A positive systems model of TCP-like congestion control: Asymptotic results

Robert Shorten Fabian Wirth Douglas Leith

April 7, 2004

Abstract

In this paper we study communication networks that employ drop-tail queueing and *Additive-Increase Multiplicative-Decrease (AIMD)* congestion control algorithms. We show that the theory of nonnegative matrices may be employed to model such networks. In particular, we show that important network properties such as: (i) fairness; (ii) rate of convergence; and (iii) throughput; can be characterised by certain non-negative matrices that arise in the study of *AIMD* networks. We demonstrate that these results can be used to develop tools for analysing the behaviour of *AIMD* communication networks. The accuracy of the models is demonstrated by means of several *NS*-studies.

1 Introduction

In this paper we describe a design oriented modelling approach that captures the essential features of networks of *AIMD* sources that employ drop-tail queues. The novelty of our approach lies in the fact that we are able to use the theory of non-negative matrices and hybrid systems to build mathematical models of communication networks. This approach is based upon a number of simple observations: (i) communication networks employing congestion control systems are feedback systems; (ii) communication systems exhibit event driven phenomena and may therefore be viewed as classical hybrid systems; and (iii) network states (queue length, window size, etc.) take only non-negative values. We show that it is possible to relate important network properties to the characteristics of the non-negative matrices that arise in the study of such communication networks. In particular, we will demonstrate that (i) bandwidth allocation amongst flows, (ii) rate of network convergence, and (iii) network throughput can all be related to properties of sets of non-negative matrices. We defer the discussion of related results in the literature to comments at the relevant places within the paper.

This paper is structured as follows. In Section 2 we develop a positive systems network model that captures the essential features of communication networks employing drop-tail queuing and*AIMD* congestion control algorithms. An exact model is presented for the case where all network sources share a uniform round-trip-time (RTT) and packet drops are synchronised. We then show how this model may be extended to the case of sources with differing RTT's and where packet drops need not be synchronised. This approach gives rise to a model of *AIMD* networks in which the network dynamics are described by a finite set of non-negative matrices. The main results of this paper are presented in Section 3. To ease exposition these results, which concern the short and long term behaviour of AIMD networks, are simply stated in this section. The use of these results to analyse network behaviour is illustrated by means of a number of case studies in Section 4. Finally, in Section 5 we present the proofs of the mathematical results as well as a number of intermediate derivations.

2 Nonnegative matrices and communication networks

A communication network consists of a number of sources and sinks connected together via links and routers. We assume that these links can be modelled as a constant propagation delay together with a queue, that the queue is operating according to a drop-tail discipline, and that all of the sources are operating a TCP-like congestion control algorithm. TCP (transmission control protocol) operates a window based congestion control algorithm. The TCP standard defines a variable *cwnd* called the congestion window. Each source uses this variable to track the number of sent unacknowledged packets that can be in transit at any time. When the window size is exhausted, the source must wait for an acknowledgement before sending a new packet. Congestion control is achieved by dynamically adapting the window size according to an additive-increase multiplicative-decrease (AIMD) law. Roughly speaking, the basic idea is for a source to gently probe the network for spare capacity by increasing the rate at which packets are inserted into the network, and to rapidly back-off the number of packets transmitted through the network when congestion is detected. We shall see that the *AIMD* paradigm with drop-tail queuing gives rise to networks whose dynamics can be accurately modelled as a positive linear system. While we are ultimately interested in general communication networks, for reasons of exposition it is useful to begin our discussion with a description of networks in which packet drops are synchronised (i.e. every source sees a drop at each congestion event). We show that many of the properties of communication networks that are of interest to network designers can be characterised by properties of a square matrix whose dimension is equal to the number of sources in the network. The approach is then extended to a model of unsynchronised networks. Even though the mathematical details are more involved, many of the qualitative characteristics of synchronised networks carry over to the non-synchronised case if interpreted in a stochastic fashion.

2.1 Synchronised communication networks

We begin our discussion by considering communication networks for which the following assumptions are valid: (i) at congestion every source experiences a packet drop; and (ii) each source has the same round-trip-time (RTT)[1]. In this case an exact model of the network dynamics may be found as follows [1]. Let $w_i(k)$ de-

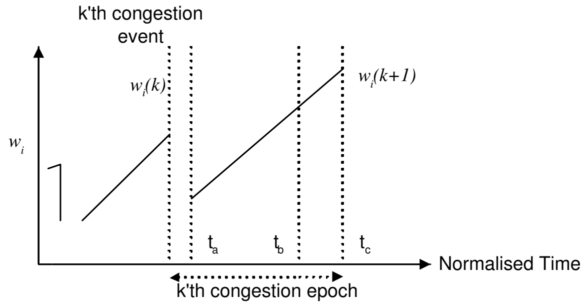


Figure 1: Evolution of window size

note the congestion window size of source i immediately before the k'th network congestion event is detected by the source. Over the k'th congestion epoch three important events can be discerned: $t_a(k), t_b(k)$ and $t_c(k)$; as depicted in Figure 1. The time $t_a(k)$ denotes the instant at which the number of unacknowledged packets in flight equals $\beta_i w_i(k)$; $t_b(k)$ is the time at which the bottleneck queue is full; and $t_c(k)$ is the time at which packet drop is detected by the sources, where time is measured in units of RTT[2]. It follows from the definition of the *AIMD* algorithm that the window evolution is completely defined over all time instants by knowledge of the $w_i(k)$ and the event times $t_a(k), t_b(k)$ and $t_c(k)$ of each congestion epoch. We therefore only need to investigate the behaviour of these quantities.

[1]One RTT is the time between sending a packet and receiving the corresponding acknowledgement when there are no packet drops.

[2]Note that measuring time in units of RTT results in a linear rate of increase for each of the congestion window variables between congestion events.

We assume that each source is informed of congestion one RTT after the queue at the bottleneck link becomes full; that is $t_c(k) - t_b(k) = 1$. Also,

$$w_i(k) \geq 0, \sum_{i=1}^{n} w_i(k) = P + \sum_{i=1}^{n} \alpha_i, \ \forall k > 0, \tag{1}$$

where P is the maximum number of packets which can be in transit in the network at any time; P is usually equal to $q_{max} + BT_d$ where q_{max} is the maximum queue length of the congested link, B is the service rate of the congested link in packets per second and T_d is the round-trip time when the queue is empty. At the $(k+1)$th congestion event

$$w_i(k+1) \quad = \quad \beta_i w_i(k) + \alpha_i[t_c(k) - t_a(k)]. \tag{2}$$

and

$$t_c(k) - t_a(k) \quad = \quad \frac{1}{\sum_{i=1}^{n} \alpha_i}[P - \sum_{i=1}^{n} \beta_i w_i(k)] + 1. \tag{3}$$

Hence, it follows that

$$w_i(k+1) \quad = \quad \beta_i w_i(k) + \frac{\alpha_i}{\sum_{j=1}^{n} \alpha_j}[\sum_{i=1}^{n}(1-\beta_i)w_i(k)], \tag{4}$$

and that the dynamics an entire network of such sources is given by

$$W(k+1) = AW(k), \tag{5}$$

where $W^T(k) = [w_1(k), \cdots, w_n(k)]$, and

$$A = \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \beta_n \end{bmatrix} + \frac{1}{\sum_{j=1}^{n} \alpha_j} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_n \end{bmatrix} \begin{bmatrix} 1-\beta_1 & 1-\beta_2 & \cdots & 1-\beta_n \end{bmatrix} \tag{6}$$

The matrix A is a positive matrix (all the entries are positive real numbers) and it follows that the synchronised network (5) is a positive linear system [2]. Many results are known for positive matrices and we exploit some of these to characterise the properties of synchronised communication networks. In particular, from the viewpoint of designing communication networks the following properties are important: (i) network fairness; (ii) network convergence and responsiveness; and (iii) network throughput. While there are many interpretations of network fairness, in this paper we concentrate on window fairness. Roughly speaking, window or pipe fairness refers to a steady state situation where n sources operating $AIMD$ algorithms have an equal number of packets P/n in flight at each congestion event; convergence refers to the existence of a unique fixed point to which the network dynamics converge; responsiveness refers to the rate at which the network converges to the fixed point; and throughput efficiency refers to the objective that the network operates at close to the bottleneck-link capacity. It is shown in [3, 4] that these properties can be deduced from the network matrix A. We briefly summarise here the relevant results in these papers.

Theorem 2.1 *[1, 4] Let A be defined as in Equation (6). Then A is a column stochastic matrix with Perron eigenvector $x_p^T = [\frac{\alpha_1}{1-\beta_1}, ..., \frac{\alpha_n}{1-\beta_n}]$ and whose eigenvalues are real and positive. Further, the network converges to a unique stationary point $W_{ss} = \Theta x_p$, where Θ is a positive constant such that the constraint (1) is satisfied; $\lim_{k\to\infty} W(k) = W_{ss}$; and the rate of convergence of the network to W_{ss} is bounded by the second largest eigenvalue of A.*

The following facts may be deduced from the above Theorem.

(i) **Fairness:** Window fairness is achieved when the Perron eigenvector x_p is a scalar multiple of the vector $[1, ..., 1]$; that is, when the ratio $\frac{\alpha_i}{1-\beta_i}$ does not depend on i. Further, since it follows for conventional TCP-flows ($\alpha = 1, \beta = 1/2$) that $\alpha = 2(1-\beta)$, any new protocol operating an *AIMD* variant that satisfies $\alpha_i = 2(1-\beta_i)$ will be TCP-fair.

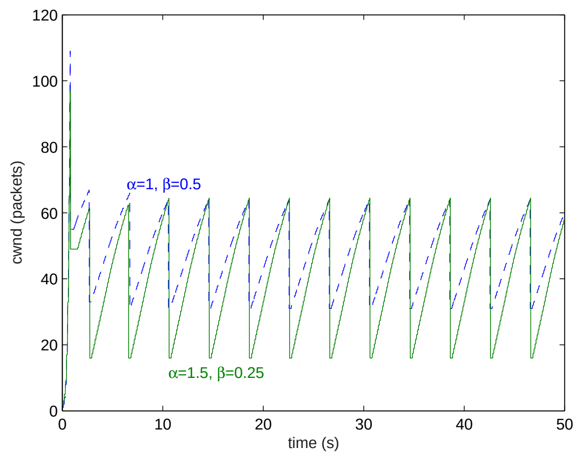


Figure 2: Example of window fairness between two TCP sources with different increase and decrease parameters (NS simulation, network parameters: 10Mb bottleneck link, 100ms delay, queue 40 packets.)

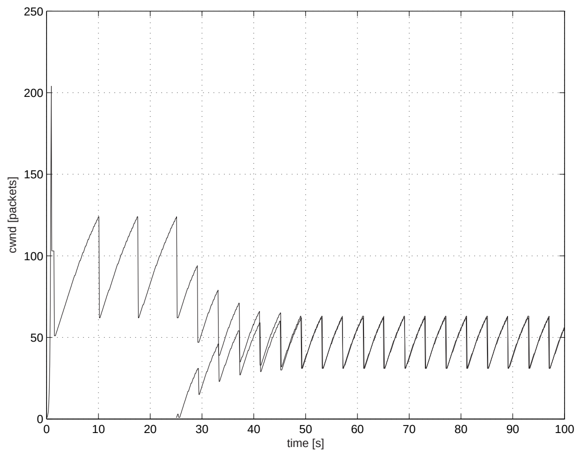


Figure 3: NS packet-level simulation ($\alpha_i = 1$, $\beta_i = 0.5$, dumb-bell with 10Mbs bottleneck bandwidth, 100ms propagation delay, 40 packet queue).

(ii) **Network responsiveness:** The magnitude of the second largest eigenvalue λ_{n-1} of the matrix A bounds the convergence properties of the entire network. It is shown in [4] that all the eigenvalues of A are real and positive and lie in the interval $[\beta_1, 1]$, where the β_i are ordered as $0 < \beta_1 \leq \beta_2 \leq .... \leq \beta_{n-1} \leq \beta_n < 1$. In particular, the second largest eigenvalue is bounded by $\beta_{n-1} \leq \lambda_{n-1} \leq \beta_n$. Consequently, fast convergence to the equilibrium state (the Perron eigenvector) is guaranteed if the largest backoff factor in the network is small. Further, we show in [4] that the network rise-time when measured in number of congestion epochs is bounded by $n_r = \log(0.95)/\log(\lambda_{n-1})$. In the special case when $\beta_i = 0.5$ for all i, $n_r \approx 4$; see for example Figure 3. Note that n_r gives the number of congestion epochs until the network dynamics have converged to 95 % of the final network state: the actual time to reach this state depends on the duration of the congestion epochs which is ultimately dependent on the α_i.

(iii) **Network throughput :** At a congestion event the network bottleneck is operating at link capacity and the total data throughput through the bottleneck link is given by

$$R(k)^- = \frac{\sum_i^n w_i(k)}{T_d + \frac{q_{max}}{B}} \tag{7}$$

where B is the link capacity, q_{max} is the bottleneck buffer size, T_d is the round-trip-time when the bottleneck queue is empty and $T_d + q_{max}/B$ is the round-trip time when the queue is full. After backoff, the data throughput is given by

$$R(k)^+ \quad = \quad \frac{\sum_i^n \beta_i w_i(k)}{T_d} \tag{8}$$

under the assumption that the bottleneck buffer empties. It is evident that if the sources backoff too much, data throughput will suffer as the queue remains empty for a period of time and the link operates below its maximum rate. A simple method to ensure maximum throughput is to equate both rates, which may be achieved by the following choice of the β_i:

$$\beta_i = \frac{T_d}{T_d + \frac{q_{max}}{B}} = \frac{RTT_{min}}{RTT_{max}}. \tag{9}$$

(iv) **Maintaining fairness :** Note that setting $\beta_i = \frac{RTT_{min}}{RTT_{max}}$ requires a corresponding adjustment of α_i if it is not to result in unfairness. Both network fairness and TCP-fairness are ensured by adjusting α_i according to $\alpha_i = 2(1 - \beta_i)$.

