becomes

\[ d\mu_j(t) = \kappa \left( \sum_{r,j \in R} \left( x_r(t) dt + \varepsilon_r^{1/2} x_r(t)^{1/2} dB_r(t) \right) - q_j(\mu_j(t)) dt \right) \]

whose linearization is equation (33) where \(G\) is a \(|J| \times |R|\) matrix with elements

\[ G_{jr} = \varepsilon_r^{1/2} x_r^{1/2} A_{jr}; \]

thus \(GG^T = AXEAT\) where \(E = \text{diag}(\varepsilon_r, r \in R)\).

**Time lags**

Consider next the system

\[ \mu_j[t+1] = \mu_j[t] + \kappa \left( \sum_{r,j \in R} x_r[t - d(j, r)] - q_j(\mu_j[t]) \right) \]

(35)

where

\[ x_r[t] = \frac{w_r}{\sum_{k \in R} \mu_k[t - d(k, r)]}. \]

(36)

A vector \(\mu\) is an equilibrium point of the system (35)-(36) if \(\mu_j[t] = \mu_j\), for \(t = \ldots, 0, 1, 2, \ldots\), satisfies these equations.

**Theorem 4** The vector \(\mu\) maximizing the strictly concave function \(V(\mu)\) is the unique equilibrium point of the system (35)-(36).
Proof. The vector $\mu$ is an equilibrium point if and only if solves
\[
\sum_{r,j \in R} \frac{w_r}{\sum_{k \in R} \mu_k} = q_j(\mu_j).
\]
But this is precisely the stationarity condition implied by the partial derivatives (30) of the function $V(\mu)$, a strictly concave function with a unique maximum. The result follows. □

Next we investigate the stability of the equilibrium point. Let $x_r = w_r/\sum_{k \in R} \mu_k$, and suppose $q_j$ is differentiable at the point $y = \mu_j$, with derivative $q_j'$. Let $\mu_j[t] = \mu_j + \xi_j[t]$. Then, linearizing the system (35)-(36) about $\mu$, we obtain
\[
\xi_j[t + 1] = \xi_j[t] - \kappa \left( \sum_{r,j \in R} x_r^2 w_r^{-1} \sum_{k \in R} \xi_k[t - d(j, r) - d(k, r)] + q_j' \xi_j[t] \right).
\]
Define the $|J| \times |J|$ matrices $(M[d], d = 0, 1, \ldots, D)$ by
\[
(M[d])_{jk} = \sum_{r} x_r A_{jr} A_{kr} I[d(j, r) + d(k, r) = d]
\]
where now $D = \max_{j, k, r} \{d(j, r) + d(k, r)\}$. Thus
\[
\sum_{d=0}^{D} M[d] = AXW^{-1}XA^T.
\]
Define the vector $\xi[t] = (\xi_j[t], j \in J)$. Then we can rewrite equation (37) in the matrix form
\[
\begin{pmatrix}
\xi[t + 1] \\
\xi[t] \\
\vdots \\
\xi[t - D + 1]
\end{pmatrix} = M \begin{pmatrix}
\xi[t] \\
\xi[t - 1] \\
\vdots \\
\xi[t - D]
\end{pmatrix}
\]
where
\[
M = \begin{pmatrix}
I - \kappa(M[0] + Q') & -\kappa M[1] & -\kappa M[2] & \ldots & -\kappa M[D] \\
I & 0 & 0 & \ldots & 0 \\
0 & I & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & 0
\end{pmatrix}.
\]
The equilibrium point $\mu$ of the system (35)-(36) is stable if the spectral radius of the matrix $M$ is less than unity. With stochastic effects, increasing the derivatives $Q'$ has broadly the same reductive effect on the covariance matrix (34) as decreasing the gain parameter $\kappa$; in contrast the destabilising effect on the matrix (38) of increasing $Q'$ is broadly the same as increasing $\kappa$. 

19
**Variants**

Several variants of the primal algorithm (5)–(6) and the dual algorithm (9)–(10) allow a similar analysis. For example, if the right hand side of equation (5) is multiplied by a positive function \( f_r(x(t), \mu(t)) \) then Theorem 1 remains valid. Similarly, if the right hand side of equation (9) is multiplied by a positive function \( f_j(x(t), \mu(t)) \) then Theorem 3 remains valid. As a simple example, we could divide the right hand side of equation (5) by \( w_r \), or of equation (9) by \( q_j(\mu_j(t)) \).

A more subtle variation is obtained if equation (9) is replaced by

\[
\frac{d}{dt} \mu_j(t) = \kappa \left( p_j \left( \sum_{r \in J} x_r(t) \right) - \mu_j(t) \right),
\]

where \( p_j \) is the inverse function of \( q_j \), and \( x_r(t) \) is again given by equation (10). Note that the expression (39) is of the same sign as expression (9), and so the proof of Theorem 3 goes through as before. Suppose \( p_j(y) \) is differentiable at the stable point with derivative \( p'_j \), and let \( \mu_j(t) = \mu_j + \xi_j(t) \). Then, linearizing the system (39) about the equilibrium point, we obtain

\[
\frac{d}{dt} \xi(t) = -\kappa (I + PAXW^{-1}XA^T) \xi(t)
\]

where \( P = \text{diag}(p'_j, j \in J) \), allowing the local convergence properties of the algorithm (39) to be studied.

**Examples**

In this section we illustrate the results of the last two sections through a discussion of some examples. The first sub-section illustrates how the functions \( p_j, j \in J \), may be determined by the detailed stochastic behaviour of resource \( j \); a simple four node network is used to facilitate comparisons between feedback mechanisms. The results of this paper are, of course, intended to apply to large-scale networks, and our second sub-section discusses the behaviour of a dual algorithm in a random network with 100 resources and 1000 routes.

**Congestion indication in a four node network**

Suppose that the total load \( y \) on a resource takes the form, on a very fine time-scale, of a Poisson stream of cells at rate \( y/\varepsilon \). Suppose that the time-axis is divided into non-overlapping slots each of length \( \tau \varepsilon \), and that a feedback signal is generated for a slot if the total number of cells arriving in that slot exceeds a threshold \( N \). (While there may well be a queue at a resource, we suppose for the moment that the feedback signals are generated by the process just described, rather than, for example, by the queue size exceeding a threshold.) Suppose that

20