Comment 1: Networks of synchronised sources and drop-tail queues have already been the subject of several studies [5, 6, 7, 8, 9]. The novelty of our approach is that we use facts from the theory of positive matrices to analyse not only the network steady-state behaviour but also the network dynamics, directly relating the qualitative properties of synchronised networks to source and network parameters.

2.2 Models of unsynchronised network

The preceding discussion illustrates the relationship between important network properties and the eigensystem of a positive matrix. Unfortunately, the assumptions under which these results are derived, namely of source synchronisation and uniform RTT, are quite restrictive (although they may, for example, be valid in many long-distance networks [10]). It is therefore of great interest to extend our approach to more general network conditions. As we will see the model that we obtain shares many structural and qualitative properties of the synchronized model described above. To distinguish variables, we will from now on denote the nominal parameters of the sources used in the previous section by α_i^s, β_i^s, $i = 1, \ldots, n$. Here the index s may remind the reader that these parameters describe the *synchronized case*, as well as that these are the parameters that are chosen by each *source*.

Consider the general case of a number of sources competing for shared bandwidth in a generic dumbbell topology (where sources may have different round-trip times and drops need not be synchronised). The evolution of the *cwnd* of a typical source as a function of time, over the $k'th$ congestion epoch, is depicted in Figure 4. As before a number of important events may be discerned, where we

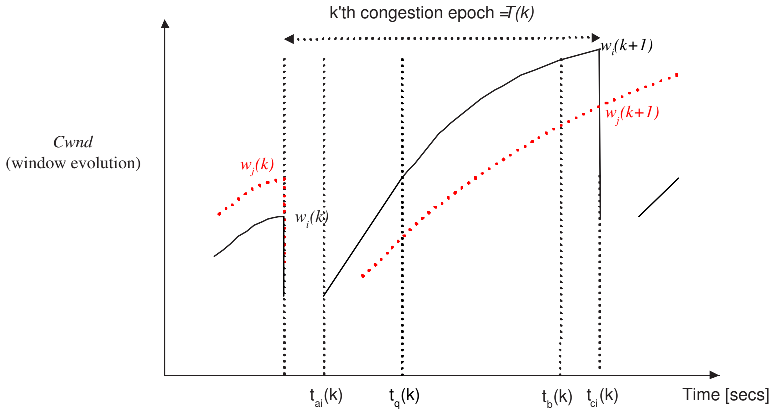


Figure 4: Evolution of window size over a congestion epoch. $T(k)$ is the length of the congestion epoch in seconds.

now measure time in seconds, rather than units of RTT. Denote by $t_{ai}(k)$ the time at which the number of packets in flight belonging to source i is equal to $\beta_i^s w_i(k)$; $t_q(k)$ is the time at which the bottleneck queue begins to fill; $t_b(k)$ is the time at which the bottleneck queue is full; and $t_{ci}(k)$ is the time at which the i'th source is informed of congestion. In this case the evolution of the i'th congestion window variable does not evolve linearly with time after t_q seconds due to the effect of the bottleneck queue filling and the resulting variation in RTT; namely, the RTT of the i'th source increases according to $RTT_i(t) = T_{d_i} + q(t)/B$ after t_q, where T_{d_i} is the RTT of source i when the bottleneck queue is empty and $0 \leq q(t) \leq q_{\max}$ denotes the number of packets in the queue. Note also that we do not assume that every source experiences a drop when congestion occurs. For example, a situation is depicted in Figure 4 where the i'th source experiences congestion at the end of the epoch whereas the j'th source does not.

Given these general features it is clear that the modelling task is more involved than in the synchronised case. Nonetheless, it is possible to relate $w_i(k)$ and $w_i(k+1)$ using a similar approach to the synchronised case by accounting for the effect of non-uniform RTT's and unsynchronised packet drops as follows.

(i) Unsynchronised source drops : Consider again the situation depicted in Figure 4. Here, the i'th source experiences congestion at the end of the epoch whereas the j'th source does not. This corresponds to the i'th source reducing its congestion window by the factor β_i^s after the $k+1$'th congestion event, and the j'th source not adjusting its window size at the congestion event. We therefore allow the back-off factor of the i'th source to take one of two values at the k'th congestion event.

$$\beta_i(k) \quad \in \quad \{\beta_i^s, 1\}, \tag{10}$$

corresponding to whether the source experiences a packet loss or not.

(ii) Non-uniform RTT : Due to the variation in round trip time, the congestion window of a flow does not evolve linearly with time over a congestion epoch. Nevertheless, we may relate $w_i(k)$ and $w_i(k+1)$ linearly by defining an average rate $\alpha_i(k)$ depending on the k'th congestion epoch:

$$\alpha_i(k) := \frac{w_i(k+1) - \beta_i(k)w(k)}{T(k)}, \tag{11}$$

where $T(k)$ is the duration of the k'th epoch measured in seconds. Equivalently we have

$$w_i(k+1) = \beta_i(k)w_i(k) + \alpha_i(k)T(k). \tag{12}$$

In the case when $q_{max} << BT_{d_i}$ $, i = 1, \dots, n$, the average α_i are (almost) independent of k and given by $\alpha_i(k) \approx \alpha_i^s / T_{d_i}$ for all $k \in \mathbb{N}, i = 1, \dots, n$.

The situation when

$$\alpha_i \approx \frac{\alpha_i^s}{T_{d_i}}, \quad i = 1, \dots, n \tag{13}$$

is of considerable practical importance and such networks are the principal concern of this paper. This corresponds to the case of a network whose bottleneck buffer is small compared with the delay-bandwidth product for all sources utilising the congested link. Such conditions prevail on a variety of networks; for example networks with large delay-bandwidth products, and networks where large jitter and/or latency cannot be tolerated. In view of (10) and (12) a convenient representation of the network dynamics is obtained as follows. At congestion the bottleneck link is operating at its capacity B, i.e.,

$$\sum_{i=1}^{n} \frac{w_i(k) - \alpha_i}{RTT_{i,max}} = B, \tag{14}$$

where $RTT_{i,max}$ is the RTT experienced by the i'th flow when the bottleneck queue is full. Note, that $RTT_{i,max}$ is independent of k. Setting $\gamma_i := (RTT_{i,max})^{-1}$

we have that

$$\sum_{i=1}^{n} \gamma_i w_i(k) = B + \sum_{i=1}^{n} \gamma_i \alpha_i \, . \tag{15}$$

By interpreting (15) at $k+1$ and inserting (12) for $w_i(k+1)$ it follows furthermore that

$$\sum_{i=1}^{n} \gamma_i \beta_i(k) w_i(k) + \gamma_i \alpha_i T(k) = B + \sum_{i=1}^{n} \gamma_i \alpha_i \, . \tag{16}$$

Using (15) again it follows that

$$T(k) = \frac{1}{\sum_{i=1}^{n} \gamma_i \alpha_i} \left(\sum_{i=1}^{n} \gamma_i (1 - \beta_i(k)) w_i(k) \right) . \tag{17}$$

Inserting this expression into (12) we finally obtain

$$w_i(k+1) = \beta_i(k) w_i(k) + \frac{\alpha_i}{\sum_{j=1}^{n} \gamma_j \alpha_j} \left(\sum_{j=1}^{n} \gamma_j (1 - \beta_j(k)) w_j(k) \right) . \tag{18}$$

and the dynamics of the entire network of sources at the k-th congestion event are described by

$$W(k+1) = A(k) W(k), \; A(k) \in \{A_1, ..., A_m\}. \tag{19}$$

where

$$\begin{aligned} A(k) \;\; &= \begin{bmatrix} \beta_1(k) & 0 & \cdots & 0 \\ 0 & \beta_2(k) & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \beta_n(k) \end{bmatrix} \\ &+ \frac{1}{\sum_{j=1}^{n} \gamma_j \alpha_j} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_n \end{bmatrix} \begin{bmatrix} \gamma_1(1-\beta_1(k)), & \ldots, & \gamma_n(1-\beta_n(k)) \end{bmatrix} , \end{aligned} \tag{20}$$

and where $\beta_i(k)$ is either 1 or β_i^s. The non-negative matrices $A_2, .., A_m$ are constructed by taking the matrix A_1,

$$A_1 \;\; = \begin{bmatrix} \beta_1^s & 0 & \cdots & 0 \\ 0 & \beta_2^s & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \beta_n^s \end{bmatrix} + \frac{1}{\sum_{j=1}^{n} \gamma_j \alpha_j} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_n \end{bmatrix} \begin{bmatrix} \gamma_1(1-\beta_1^s), & \ldots, & \gamma_n(1-\beta_n^s) \end{bmatrix} ,$$

and setting some, but not all, of the β_i to 1. This gives rise to $m = 2^n - 1$ matrices associated with the system (19) that correspond to the different combinations of source drops that are possible. We denote the set of these matrices by $\mathcal{A}$.

Example 1 (Two Unsynchronised Flows) *We illustrate the application of the model (19) for a network of two TCP flows. Possible combinations of packet drops are: (i) both flows experience a packet drop at a congestion event; (ii) flow 1 experiences a drop only; and (iii) flow 2 experiences a drop only . The associated set $\mathcal{A}$ is $\{A_1, A_2, A_3\}$ with*

$$
\begin{aligned}
A_1 &= \begin{bmatrix} \beta_1^s + \frac{\gamma_1\alpha_1}{\gamma_1\alpha_1+\gamma_2\alpha_2}(1-\beta_1^s) & \frac{\gamma_1\alpha_1}{\gamma_1\alpha_1+\gamma_2\alpha_2}(1-\beta_2^s) \\ \frac{\gamma_2\alpha_2}{\gamma_1\alpha_1+\gamma_2\alpha_2}(1-\beta_1^s) & \beta_2^s + \frac{\gamma_2\alpha_2}{\gamma_1\alpha_1+\gamma_2\alpha_2}(1-\beta_2^s) \end{bmatrix} \\
A_2 &= \begin{bmatrix} \beta_1^s + \frac{\gamma_1\alpha_1}{\gamma_1\alpha_1+\gamma_2\alpha_2}(1-\beta_1^s) & 0 \\ \frac{\gamma_2\alpha_2}{\gamma_1\alpha_1+\gamma_2\alpha_2}(1-\beta_1^s) & 1 \end{bmatrix} \\
A_3 &= \begin{bmatrix} 1 & \frac{\gamma_1\alpha_1}{\gamma_1\alpha_1+\gamma_2\alpha_2}(1-\beta_2^s) \\ 0 & \beta_2^s + \frac{\gamma_2\alpha_2}{\gamma_1\alpha_1+\gamma_2\alpha_2}(1-\beta_2^s) \end{bmatrix}
\end{aligned}
$$

Finally we note that another, sometimes very useful, representation of the network dynamics can be obtained by considering the evolution of scaled window sizes at congestion; namely, by considering the evolution of $W_\gamma^T(k) = [\gamma_1 w_1(k), \gamma_2 w_2(k), ..., \gamma_n w_n(k)]$. Here one obtains the following description of the network dynamics:

$$
W_\gamma(k+1) = \bar{A}(k) W_\gamma(k),\ \bar{A}(k) \in \bar{\mathcal{A}} = \{\bar{A}_1, ..., \bar{A}_m\},\ \ m = 2^n - 1, \tag{21}
$$

where the $\bar{A}_i$ are obtained by the similarity transformation associated with the change of variables, in particular

$$
\bar{A}_1 = \begin{bmatrix} \beta_1^s & 0 & \cdots & 0 \\ 0 & \beta_2^s & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \beta_n^s \end{bmatrix} + \frac{1}{\sum_{j=1}^n \gamma_j\alpha_j} \begin{bmatrix} \gamma_1\alpha_1 \\ \gamma_2\alpha_2 \\ \cdots \\ \gamma_n\alpha_n \end{bmatrix} \begin{bmatrix} 1-\beta_1^s & 1-\beta_2^s & \cdots & 1-\beta_n^s \end{bmatrix}.
$$

As before the non-negative matrices $\bar{A}_2, .., \bar{A}_m$ are constructed by taking the matrix $\bar{A}_1$ and setting some, but not all, of the β_i^s to 1. All of the matrices in the set $\bar{\mathcal{A}}$ are now column stochastic; for convenience we use this representation of the network dynamics to prove the main mathematical results presented in this paper.

<u>Comment 2:</u> Before proceeding we note that networks of unsynchronised sources have also been the subject of wide study in the TCP community: see [11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 20] and the accompanying references for further details. While most of this work has concentrated on developing and analysing

TCP models that are based upon fluid analogies, several authors have recently developed hybrid systems models of networks with a single bottleneck link which employ AIMD congestion control mechanisms: most notably by Hespanha [22] and Baccelli and Hong [23]. We note that the model derived in [23] is similar to the model presented here. In particular, under mild assumptions, the sets of solutions of the model in [23] and of (19) coincide, so that the results of that reference are immediately applicable to the model presented here. However, whereas the model derived by Baccelli and Hong is also a random matrix model, their model is both affine and the homogeneous (linear) part is characterised by general matrices (namely, not by non-negative matrices). We will shall see in Section 3 that the properties of linearity (no affine term) and nonnegativity are useful in characterising and interpreting the asymptotic properties of AIMD networks.

3 Main results

The ultimate objective of our work is to use the network model developed in Section 2 to establish design principles for the realisation of*AIMD* networks. In this section we present two results, both of which are derived from our network model in Section 2, which in a sense characterise the asymptotic behaviour of both long and short lived flows.

Preamble to main results

It follows from (19) that $W(k) = \Pi(k)W(0)$, where $\Pi(k) = A(k)A(k-1)....A(0)$. Consequently, the behaviour of $W(k)$, as well as the network fairness and convergence properties, are governed by the properties of the matrix product $\Pi(k)$. The objective of this section is to analyse the average behaviour of $\Pi(k)$ with a view to making concrete statements about these network properties. To facilitate analytical tractability we will make two mild simplifying assumptions.

Assumption 3.1 *The probability that $A(k) = A_i$ in (19) is independent of k and equals ρ_i.*

Comment 3: In other words Assumption 3.1 says that the probability that the network dynamics are described by $W(k+1) = A(k)W(k), A(k) = A_i$ over the k'th congestion epoch is ρ_i and that the random variables $A(k), k \in \mathbb{N}$ are independent and identically distributed (i.i.d.).

Given the probabilities ρ_i for $i \in \{1, ..., 2^n - 1\}$, one may then define the probability λ_j that source j experiences a backoff at the k'th congestion event as follows:

$$\lambda_j = \sum \rho_i \, ,$$

where the summation is taken over those i which correspond to a matrix in which the j'th source sees a drop. Or to put it another way, the summation is over those indices i for which the matrix A_i is defined with a value of $\beta_j \neq 1$.

Assumption 3.2 *We assume that $\lambda_j > 0$ for all $j \in \{1, ..., n\}$.*

Simply stated, Assumption 3.2 states that almost surely all flows must see a drop at some time (provided that they live for a long enough time).

Comment 4: A consequence of the above assumptions is that the probability that source j experiences a drop at the k'th congestion event is not independent of the other sources. For example, if the first $n-1$ sources do not see a drop then this implies that source n must see a drop (in accordance with the usual notion of a congestion event, we require at least one flow to see a drop at each congestion event). Hence, the events cannot be independent.

We now present two results that characterise the expected behaviour of *AIMD*-networks that satisfy Assumptions 3.1 and 3.2. The first characterises the ensemble average behaviour of flows, while the second characterises the time average behaviour. Both results indicate that the expected number of packets that are in-flight in the network is determined solely by the network parameters and the probability ρ_j.

Result 1. Ensemble average behaviour of TCP-flows

Theorem 3.1 *Consider the stochastic system defined in the preamble. Let $\Pi(k)$ be the random matrix product arising from the evolution of the first k steps of this system:*

$$\Pi(k) = A(k)A(k-1)....A(0).$$

Then, the expectation of $\Pi(k)$ is given by

$$E(\Pi(k)) = (\sum_{i=1}^{m} \rho_i A_i)^k; \quad (22)$$

and the asymptotic behaviour of $E(\Pi(k))$ satisfies

$$\lim_{k\to\infty} E(\Pi(k)) \;\; = \;\; x_p y_p^T, \quad (23)$$

where the vector x_p is given by $x_p^T = \Theta(\frac{\alpha_1}{\lambda_1(1-\beta_1)}, \frac{\alpha_2}{\lambda_2(1-\beta_2)}, ..., \frac{\alpha_n}{\lambda_n(1-\beta_n)})$, $y_p^T = (\gamma_1, ..., \gamma_n)$. Here $\Theta \in \mathbb{R}$ is chosen such that equation (15) is satisfied if w_i is replaced by $x_{pi} = \Theta\alpha_i/(\lambda_i(1-\beta_i))$.

Comment 5: Theorem 3.1 characterises the ensemble average behaviour of the congestion variable vector $W(k)$. The congestion variable vector of a network of flows starting from initial condition $W(0)$ and evolving for k congestion epochs is given by $W(k) = \Pi(k)W(0)$. The average window vector over many repetitions is given by $E(\Pi(k))W(0)$. Theorem 3.1 provides an expression for calculating this average in terms of the network parameters and the probabilities ρ_i. Furthermore, we have that as k becomes large $E(\Pi(k))W(0)$ tends asymptotically to $x_p y_p^T W(0)$. The rate of convergence of $E(\Pi(k))W(0)$ to this limiting value is bounded by the second largest eigenvalue of the matrix $\sum_{n=1}^{m} \rho_i A_i$. Note that $y_p^T W(0)$ is a constant independent of $W(0)$ (due to the constraint (15)). Hence, Theorem 3.1 states that the ensemble average window vector tends to a scalar multiple of x_p. When the $\lambda_i, i = 1, ..., n$ are equal, x_p is identical to the Perron eigenvector obtained in the case of synchronised networks; that is, the ensemble average in the unsynchronised case is identical to the fixed point in the deterministic situation where packet drops are synchronised. Many of the deduced properties for synchronised networks therefore carry over with a stochastic interpretation to the unsynchronised case: in particular, window fairness.

Comment 6: Theorem 3.1 is concerned with the expected behaviour of the source congestion windows at the k'th congestion epoch. For k sufficiently large the expected throughput before backoff can be approximated as $\sum_{i=1}^{k} \frac{\alpha_i}{\lambda_i(1-\beta_i)RTT_{i,max}}$. The worst case throughput after backoff (which occurs when the queue is on average empty after backoff) is approximately $\sum_{i=1}^{k} \frac{\alpha_i}{\lambda_i(1-\beta_i)RTT_{i,min}}$. An immediate consequence of this observation is that the bottleneck link is guaranteed to be operating at capacity (on average) for k large enough if $\beta_i = \frac{RTT_{i,min}}{RTT_{i,max}}$.

Result 2. Time average behaviour of flows

We now present the following theorem which is concerned with networks characterised by long-lived flows.

Theorem 3.2 *Consider the stochastic system defined in the preamble and let*

$$\overline{W}(k) \quad := \quad \frac{1}{k+1}\sum_{i=0}^{k} W(i) = \left(\frac{1}{k+1}\sum_{i=0}^{k}\Pi(i)\right) W(0),$$

and where

$$\Pi(k) = A(k)A(k-1)....A(0).$$

Then, with probability one

$$\lim_{k\to\infty} \overline{W}(k) \quad = \quad x_p y_p^T W(0) = \left(\sum_{j=1}^{n} \gamma_j w_j(0)\right) x_p, \qquad (24)$$

where the vectors x_p and y_p are as defined in Theorem 3.1.

Comment 7: Theorem 3.2 states that the time-average vector of window sizes almost surely converges asymptotically to a scalar multiple of x_p. Hence, x_p determines the time-averaged relative number of unacknowledged packets in the network from each source at each congestion event.

Comment 8: In view of Comment 6, it again follows that asymptotically, the time-averaged throughput through the bottleneck link will approach the capacity B for k sufficiently large if $\beta_i = \frac{RTT_{i,min}}{RTT_{i,max}}$.

4 Model Validation

The mathematical results derived in Section 3 are surprisingly simple when one considers the potential mathematical complexity of the unsynchronised network model (19). The simplicity of these results is a direct consequence of Assumptions 3.1 and 3.2. The objective of this section is therefore twofold; (i) to validate the unsynchronised model (19) in a general context; and (ii) to validate the analytical predictions of the model and thereby confirm that the aforementioned assumptions are appropriate in practical situations.

4.1 Two Unsynchronised Flows

We first consider the behaviour of two TCP flows in the dumbbell topology shown in Figure 5. Our analytic results are based upon two fundamental assumptions: (i) that the dynamics of the evolution of the source congestion windows can be accurately modelled by equation (19); and (ii) the allocation of packet drops amongst the sources at congestion can be described by random variables. We consider each of these assumptions in turn.

(i) *Accuracy of dynamics model.* A comparison of the predictions the model (19) against the output of a packet-level NS simulation is depicted in Figure 6. Here, the pattern of packet drops observed in the simulation is used to select the appropriate matrix $A(k)$ from the set $\mathcal{A}$ at each congestion event when evaluating (19). As can be seen, the model output is very accurate. Also plotted in Figure 7 is the evolution of the linear combination $\sum_{i=1}^{n} \gamma_i w_i$ where the γ_i are defined in Equation (15). It can be seen that $\sum_{i=1}^{n} \gamma_i w_i$ has the same value at each congestion event thereby validating the constraint (15) used in the model.

(ii) *Validity of random drop model.* It is well known that networks of TCP

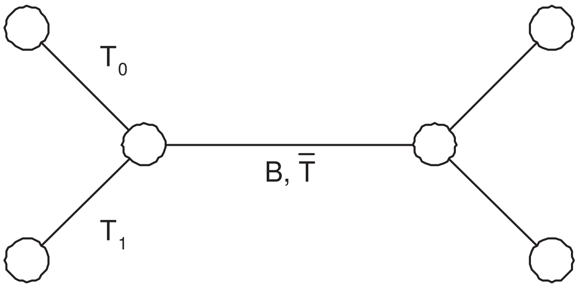


Figure 5: Dumbbell topology.

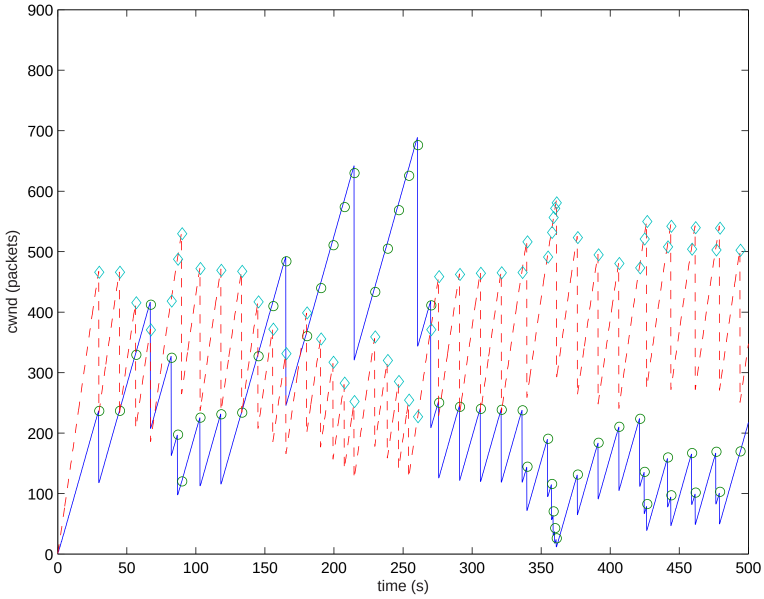


Figure 6: Predictions of the network model compared with packet-level NS simulation results. Key: ∘ flow 1 (model), ◊ flow 2 (model), - flow 1 (NS), – flow 2 (NS). Network parameters: B=100Mb, q_{max}=80 packets, $\bar{T}$=20ms, T_0=102ms; T_1=42ms; no background web traffic.

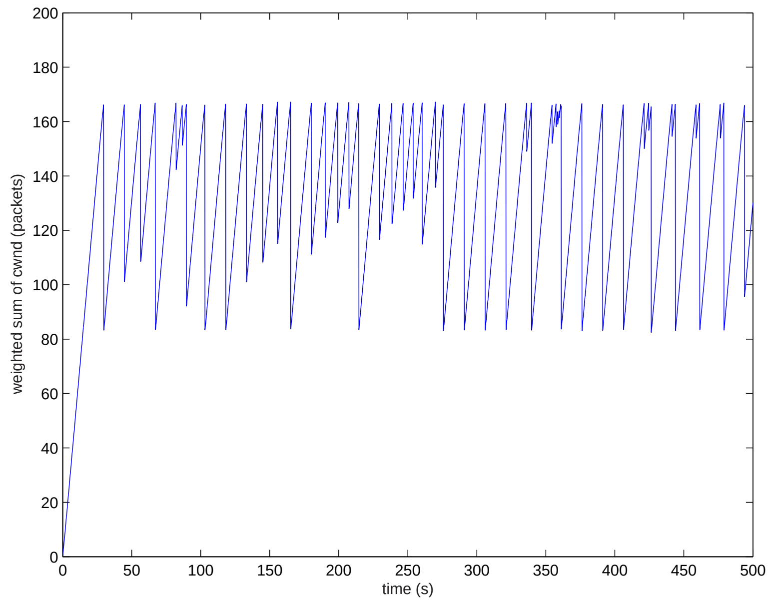


Figure 7: Evolution of $\sum_{i=1}^{n} \gamma_i w_i$. Network parameters: B=100Mb, q_{max}=80 packets, $\bar{T}$=20ms, T_0=102ms; T_1=42ms; no background web traffic.

flows with drop-tail queues can exhibit a rich variety of deterministic drop-behaviours [24]. However, most real networks carry at least a small amount web traffic. In Figure 8 we plot NS simulation results where the mean congestion window as the level of background web traffic is varied (background information on the web traffic generator in NS is described in [25]). To illustrate the impact of small amounts of web traffic, these results are given for a network condition where phase effects are particularly pronounced: the congestion window time histories with no web traffic are shown in Figure 9. It can be seen that the time histories appear to jump between two persistent regimes. In the first regime flow 1, which has a propagation delay of 122ms, achieves a larger congestion window than flow 2, which has a propagation time of only 62ms, see Figure 9(b). The reverse reverse is true in the second regime, see Figure 9(c). The impact of background web traffic is evident from Figure 10: despite its small volume, the effect of this traffic is enough to disrupt the coherent structure associated with phase effects and other complex phenomena previously observed in simulations of unsynchronised networks [24].

From the packet-based simulation results we can determine the proportion of congestion events corresponding to both flows simultaneously seeing a packet drop, flow 1 seeing a drop only, and flow 2 seeing a drop only. Using these estimates of the probabilities ρ_i, the mean congestion window

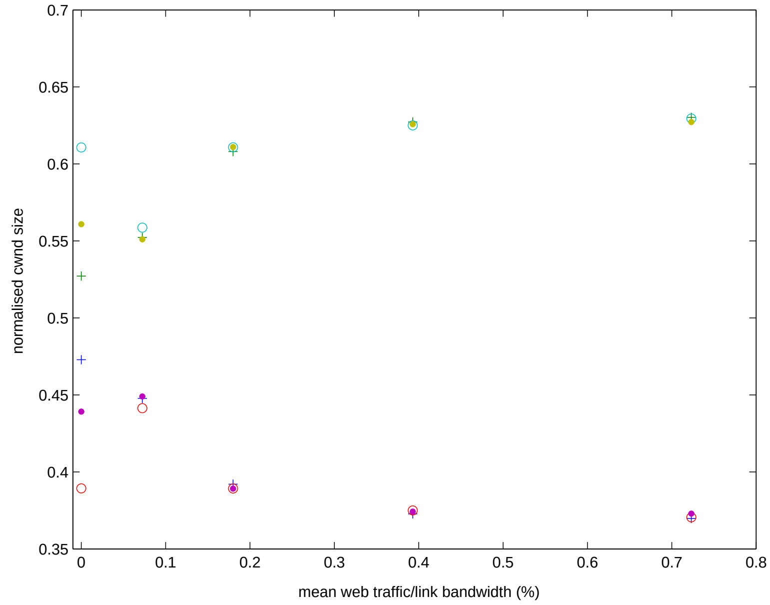
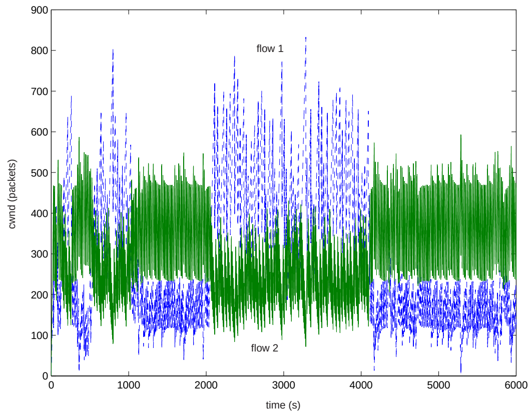
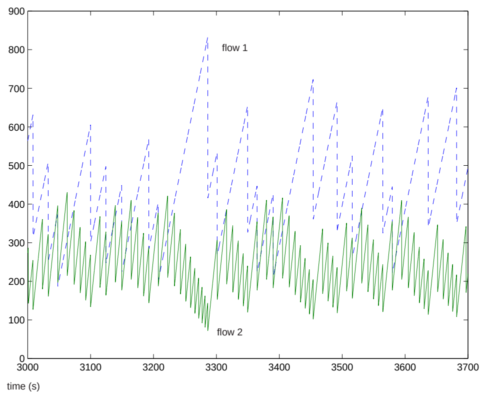


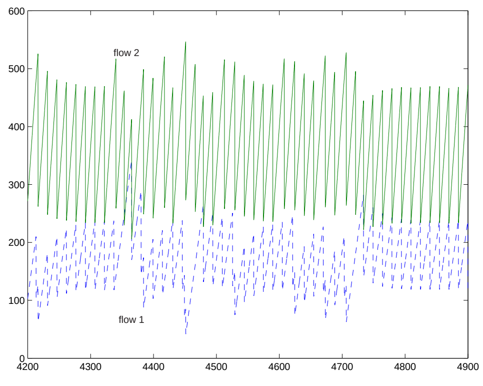
Figure 8: Variation of mean $w_i(k)$ with level of background web traffic in dumbbell topology of Figure 5. Key: $+NS$ simulation result; $\cdot$ mathematical model (19); $\circ$ Theorem 3.2. Network parameters: B=100Mb, q_{max}=80 packets, $\bar{T}$=20ms, T_0=102ms; T_1=42ms.



(a)

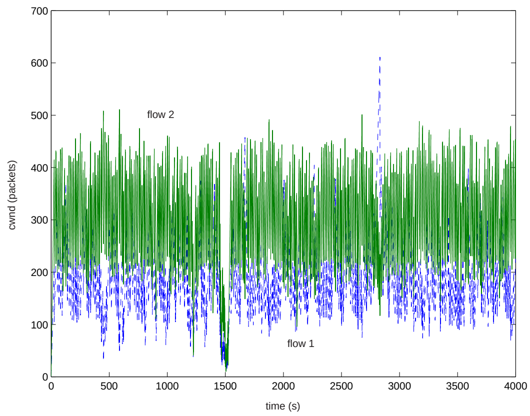


(b) $t = 3000s$ to $t = 3750s$

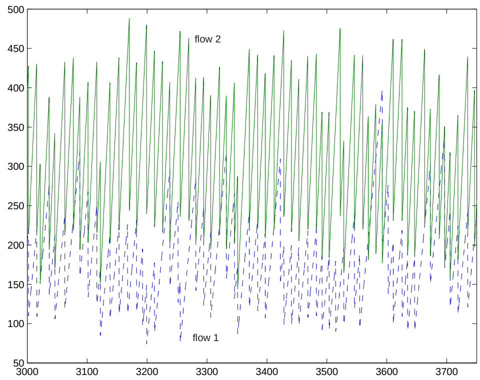


(c) $t = 4200s$ to $t = 4900s$

Figure 9: Congestion window time history corresponding to results in Figure 8 with no web traffic. Network parameters: B=100Mb, q_{max}=80 packets, $\bar{T}$=20ms, T_0=102ms; T_1=42ms.



(a)



(b) $t = 3000s$ to $t = 3750s$

Figure 10: Congestion window time history corresponding to results in Figure 8 with 0.4% web traffic. Network parameters: B=100Mb, q_{max}=80 packets, $\bar{T}$=20ms, T_0=102ms; T_1=42ms.

can be estimated using expression (23) from Theorem 3.2. The resulting estimates are shown in Figure 11, and are also presented in tabular form in Table 1. It can be seen that there is close agreement between the packet-level simulation results and the predictions obtained using (23). The actual convergence of the simulation data to the mean values (i.e. the time average as a function of the length of the time history) is depicted in Figure 12.

Also shown in Figure 11 are the analytic predictions for the case where each source has an equal probability of backing off: namely, when $\lambda_i = \frac{1}{n}\ \forall i$. The corresponding ratio of the elements of the average congestion window vector is the same as that under the assumption of source synchronisation (it is important to note that patterns of packet drop other than synchronised drops can lead to the same distribution as long as the proportion of backoff events experienced by the two flows is the same). Observe that the resulting predictions are an accurate estimate of the mean congestion window size and that as the level of web traffic increases the mean window size approaches that in the synchronised case (see Figure 8).

Before proceeding we also present results from several other two-flow networks in Figures 13 and 14. As can be seen from the figures, the predictions of Theorem 3.2 and the NS-simulations are consistently in close agreement.

T_1 (ms)	Flow 1			Flow 2		
	NS Simulation	Model	Theorem 3.2	NS Simulation	Model	Theorem 3.2
2.0	0.1924	0.1908	0.1895	0.8076	0.8092	0.8105
12.0	0.2762	0.2757	0.2736	0.7238	0.7243	0.7264
22.0	0.3253	0.3235	0.3237	0.6747	0.6765	0.6763
42.0	0.3691	0.3654	0.3651	0.6309	0.6346	0.6349
62.0	0.4226	0.4230	0.4239	0.5774	0.5770	0.5761
82.0	0.4599	0.4605	0.4600	0.5401	0.5395	0.5400
102.0	0.4866	0.4901	0.4943	0.5134	0.5099	0.5057
122.0	0.5156	0.5071	0.5082	0.4844	0.4929	0.4918
142.0	0.5461	0.5406	0.5378	0.4539	0.4594	0.4622
162.0	0.5877	0.5813	0.5825	0.4123	0.4187	0.4175
252.0	0.6652	0.6627	0.6609	0.3348	0.3373	0.3391

Table 1: Tabular data for Figure 11. T_1 is the fixed delay associated with source 1 that is depicted in Figure 5. The first column for each flow gives the actual average window size as predicted by the NS simulator; the second column gives the predictions of the model (19); and the third column gives the long-time average predictions of Theorem 3.2.

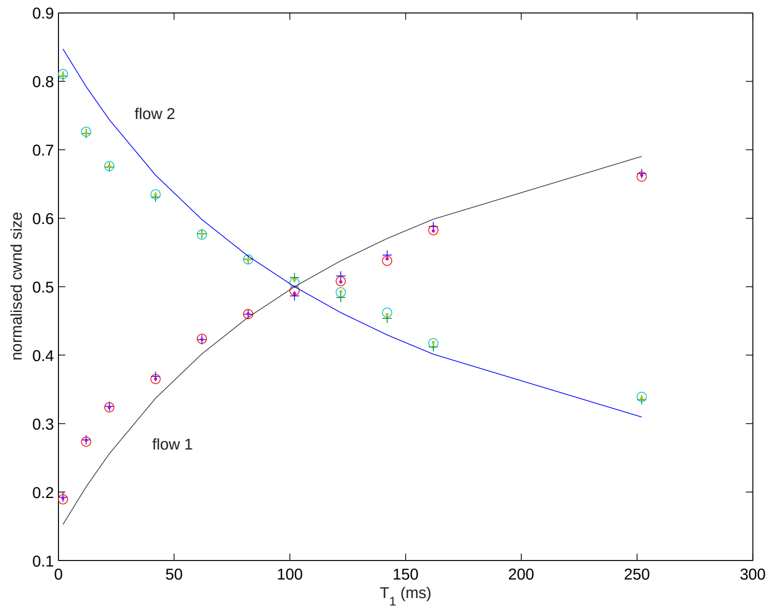


Figure 11: Variation of mean $w_i(k)$ with propagation delay T_1 in dumbbell topology of Figure 5. Key: $+NS$ simulation result; $\cdot$ mathematical model (19); $\circ$ Theorem 3.2; solid lines correspond to synchronised case. Network parameters: B=100Mb, q_{max}=80 packets, $\bar{T}$=20ms, T_0=102ms; approximately 0.5% bidirectional background web traffic.

4.2 Many Unsynchronised Flows

The foregoing results are for networks with two competing TCP sources. We note briefly that we have also validated our results against packet-level simulations for networks of up to five flows. As in the two flow case, and the simulation and analytical predictions are in close agreement; a sample of the results that we have collected is depicted in Figures 15-16.

4.3 Model Approximations

Predictions based upon the model (19) rely on knowledge of the rate α_i at which each of the sources increases its window size. In the case of networks with small queue sizes, Equation (13) gives a good approximation of these rates. However, this approximation neglects the curvature in the *cwnd* evolution induced by time-varying round-trip time and can therefore be expected to become less accurate as the queue provisioning increases. This behaviour is confirmed by the simulation results in Figure 17. For the network depicted in this example, it can be seen that the approximation (13) is accurate for queue sizes less than half the

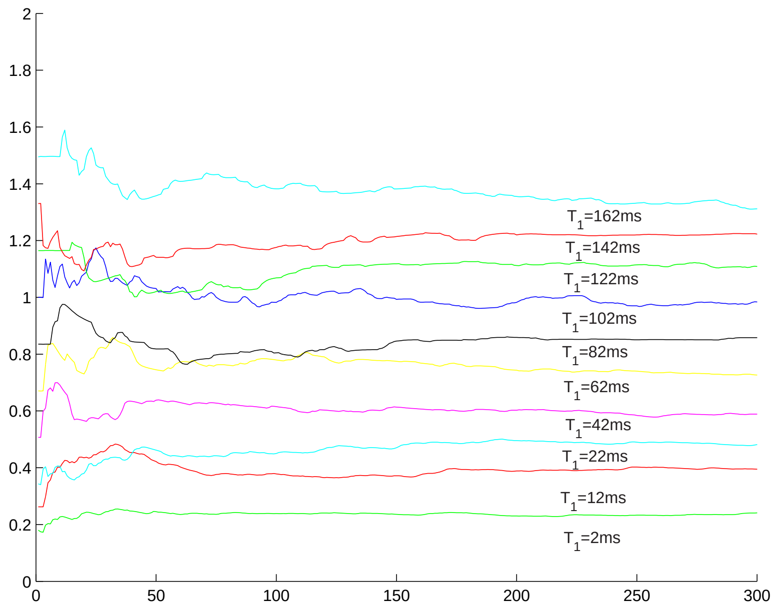


Figure 12: Convergence of the empirical mean of the window size to asymptotic values shown in Figure 11. *NS* simulation results; network parameters: B=100Mb, q_{max}=80 packets, $\bar{T}$=20ms, T_0=102ms.

effective delay-bandwidth product, but its accuracy degrades for larger queue sizes. We emphasize that the loss of predictive power is due to the validity of the approximation (13) and not the fidelity of the network model (19); a more accurate estimate of α_i would lead to better model performance. Techniques for approximating α_i when the queue is not small have already been explored in [26]. A sample result illustrating the effect of using approximations developed in this paper are shown in Figure 18.

The model (19) also neglects the fact that the number of packets in flight for TCP flows is quantised: namely, restricted to integer values, owing to the packet based nature of the traffic. Hence, the accuracy of the model (19) can be expected to degrade under network conditions where the peak window size w_i of a flow is small. This effect can be seen in the simulation results shown in Figure 19 and Figure 20.

5 Mathematical derivations

Theorem 3.1 and Theorem 3.2 follow from several interesting properties of the set of matrices $\mathcal{A} = \{A_1, ..., A_m\}$. Roughly speaking, these results may be classified as being algebraic or stochastic in nature. The purpose of this section is to

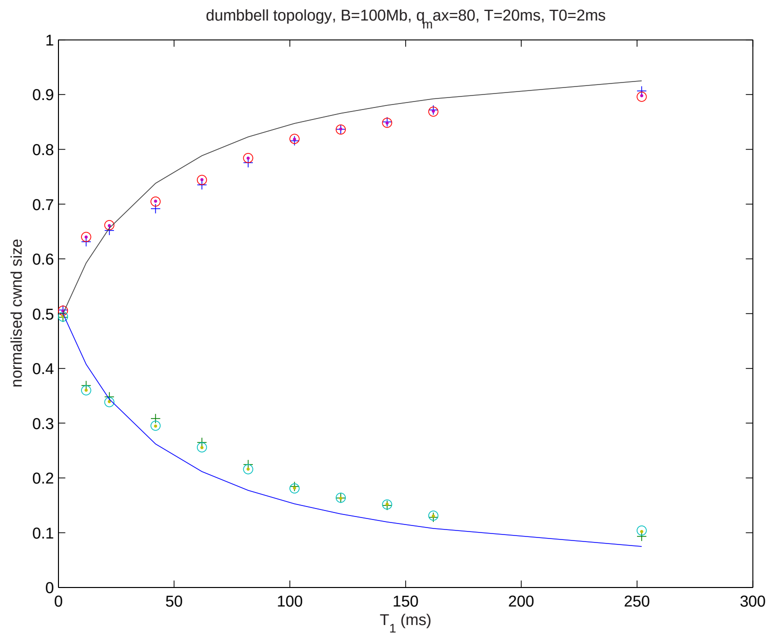


Figure 13: Variation of mean $w_i(k)$ with propagation delay T_1 in dumbbell topology of Figure 5. Key: $+NS$ simulation result; $\cdot$ mathematical model (19); $\circ$ Theorem 3.2; solid lines correspond to synchronised case. Network parameters: B=100Mb, q_{max}=80 packets, $\bar{T}$=20ms, T_0=2ms; approximately 0.5% bidirectional background web traffic.

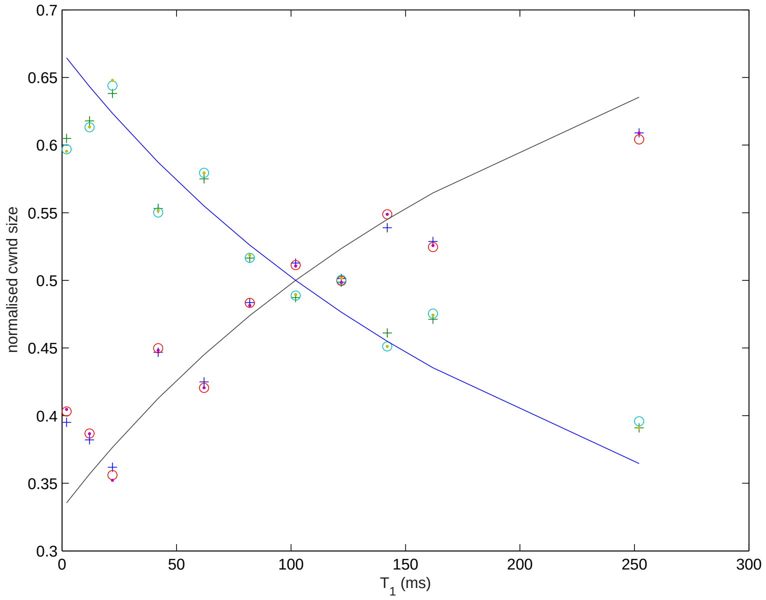


Figure 14: Variation of mean $w_i(k)$ with propagation delay T_1 in dumbbell topology of Figure 5. Key: $+NS$ simulation result; $\cdot$ mathematical model (19); $\circ$ Theorem 3.2; solid lines correspond to synchronised case. Network parameters: B=100Mb, q_{max}=80 packets, $\bar{T}$=100ms, T_0=102ms; approximately 0.5% bidirectional background web traffic.

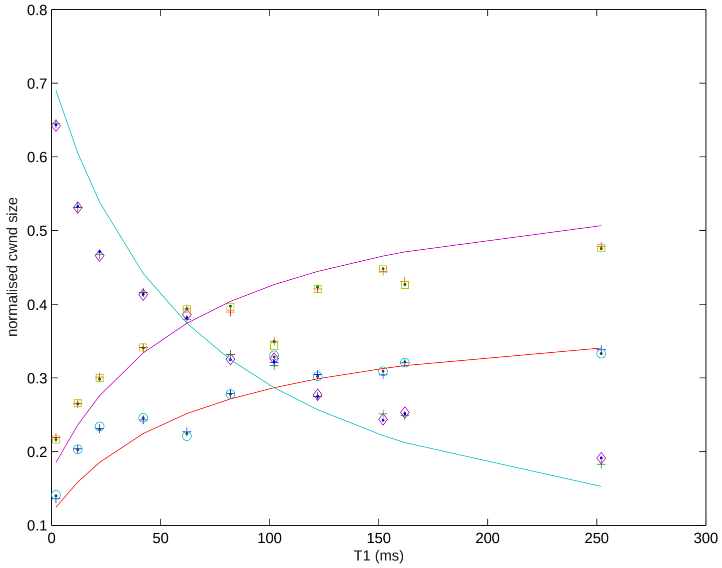


Figure 15: Variation of mean $w_i(k)$ with propagation delay T_1 in dumbbell topology with three TCP flows. Key: $+NS$ simulation result; $\cdot$ mathematical model (19); $\circ$, $\diamond$, $\square$ Theorem 3.2 flows 1, 2 and 3 respectively; solid lines correspond to synchronised case. Network parameters: B=100Mb, q_{max}=80 packets, $\bar{T}$=20ms, T_0=102ms, T_2=62ms; approximately 0.5% bidirectional background web traffic.

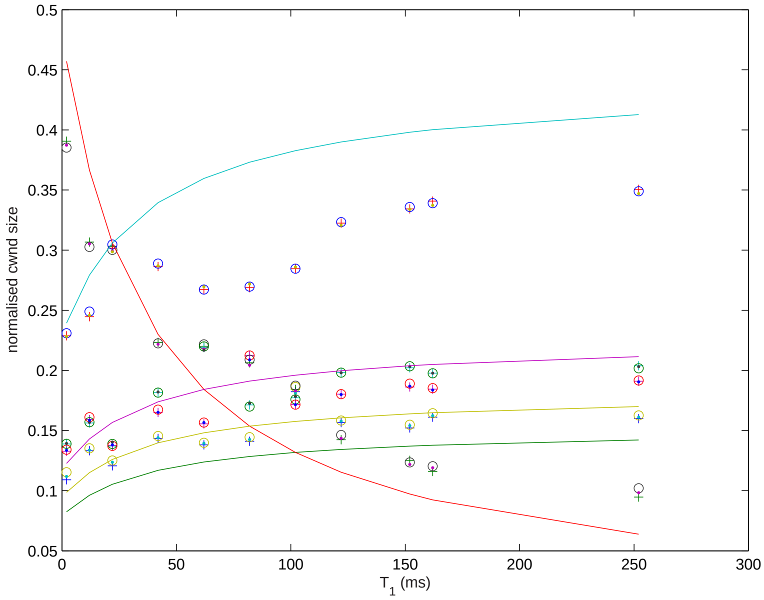


Figure 16: Variation of mean $w_i(k)$ with propagation delay T_1 in dumbbell topology with five TCP flows. Network parameters: B=100Mb, q_{max}=80 packets, $\bar{T}$=20ms, T_0=102ms, T_2=62ms; approximately 0.5% bidirectional background web traffic.

elucidate these properties and to use them to prove the results given in Section 3.

It was noted before that the matrices in the set $\mathcal{A}$ are not column stochastic. However, the matrices in this set are simultaneously similar to a set of column stochastic matrices under the transformation $\Gamma = \text{diag}[\gamma_1, ..., \gamma_n]$. For $A \in \mathcal{A}$, determined by a choice of parameters $\beta_1(A), \ldots, \beta_n(A)$ we have

$$\Gamma A \Gamma^{-1}$$

$$= \begin{bmatrix} \beta_1(A) & 0 & \cdots & 0 \\ 0 & \beta_2(A) & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \beta_n(A) \end{bmatrix} + \frac{1}{\sum_{j=1}^n \gamma_j \alpha_j} \begin{bmatrix} \gamma_1\alpha_1 \\ \gamma_2\alpha_2 \\ \cdots \\ \gamma_n\alpha_n \end{bmatrix} \begin{bmatrix} (1-\beta_1(A)), & \cdots, & (1-\beta_n(A)) \end{bmatrix},$$

and setting $\hat{\alpha}_j := \gamma_j \alpha_j, j = 1, \ldots, n$ we have

$$= \begin{bmatrix} \beta_1(A) & 0 & \cdots & 0 \\ 0 & \beta_2(A) & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \beta_n(A) \end{bmatrix} + \frac{1}{\sum_{j=1}^n \hat{\alpha}_j} \begin{bmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \cdots \\ \hat{\alpha}_n \end{bmatrix} \begin{bmatrix} (1-\beta_1(A)), & \cdots, & (1-\beta_n(A)) \end{bmatrix}.$$

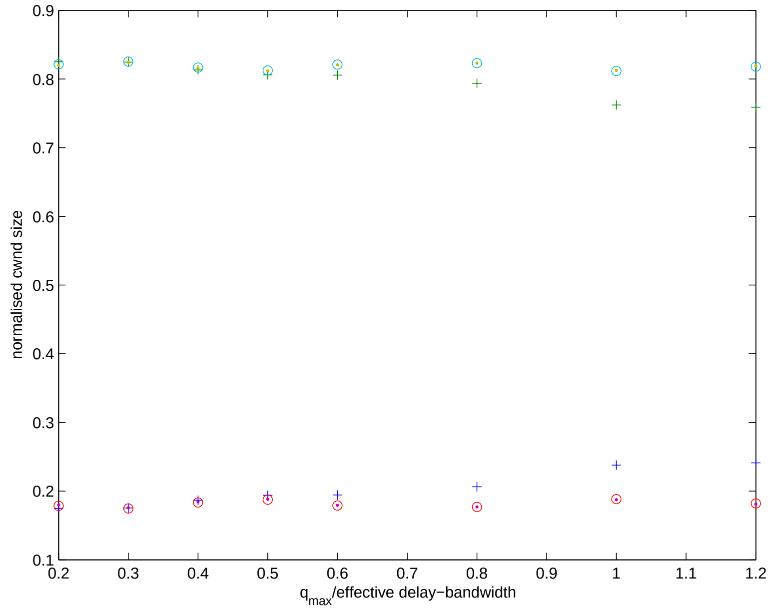


Figure 17: Accuracy of approximation (13) as queue provisioning is varied. The sum of the congestion window at each congestion event is 200 packets (the effective delay-bandwidth product). Key: $+NS$ simulation result; $\cdot$ mathematical model (19); $\circ$ Theorem 3.2; network parameters: B=100Mb, $\bar{T}$=20ms, T_0=102ms, $T_1 = 2ms$.

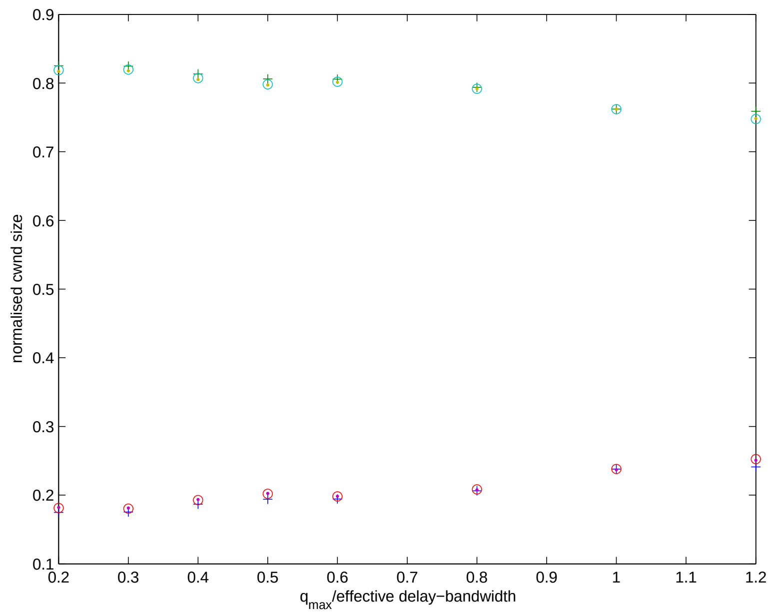


Figure 18: Accuracy of refined approximation to (11) from [26] as queue provisioning is varied. Effective delay-bandwidth product is 200 packets. Key: + *NS* simulation result; · mathematical model (19); ∘ Theorem 3.2; network parameters: B=100Mb, $\bar{T}$=20ms, T_0=102ms, $T_1 = 2ms$.

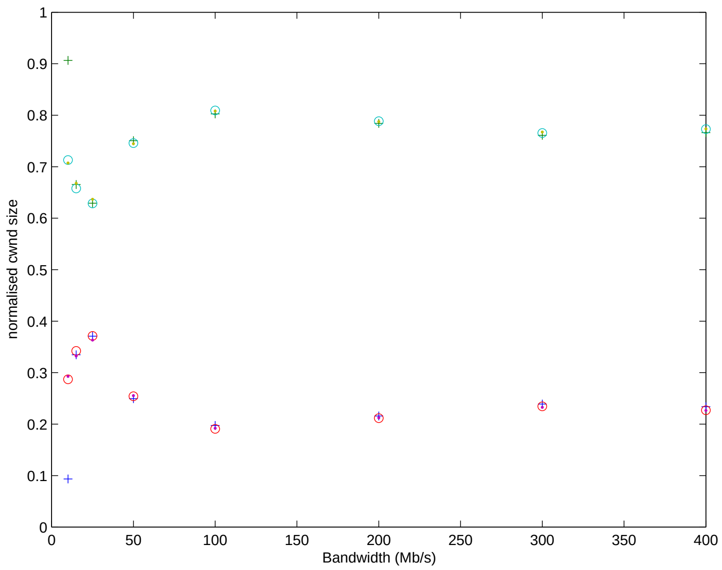


Figure 19: Sensitivity of model accuracy to effective pipe size. Bottleneck bandwidth B is varied and q_{max} is varied in proportion. Key: $+NS$ simulation result; $\cdot$ mathematical model (19); $\circ$ Theorem 3.2; network parameters: $q_{max} = 0.5B\bar{T}$, $\bar{T}$=20ms, T_0=102ms, $T_1 = 2ms$.

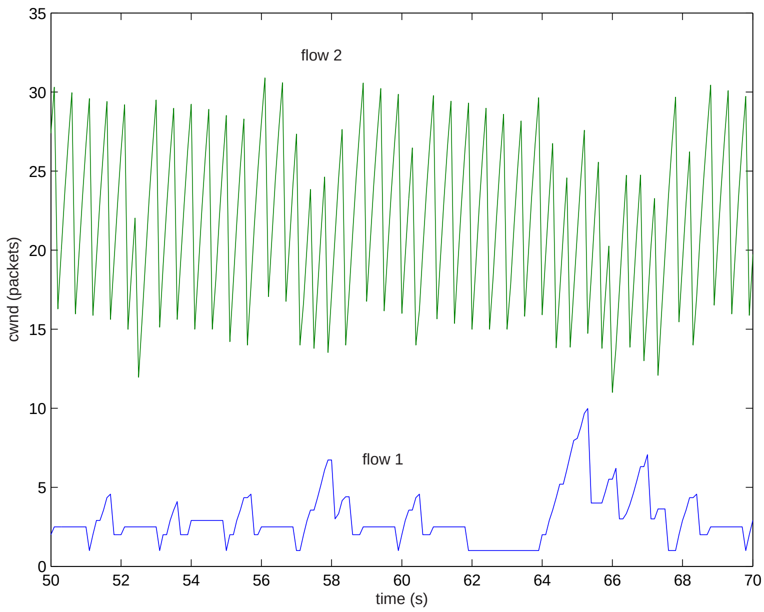


Figure 20: Typical *cwnd* time histories for small bottleneck link bandwidth. *NS* simulation; network parameters: B=10Mbs, q_{max} = 10 packets, $\bar{T}$=20ms, T_0=102ms, $T_1 = 2ms$.

It is easy to see that the transformed matrices are column stochastic. We shall exploit this observation in the sequel as column stochastic matrices are easier to deal with than nonstochastic ones. In view of this fact we note that the Perron eigenvector of $\Gamma A_1 \Gamma^{-1}$ is given by $\bar{x}_p^T = (\frac{\hat{\alpha}_1}{\lambda_1(1-\beta_1)}, \frac{\hat{\alpha}_2}{\lambda_2(1-\beta_2)}, ..., \frac{\hat{\alpha}_n}{\lambda_n(1-\beta_n)})$, and that the corresponding Perron eigenvector of A_1 is $x_p^T = (\frac{\alpha_1}{\lambda_1(1-\beta_1)}, \frac{\alpha_2}{\lambda_2(1-\beta_2)}, ..., \frac{\alpha_n}{\lambda_n(1-\beta_n)})$. In the sequel we will derive results that are expressed in terms of $\bar{x}_p$. These correspond to the dynamics of the system (21) and refer to the stochastic properties of the vector $\overline{W}_\gamma(k)$. The corresponding results for the system (19) are directly deduced from these results by similarity.

5.1 Algebraic properties of the set $\mathcal{A}$

We will from now on assume without loss of generality that the matrices in the set $\mathcal{A}$ are column stochastic, which corresponds to the case $\gamma_1 = \ldots = \gamma_n = 1$. Should this not be the case we can always apply the transformation Γ to obtain this property, which just amounts to a rescaling of the α_i. In the derivation of the main results of this paper we make frequent use of the fact that the matrices in the set $\mathcal{A}$, and products of matrices in this set, are nonnegative and in particular column stochastic. This observation implies the existence of an $n-1$ dimensional subspace that is invariant under $\mathcal{A}$. We will also see that the matrices in this set can be simultaneously transformed into block triangular form with an $n-1$ dimensional symmetric block. Given these observations, we will then show under mild assumptions that the distance of a matrix product of length k, constructed from matrices in $\mathcal{A}$, from the set of rank-1 matrices converges asymptotically to zero as k increases.

Lemma 5.1 *Consider the set $\mathcal{A}$. Then, there exists an $n-1$ dimensional subspace invariant under $\mathcal{A}$.*

Proof. The row vector $v := [1, ..., 1]$ is a left eigenvector of all of the matrices in the set $\mathcal{A}$ as they are column stochastic. This implies that the $n-1$ dimensional subspace orthogonal to v is invariant under $\mathcal{A}$ [27]. ■

Lemma 5.2 *Consider the set of matrices $\mathcal{A}$. There exists a real non-singular transformation T such that for all $A \in \mathcal{A}$ we have*

$$T^{-1}AT \;=\; \begin{bmatrix} S & B \\ 0 & 1 \end{bmatrix}, \tag{25}$$

where $S \in \mathbb{R}^{(n-1)\times(n-1)}$ is symmetric, so that in particular the eigenvalues of S are real and of absolute value ≤ 1.

Proof. We denote $\alpha = \begin{bmatrix} \alpha_1 & \dots & \alpha_n \end{bmatrix}^T$ and $c_\alpha := (\sum_{j=1}^{n} \alpha_j)^{-1}$. Let $A = \Lambda + c_\alpha \alpha \beta^T \in \mathcal{A}$, where Λ is the diagonal matrix with entries equal to 1 or β_i^s and β is the corresponding vector with entries 0 or $1-\beta_i^s$ and where we denote . Consider the diagonal matrix

$$D \;=\; \begin{bmatrix} \frac{1}{\sqrt{\alpha_1}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{\alpha_2}} & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \dots & \frac{1}{\sqrt{\alpha_n}} \end{bmatrix}. \tag{26}$$

Then DAD^{-1} is a non-negative matrix with a left eigenvector given by $z_p = vD^{-1}$, (with v defined in the proof of the previous lemma). Further, it follows that, $DAD^{-1} = \Lambda + Dc_\alpha \alpha \beta^T D^{-1}$ and by inspection $D\alpha = z_p$.

We now chose an orthogonal matrix M whose last column is $z_p/\|z_p\|$. Then e_n^T (the n-th unit row vector) is a left eigenvector of $M^T DAD^{-1}M$, and furthermore

$$M^T DAD^{-1}M = M^T \Lambda M + c_\alpha M^T z_p \beta^T D^{-1} M\,.$$

Now as z_p is a multiple of the last column of M it follows that $M^T z_p = \|z_p\| e_n$ and hence the entries of $M^T z_p \beta^T D^{-1}$ are nonzero only in the last row. Thus using that e_n^T is a left eigenvector we have

$$M^T DAD^{-1} M = \begin{bmatrix} S & B \\ 0 & 1 \end{bmatrix}, \tag{27}$$

where $S \in \mathbb{R}^{(n-1)\times(n-1)}$ is equal to the upper left $(n-1)$-minor of $M^T \Lambda M$ and thus symmetric. The assertion follows by setting $T = D^{-1}M$. The eigenvalues of S are bounded in absolute value by 1 as the matrix A is column stochastic and thus has spectral radius equal to 1. ■

We denote the the of matrices S that appear as the upper left block in (25) by $\mathcal{S}$.

Corollary 5.1 *Consider the system (19). Then for each $S \in \mathcal{S}$ the function $V(z((k)) = z^T(k)z(k)$ is a quadratic Lyapunov function for the dynamic system*

$$\Sigma : z(k+1) = Sz(k), \tag{28}$$

i.e., for all solutions of Σ we have $V(z(k+1)) - V(z(k)) \le 0$ for all k.

Proof. The assertion follows immediately as the matrices $\{S_1, S_2, .., S_m\}$ are symmetric column stochastic matrices. ■

There are some interesting consequences of Corollary 5.1 for products of matrices from the set $\mathcal{S}$. As these matrices are symmetric and of norm less or equal to 1 they form what is called a *paracontracting* set of matrices. This property is defined by the requirement that

$$Sx \neq x \Leftrightarrow \|Sx\| < \|x\|, \quad \forall x \in \mathbb{R}^{n-1}, S \in \mathcal{S}. \tag{29}$$

This is true for our set $\mathcal{S}$, as the matrices $S \in \mathcal{S}$ are symmetric and of spectral radius at most 1. It is know [28], that finite sets of matrices that are paracontracting have *left convergent products*, i.e., for any sequence $\{S(k)\}_{k\in\mathbb{N}}$ in $\mathcal{S}$ the following limit exists

$$\lim_{k\to\infty} S(k)S(k-1)\dots S(0). \tag{30}$$

For related literature on paracontracting sets of matrices we refer to [29, 28] and references therein.

In the following we prove results on the convergence of products of the matrices in $\mathcal{A}$ to the set of column-stochastic matrices of rank 1. To this end it will be convenient to introduce a notation that identifies each matrix $A \in \mathcal{A}$ with the sources that do not see a drop in that congestion event. Let $\mathcal{I} \subset \{1, 2, \dots, n\}$ be the index set of sources not experiencing congestion at a congestion event. (Clearly, $\mathcal{I} = \{1, 2, \dots, n\}$ can be ignored, as this means that there is no congestion).

The matrix corresponding to an index set $\mathcal{I}$ is given by

$$A_{\mathcal{I}} = \operatorname{diag}(\beta_1(\mathcal{I}), \dots, \beta_n(\mathcal{I})) + c_\alpha \alpha \begin{bmatrix} 1 - \beta_1(\mathcal{I}) & \dots & 1 - \beta_n(\mathcal{I}) \end{bmatrix},$$

where $\beta_i(\mathcal{I}) = 1$, if $i \in \mathcal{I}$ and $\beta_i(\mathcal{I}) = \beta_i^s$ otherwise and $c_\alpha := (\sum_{j=1}^n \alpha_j)^{-1}$. We now recover our set of possible matrices by

$$\mathcal{A} := \{A_{\mathcal{I}} \mid \mathcal{I} \subsetneq \{1, 2, \dots, n\}\}, \tag{31}$$

which results in a set of $2^n - 1$ matrices, as it should. Note that all $A \in \mathcal{A}$ are column stochastic, so that they have an eigenvalue equal to 1 equal to the spectral radius.

If $\mathcal{I} \neq \emptyset$, i.e., if at least one source does not experience congestion, then the dimension of the eigenspace corresponding to 1 is equal to the number of sources not seeing the congestion event. To see this consider first the case that the first k sources $k \in \{1, \dots, n-1\}$ do not see a drop and the others do. In this case

$$A_{\{1,\dots,k\}} = \begin{bmatrix} I_{k\times k} & B \\ 0 & C \end{bmatrix}, \tag{32}$$

where $B > 0$ by definition. As the matrix is column stochastic this means that all columns of C sum to a value strictly less than one, and hence $r(C) \leq \|C\|_1 < 1$ and the claim follows for $A_{\{1,\dots,k\}}$. Now an arbitrary matrix $A_{\mathcal{I}}$, $\mathcal{I} \neq \emptyset$ may be brought into the form (32) by permutation of the index set and we have shown the desired property.

Note also that the eigenspace of $A_{\mathcal{I}}$ associated to the eigenvalue 1 is given by

$$V_{\mathcal{I}} = \operatorname{span}\{e_i \mid i \in \mathcal{I}\}, \tag{33}$$

where e_i denotes the i-th unit vector.

Let us briefly discuss the eigenspaces of $S_{\mathcal{I}}$ corresponding to the eigenvalue 1, which we denote by $V(S_{\mathcal{I}})$. If $\mathcal{I} = \emptyset, \{1\}, \dots \{n\}$, then as we have seen in (33), the multiplicity of 1 as an eigenvalue of $A_{\mathcal{I}}$ is 1, so from (25) we have that $r(S_{\mathcal{I}}) < 1$. In this case we will (with slight abuse of notation) set $V(S_{\mathcal{I}}) = \{0\}$. We denote the subspace orthogonal to $[1, \dots, 1]$ by $[1, \dots, 1]^\perp$. Recall from Lemma 5.1, that this is an invariant subspace of $A \in \mathcal{A}$. In general, we see from (33) and Lemma 5.1 that if $\mathcal{I} = \{k_1, k_2, \dots, k_l\} \subsetneq \{1, \dots, n\}$ then a basis for

$$\begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}^\perp \cap V_{\mathcal{I}} \tag{34}$$

is given, e.g. by

$$e_{k_1} - e_{k_2}, e_{k_1} - e_{k_3}, \dots, e_{k_1} - e_{k_l}.$$

Hence the eigenspace of $V(S_{\mathcal{I}})$ is spanned by

$$M^T D^{-1}(e_{k_1} - e_{k_2}), M^T D^{-1}(e_{k_1} - e_{k_3}), \dots, M^T D^{-1}(e_{k_1} - e_{k_l}). \tag{35}$$

From this it follows that

$$V(S_{\mathcal{I}_1}) \cap V(S_{\mathcal{I}_2}) = V(S_{\mathcal{I}_1 \cap \mathcal{I}_2}), \tag{36}$$

justifying our abuse of notation above. In particular, $V(S_{\mathcal{I}_1}) \cap V(S_{\mathcal{I}_2}) \neq \{0\}$ if and only if $\mathcal{I}_1 \cap \mathcal{I}_2$ contains at least 2 elements.

Proposition 5.1 *Let $\{S(k)\}_{k\in\mathbb{N}} \subset \mathcal{S}$ be a sequence with associated index sets $\mathcal{I}(k)$. The following statements are equivalent:*

(i) For all $z_0 \in \mathbb{R}^{n-1}$ it holds that

$$\lim_{k\to\infty} S(k)S(k-1)\dots S(0)z_0 = 0\,.$$

(ii) for all but one $l \in \{1,\dots,n\}$ it holds that for each $k \in \mathbb{N}$, there is an $k_1 > k$ with $l \notin \mathcal{I}(k_1)$.

(iii) If $\{A_1,\dots,A_s\} \subset \mathcal{A}$ are the matrices that appear infinitely often in the sequence $A_{\mathcal{I}(k)}$, then

$$\hat{A} := \frac{1}{s}\sum_{l=1}^{s} A_l$$

is a matrix that with the exception of at most one column has strictly positive entries.

If in (iii) the k-th column of $\hat{A}$ has zero entries then this column is equal to e_k.

Proof. (ii)⇔(iii): Note first that the k-th column of $\hat{A}$ is not equal to e_k, if and only if for one of the matrices A_l, $l = 1,\dots,m$ the corresponding column is not a unit vector. The assumption on the matrix A_l implies that the k-th source experiences a drop infinitely many times. Under assumption (iii) this is true for all but at most one column, which implies (ii). Conversely, under the assumption (ii) the k-th source experiences a drop infinitely many times. As there are only finitely many matrices in which the k-th column is not equal to e_k one of these appears infinitely often in the sequence of matrices and therefore in the definition of $\hat{A}$. This implies (iii).

(i)⇒(ii): If (ii) does not hold, then (without loss of generality) there is a $s \geq 0$ such that for all $t \geq s$ we have $\{1,2\} \subset \mathcal{I}(k)$. This implies that for all $t \geq s$ the matrix $S(k)$ has the eigenspace $M^T D^{-1}(e_1 - e_2)$ as an eigenspace corresponding to the eigenvalue one. Hence any z_0 such that $S(s)\dots S(0)z_0$ is a multiple of $M^T D^{-1}(e_1 - e_2)$ does not satisfy (i). Such a z_0 exists as all the matrices in $\mathcal{A}$ are invertible. This shows the assertion.

(ii)⇒(i): Denote $z(k) := S(k-1)\dots S(0)z_0$. Using paracontractivity of $\mathcal{S}$ and (30) it follows that

$$z_\infty := \lim_{k\to\infty} z(k)$$

exists. If $z_\infty = 0$ there is nothing to show. Otherwise, we claim that for some k_0 sufficiently large, it follows that $S_k z_\infty = z_\infty$ for all $k \geq k_0$. To this end note that because $\mathcal{S}$ is finite, there exists a constant $0 < r < 1$ such that for all $S \in \mathcal{S}$ we have

$$Sz_\infty \neq z_\infty \Rightarrow \|Sz_\infty\| < r\|z_\infty\|\,.$$

By convergence this implies for all $S \in \mathcal{S}$ and all k sufficiently large that

$$Sz_\infty \neq z_\infty \Rightarrow \|Sz(k)\| < \frac{1+r}{2}\|z_\infty\| < \|z_\infty\|\,.$$

On the other hand the sequence $\|z(k)\|$ is decreasing, so it follows that $\|z(k)\| \geq \|z_\infty\|$ for all $k \in \mathbb{N}$. This implies that for k sufficiently large it must hold that $S(k)z_\infty = z_\infty$. This, however, means that z_∞ lies in the eigenspace $V(S(k))$ for all k large enough. From (36) it follows that at least two sources do not see a drop for all k large enough.

■

For the statement of the next result we denote the set of column stochastic matrices of rank 1 by $\mathcal{R}$. Note that the matrices in $\mathcal{R}$ are of the form

$$\eta \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}\,,$$

where η is a nonnegative vector, the entries of which sum to 1. In particular, the matrices in $\mathcal{R}$ are idempotent, because $\begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}\eta = 1$.

Theorem 5.1 *Let $\{A(k)\}_{k\in\mathbb{N}} \subset \mathcal{A}$ be a sequence with associated index sets $\mathcal{I}(k)$. Then each of the statements of Proposition 5.1 is equivalent to*

$$\lim_{k\to\infty} \operatorname{dist}\left(A(k)A(k-1)\dots A(0), \mathcal{R}\right) = 0\,. \tag{37}$$

Proof. Consider Proposition 5.1(ii). It follows from Corollary 5.1 that the system (21) can be transformed to an equivalent system (25). This implies that for each t the product $A(k)A(k-1)\dots A(0)$ is similar to

$$T(k) := \begin{bmatrix} S(k)S(k-1)\dots S(0) & * \\ 0 & 1 \end{bmatrix}\,. \tag{38}$$

By Proposition 5.1 it follows that $S(k)S(k-1)\dots S(0) \to 0$. As the distance of $T(k)$ to a matrix of rank 1 is upper bounded by $\|S(k)S(k-1)\dots S(0)\|$ this implies that the distance of $A(k)A(k-1)\dots A(0)$ to the set of matrices of rank 1 converges to zero. As each of these matrices is column stochastic any limit point of the sequence $\{A(k)A(k-1)\dots A(0)\}$ is column stochastic.

Conversely, it is clear that the (37) implies Proposition 5.1 (i). This shows the assertion. ∎

The minor drawback of Proposition 5.1 is that no rate of convergence is supplied. Indeed, the reader may convince himself that the rate of convergence may be made arbitrarily slow by considering sequences that have repetitions of the same matrices for longer and longer intervals as $k \to \infty$. It is therefore useful to provide conditions that guarantee an exponential decay. One such condition is provided in the following proposition.

Proposition 5.2 *For every $a \in \{n, n+1, n+2, \dots\}$ there exists a constant $r_a < 1$ with the following property. For any sequence of index sets $\mathcal{I}(k)$ such that for all $l \in \mathbb{N}$ and all $i \in \{1, \dots, n\}$ there is a $b \in \{la, la+1, \dots, (l+1)a-1\}$ with $i \notin \mathcal{I}(s)$ it holds that*

$$\|S(k-1)\dots S(k')\| \leq r_k^{-2k+1} r_k^{(k-k')}, \quad \forall k \geq k' \geq 0 \tag{39}$$

$$\operatorname{dist}(A(k-1)\dots A(k'), \mathcal{R}) \leq \|T\|\|T^{-1}\| r_k^{-2k+1} r_k^{(k-k')}, \quad \forall k \geq k' \geq 0 \tag{40}$$

with T defined by (25)*.*

Proof. Fix $a \in \{n, n+1, n+2, \dots\}$ and consider a finite sequence $\mathcal{I}(0), \dots, \mathcal{I}(a-1)$ of index sets such that for all $i \in \{1, \dots, n\}$ there is an $0 \leq b \leq k-1$ with $i \notin \mathcal{I}(b)$. If we continue this sequence periodically then it clearly satisfies the assumption of Proposition 5.1, so that for the associated sequence $\{S(k)\}_{k\in\mathbb{N}}$ we have $\lim_{k\to\infty} S(k-1)S(k-2)\dots S(0) \to 0$. As the sequence of matrices is periodic this implies that $r(S(a-1)S(a-2)\dots S(0)) < 1$. Using that for all $\mathcal{I}$ we have $\|S_{\mathcal{I}}\| = 1$ and $\|S_{\mathcal{I}}x\| = \|x\|$ if and only if $S_{\mathcal{I}}x = x$ the inequality on the spectral radius implies that $\|S(k-1)S(k-2)\dots S(0)\| \leq \rho < 1$ for some constant ρ depending on the sequence.

By taking the maximum over all the (finitely many) sequences of length a satisfying the drop condition and denoting it by r_a, we obtain that

$$\|S(k-1)\dots S(0)\| \leq r_a^{k/a},$$

so that the first claim follows.

The second claim follows as the distance of $A(k)A(k-1)\dots A(0)$ to the set of matrices of rank 1 is upper bounded by $\|T^{-1}\begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix} T\|$. ∎

Note that any actual flow on a real network has to satisfy the assumption on the drops seen described in the previous proposition. The reason for this is that if a flow does not see a drop it will continue to increase the amount of packages sent by a constant rate. Eventually this leads to the case the the amount of packages sent exceeds the capacity of the pipe if no drops are seen. But at this point the source necessarily sees a drop. This very coarse argument shows that all realistic flows will satisfy the assumptions of the previous proposition for some k.

5.2 Stochastic properties of the set $\mathcal{A}$

We now proceed to give a number of results that relate to random products of matrices from the set $\mathcal{A}$. In this section we assume that Assumptions 1 and 2 hold.

We first note that under our assumptions that the expectation of $\mathcal{A}$ is a positive matrix that is column stochastic with Perron eigenvector $\bar{x}_p^T = (\frac{\alpha_1}{\lambda_1(1-\beta_1)}, \frac{\alpha_2}{\lambda_2(1-\beta_2)}, ..., \frac{\alpha_n}{\lambda_n(1-\beta_n)})$. We then proceed to show that the expectation of

$$\Pi(k) = A(k)A(k-1)\dots A(0),$$

is also a column stochastic matrix with the same Perron eigenvector. The second result in this section concerns the asymptotic behaviour of the expectation of $\Pi(k)$. These results immediately yield Theorem 3.1 and Theorem 3.2, using the transformation Γ, if necessary.

The final results in this section revisit the convergence of $\Pi(k)$ to the set of rank-1 idempotent matrices. We show that for all $\delta > 0$ the probability of $\Pi(k)$ being at least a distance δ from the rank-1 idempotent matrices goes to zero as k becomes large.

In the following we will use the notation $A_{\mathcal{I}} = \Lambda_{\mathcal{I}} + c_\alpha \alpha \beta(\mathcal{I})^T$, where Λ denotes the diagonal matrix, $c_\alpha := (\sum_{j=1}^n \alpha_j)^{-1}$ and $\beta(\mathcal{I})$ is the vector with entries $1 - \beta_i(\mathcal{I})$.

Lemma 5.3 *Assume that $\lambda_i > 0$ for $i = 1, \dots, n$ then the expectation*

$$E(A) = \sum_{\mathcal{I}} \rho_{\mathcal{I}} A_{\mathcal{I}}$$

is positive, column stochastic, and a Perron eigenvector for it is given by

$$\bar{x}_p^T = \left(\frac{\alpha_1}{\lambda_1(1-\beta_1)}, \frac{\alpha_2}{\lambda_2(1-\beta_2)}, ..., \frac{\alpha_n}{\lambda_n(1-\beta_n)} \right). \tag{41}$$

Proof. By definition of the expectation we have

$$\begin{aligned} E(A) = \sum_{\mathcal{I}} \rho_{\mathcal{I}} A_{\mathcal{I}} &= \sum_{\mathcal{I}} \rho_{\mathcal{I}} \Lambda_{\mathcal{I}} + c_\alpha \sum_{i=1}^{m} \rho_{\mathcal{I}} \alpha \beta(\mathcal{I})^T \\ &= E(\Lambda) + c_\alpha \alpha E(\beta)^T. \end{aligned} \tag{42}$$

The i'th diagonal entry of the diagonal matrix $E(\Lambda)$ is

$$E(\Lambda_{i,i}) = \lambda_i\beta_i + (1 - \lambda_i) \tag{43}$$

and the i'th entry of $E(\beta)$ is

$$E(\beta_i) = \lambda_i(1 - \beta_i). \tag{44}$$

Hence, the matrix $E(A)$ is of the form of A_1 defined in Equation (21) with the same vector α and β_i replaced by $\tilde{\beta}_i := 1-\lambda_i(1-\beta_i) \in (0, 1)$. It follows by Theorem 2.1 that a Perron eigenvector of $E(A)$ is given by $\bar{x}_p^T = (\frac{\alpha_1}{\lambda_1(1-\beta_1)}, \frac{\alpha_2}{\lambda_2(1-\beta_2)}, \ldots, \frac{\alpha_n}{\lambda_n(1-\beta_n)})$.
■

Lemma 5.4 *Consider the random system (19) subject to Assumptions 3.1 and 3.2. The expectation of* $\Pi(k)$ *is:*

$$E(\Pi(k)) = E(A)^k = (\sum_{\mathcal{I}} \rho_{\mathcal{I}} A_{\mathcal{I}})^k. \tag{45}$$

Proof. By independence we have that the expectation of the product is the product of the expectations. This implies the equality. ■

Proposition 5.3 *Consider the random system (19) subject to Assumptions 3.1 and 3.2. Then, with probability one,*

$$\lim_{k\to\infty} \text{dist}\,(A(k)A(k-1)\ldots A(0), \mathcal{R}) = 0\,.$$

.

Proof. Under the assumptions that the λ_j are positive and the independence assumptions, with probability one each source will see infinitely many drops. Now the result follows from Theorem 5.1. ■

For convenience we refine the above proposition as the following lemma, which is a consequence of Proposition 5.2.

Lemma 5.5 *For all* $\delta > 0$*, the expectation of* $\text{dist}(\Pi(k), \mathcal{R})$ *exponentially tends to zero as* $k \to \infty$*, i.e. there are constants* $0 \le \mu < 1$*,* $C \ge 1$ *such that*

$$E(\text{dist}(\Pi(k), \mathcal{R})) \le C\mu^k\,.$$

5.3 Proof of Theorem 3.2

We now proceed to present the main result of this paper, Theorem 3.2. This result forms part of the theory of products of random matrices and could be shown by introducing the necessary concepts from that theory as discussed in [30]. We prefer to give a proof that relies on fairly elementary arguments in order to keep the main ideas accessible.

To aid exposition we first give an outline of this proof.

Outline of proof: We are interested in the asymptotic behaviour of the average window variable $\overline{W}(k)$,

$$\begin{aligned}\overline{W}(k) &= \frac{1}{k}\sum_{i=0}^{k-1} W(i) \\ &= \frac{1}{k}\left(\Pi(k-1)+...+\Pi(0)\right)W(0) \\ &= \frac{1}{k}\left(\sum_{i=0}^{k-1}\Pi(k)\right)W(0),\end{aligned}$$

as k tends to infinity. Our proof consists of five main steps.

Step 1 : We first show that $\overline{W}(k)$ can be approximated as

$$\begin{aligned}\overline{W}(k) &= \frac{1}{k+1}(\Pi(k)+...+\Pi(0))W(0) \\ &= \frac{1}{k+1}(R(k)+\Delta(k)+\ldots+R(l)+\Delta(l)+\Pi(l-1)+...+\Pi(0))W(0) \\ &\approx \frac{1}{k+1}(R(k)+\ldots+R(l)+\Pi(l-1)+...+\Pi(0))W(0),\end{aligned}$$

where $R(k),...,R(l)$ are column stochastic rank-1 matrices and $\Delta(k),..,\Delta(l)$ are small.

Step 3 : It is then seen that $R(k)+...+R(l)$ can be approximated as $(\sum_{i=1}^{m}\rho_i A_i)^l$.

Step 4 : And it follows that

$$\lim_{k\to\infty}\overline{W}(k) = x_p\bar{y}_p^T W(0),$$

where $\bar{y}_p^T = (1,...,1)$.

We now present the proof in detail:

Proof. (of Theorem 3.2)

In the sequel it will be convenient to measure the distance in terms of the size of the block on the $n-1$-dimensional invariant subspace. To this end recall that each product $\Pi(k)$ is similar to a matrix of the form

$$\Pi(k) = T\begin{bmatrix} S(k) & B(k) \\ 0 & 1\end{bmatrix}T^{-1},$$

and using this representation we may define $\mathrm{dist}(\Pi(k), \mathcal{R}) := \|S(k)\|$. In particular, this implies that for all products $\Pi(k)$ and all $A \in \mathcal{A}$ we have

$$\mathrm{dist}(A\Pi(k), \mathcal{R}) \leq \mathrm{dist}(\Pi(k), \mathcal{R}) .$$

Consider now a fixed $k_0 \in \mathbb{N}$. For $k \geq k_0$ we factorize

$$\Pi(k) = \Phi(k)\Psi(k) ,$$

where $\Phi(k)$ is the leading part of length k_0, i.e.,

$$\Phi(k) := A(k)A(k-1)\dots A(k-k_0+1) , \text{ and } \Psi(k) := A(k-k_0)\dots A(0) .$$

Let $k \geq k_0$ and write $k = mk_0 + l$, with $0 \leq l < k_0$. We now rewrite $\overline{W}(k)$ as sums of products where the first k_0 factors in each sum do not overlap. Consider

$$\begin{aligned} \overline{W}(k) = \frac{1}{k+1}(\Pi(k) + ... + \Pi(0))W(0) = \\ \left(\frac{1}{k+1}\sum_{i=0}^{k_0-1}\sum_{j=2}^{m}\Pi(jk_0+l-i)\right)W(0) + \frac{1}{k+1}\sum_{j=0}^{k_0+l}\Pi(j)W(0) = \\ \left(\frac{1}{k+1}\sum_{i=0}^{k_0-1}\sum_{j=2}^{m}\Phi(jk_0+l-i)\Psi(jk_0+l-i)\right)W(0) + \frac{1}{k+1}\sum_{j=0}^{k_0+l}\Pi(j)W(0) . \end{aligned} \tag{46}$$

Now by assumption for fixed $0 \leq i \leq k_0 - 1$ the random variables $\Phi(jk_0 + l - i)$, $j = 1, \dots, m$ are independent and identically distributed. By Lemma 5.4 the expectation of these random variables is given by

$$E(\Pi(k_0)) = E(A)^{k_0} .$$

We now investigate the individual sums that are in the inside. As they are all of the same form, we consider the sum for $i = l$ to keep notation simple. We may write

$$\Phi(jk_0) = R(jk_0) + \Delta(jk_0) ,$$

where $R(jk_0)$ is the column stochastic, rank one matrix in closest to $\Phi(jk_0)$ and $\Delta(jk_0)$ is an error term. With this notation we have

$$\sum_{j=2}^{m}\Phi(jk_0)\Psi(jk_0) = \sum_{j=2}^{m}\left[R(jk_0) + \Delta(jk_0)\right]\Psi(jk_0) .$$

Now $R(jk_0)\Psi(jk_0) = R(jk_0)$ as $R(jk_0)$ is of the form $\eta\left[1,\ldots,1\right]$ and $\Psi(jk_0)$ is column stochastic. Thus we may continue

$$\sum_{j=2}^{m}\left[R(jk_0)+\Delta(jk_0)\right]\Psi(jk_0) = \sum_{j=2}^{m} R(jk_0) + \sum_{j=2}^{m}\Delta(jk_0)\Psi(jk_0) = \sum_{j=2}^{m}\Phi(jk_0) + \sum_{j=2}^{m}\Delta(jk_0)(\Psi(jk_0)-I)\,.$$

We note for further reference that

$$\sum_{j=2}^{m}\|\Delta(jk_0)(\Psi(jk_0)-I)\| \le \sum_{j=2}^{m}\|\Delta(jk_0)\|C\,,$$

with a constant C uniformly bounding $\|\Pi(k)-I\|$ over all matrix products from $\mathcal{A}$.

Denote $d(k_0) := E(\mathrm{dist}(\Pi(k_0),\mathcal{R}))$. By Lemma 5.5 we have that $d(k_0)\to 0$ as $k_0\to\infty$. Noting that $\|\Delta(jk_0)\| = \mathrm{dist}(\Phi(jk_0),\mathcal{R})$ by the strong law of large numbers we have

$$P\left(\frac{1}{m-1}\sum_{j=2}^{m}(\|\Delta(jk_0)\|-d(k_0))\ge\varepsilon\right) \le \frac{1}{(m-1)\varepsilon^2}\mathrm{Var}(\mathrm{dist}(\Pi(k_0),\mathcal{R})) \quad (47)$$

Applying the law of large numbers again, we also have that

$$P\left(\|\frac{1}{m-1}\sum_{j=2}^{m}\Phi(jk_0)-E(A)^{k_0}\|\ge\varepsilon\right) \le \frac{1}{(m-1)\varepsilon^2}\|\mathrm{cov}(\Pi(k_0))\|\,. \quad (48)$$

Consider now a fixed $\delta>0$ and choose k_0 large enough so that $d(k_0)C<\delta/4$ then we have for all m large enough that

$$\begin{aligned}
& P(\|\frac{1}{m-1}\sum_{j=2}^{m}\Pi(jk_0)-E(A)^{k_0}\|\ge\delta) \\
\le\ & P(\|\frac{1}{m-1}\sum_{j=2}^{m}\Phi(jk_0)-E(A)^{k_0}\| + \frac{1}{m}\sum_{j=2}^{m-1}\|\Delta(jk_0)(\Psi(jk_0)-I)\|\ge\delta) \\
\le\ & P(\|\frac{1}{m-1}\sum_{j=2}^{m}\Phi(jk_0)-E(A)^{k_0}\|\ge\frac{\delta}{2}) + P(\frac{1}{m-1}\sum_{j=2}^{m}\|\Delta(jk_0)\|C\ge\frac{\delta}{2}) \\
\le\ & P(\|\frac{1}{m-1}\sum_{j=2}^{m}\Phi(jk_0)-E(A)^{k_0}\|\ge\frac{\delta}{2}) + P(\frac{1}{m-1}\sum_{j=2}^{m}\|\Delta(jk_0)\|-d(k_0)\ge\frac{\delta}{4C}) \\
\le\ & \frac{4}{(m-1)\delta^2}\mathrm{Var}(\mathrm{dist}(\Pi(k_0),\mathcal{R})) + \frac{16C^2}{(m-1)\delta^2}\|\mathrm{cov}(\Pi(k_0))\|\,,
\end{aligned}$$

where we have used (47) and (48) in the last step. It follows that for all $\delta > 0$ and k_0 such that $d(k_0)C < \delta/4$ we have

$$\lim_{m\to\infty} P(\|\frac{1}{m-1}\sum_{j=2}^{m}\Pi(jk_0) - E(A)^{k_0}\| > \delta) \to 0\,. \tag{49}$$

Repeat these arguments in the same manner for the terms $\sum_{j=2}^{m}\Phi(jk_0+l-i)\Psi(jk_0+l-i)$, $i = 0,\dots,k_0-1$ in (46) and note that

$$\frac{1}{k+1}\sum_{i=0}^{k_0-1}\sum_{j=2}^{m}\Phi(jk_0+l-i)\Psi(jk_0+l-i) =$$
$$\frac{m-1}{k+1}\sum_{i=0}^{k_0-1}\frac{1}{m-1}\sum_{j=2}^{m}\Phi(jk_0+l-i)\Psi(jk_0+l-i)\,,$$

and $\lim_{m\to\infty}(m-1)/(k+1) = k_0$. As the remaining term in (46) converges to 0 as $k \to \infty$ we obtain from (46) and (49)

$$\lim_{k\to\infty} P(\|\frac{1}{k+1}\sum_{j=0}^{k}\Pi(j) - E(A)^{k_0}\| > k_0\delta) \to 0\,,$$

provided that $d(k_0)C < \delta/4$.

Now $E(A)^{k_0} \to x_p y_p^T$ and using Lemma 5.5 $k_0 d(k_0) \to 0$ as $k_0 \to \infty$. Thus with probability one we have

$$\lim_{k\to\infty}\frac{1}{k+1}\sum_{j=0}^{k}\Pi(k) = x_p y_p^T\,.$$

The claim now follows. ■

6 Conclusions

In this paper we have presented, and validated using the network simulator NS, a random matrix model that describes the behaviour of a network of n AIMD flows that compete for shared bandwidth via a bottleneck router employing drop-tail queuing. We have used this model to relate several important network properties to properties of sets of nonnegative matrices that arise in the study of such networks. We have also derived under simplifying assumptions a number of analytic results that characterise the asymptotic time-average and ensemble-average throughput of such networks.

The obtained results are interesting for a number of reasons. Firstly, they suggest that for long lived flows, the average throughput of each flow may be controlled in a precise manner by adjusting the $AIMD$ parameters of each flow (α_i, β_i), or by adjusting the drop probabilities λ_i (or by a combination of both). This observation suggests that the results of this paper may be used as the basis for algorithms that prioritise network traffic in a precise manner. A second interesting observation that arises as a result of the work is that the presence of random background web traffic appears to have the effect of rendering Assumptions 3.1 and 3.2 valid. Future work will investigate whether this conjecture can be validated or refuted statistically for a number of network topologies.

Finally, we note that while our model has been used to characterise the average behaviour of the source congestion windows, much analytical and experimental work remains to be done. From an analytic perspective, characterising the distributions of the congestion window vector, and characterising the rates of convergence to the expected values of the distributions is of great practical importance, as is extending the model to include the effect of slow-start behaviour and of multiple short lived flows. Finally, we note that while the fidelity of the model has been established using NS, we are currently in the process of validating the model and its predictions using measurements from real networks.

Acknowledgements

This work was supported by Science Foundation Ireland grant 00/PI.1/C067. This work was also partially supported by the Enterprise Ireland grant SC/2000/084/Y. Enterprise Ireland is not responsible for any use of data appearing in this publication.

References

[1] R. Shorten, D. Leith, J. Foy, and R. Kilduff, "Analysis and design of synchronised communication networks." *Submitted to* Automatica, 2003.

[2] A. Berman and R. Plemmons, *Nonnegative matrices in the mathematical sciences*. SIAM, 1979.

[3] R. Shorten, D. Leith, J. Foy, and R. Kilduff, "Analysis and design of synchronised communication networks," in *Proceedings of 12th Yale Workshop on Adaptive and Learning Systems*, 2003.

[4] A. Berman, R. Shorten, and D. Leith, "Positive matrices associated with synchronised communication networks." Submitted to *Linear Algebra and its Applications*, 2003.

[5] D. Hong and D. Lebedev, "Many TCP user asymptotic analysis of the AIMD model," Tech. Rep. INRIA Technical Report 4229, INRIA Rocquencourt, 2001.

[6] F. Bacelli and D. Hong, "Interaction of TCP flows as billiards," Tech. Rep. INRIA Technical Report 4437, INRIA Rocquencourt, 2002.

[7] P. Brown, "Resource sharing of TCP connections with different round trip times," in *Proceedings of IEEE INFOCOM*, (Tel Aviv, Israel), March 2000.

[8] E. Altman, T. Jimenez, and R. Nunez-Queija, "Analysis of two competing TCP/IP connections," *Perform. Evaluation*, vol. 49, no. 1-4, pp. 43–55, 2002.

[9] J. Hespanha, S. Hohacek, K. Obrarzka, and J. Lee, "Hybrid model of TCP congestion control," in *Hybrid Systems: Computation and Control*, pp. 291–304, 2001.

[10] L. Xu, K. Harfoush, and I. Rhee, "Binary increase congestion control for fast long-distance networks." To appear in Proceedings of IEEE INFOCOM, 2004.

[11] F. P. Kelly, "Mathematical modelling of the internet," in *Proceedings of ICIAM 99, 4th International Congress of Industrial Applied Mathematics*, (Edinburgh, UK), July 1999.

[12] S. Low, F. Paganini, and J. Doyle, "Internet congestion control," *IEEE Control Systems Magazine*, vol. 32, no. 1, pp. 28–43, 2002.

[13] S. Mascolo, "Congestion control in high speed communication networks using the Smith principle," *Automatica*, vol. 35, pp. 1921,1935, 1999.

[14] S. S. Kunniyur and R. Srikant, "Stable, Scalable, Fair Congestion Control and AQM schemes that achieve high utilisation in the internet," *IEEE Transactions on Automatic Control*, vol. 48, no. 11, pp. 2024–2029, 2003.

[15] L. Massoulie, "Stability of distributed congestion control with heterogeneous feedback delays," *IEEE Transactions on Automatic Control*, vol. 47, no. 6, pp. 895–902, 2002.

[16] C. Hollot, V. Misra, D. Towsley, and W. Gong, "Analysis and design of controllers for aqm routers supporting TCP flows," *IEEE Transactions on Automatic Control*, vol. 47, no. 6, pp. 945–959, 2002.

[17] G. Vinnicombe, "On the stability of networks operating TCP-like congestion control," tech. rep., Cambridge Univ. Statistical Laboratory Research Report, 2000-398., 2000.

[18] R. Johari and D. Tan, "End-to end congestion control for the internet: delays and stability," *IEEE/ACM Transactions on Networking*, vol. 9, no. 6, pp. 818–832, 2001.

[19] Y. Chait, C. V. Hollot, V. Misra, H. Han, and Y. Halevi, "Dyanamic analysis of congested TCP networks," in *Proceedings of American Control Conference*, 1999.

[20] C. V. Hollot and Y. Chait, "Non-linear stability analysis of a class of TCP/AQM networks," in *Proceedings of IEEE Conference on Decision and Control*, 2001.

[21] C. V. Hollot, V. Misra, D. Towsley, and W. Gong, "A control theoretic analysis of red," in *Proceedings of IEEE INFOCOM*, (Anchorage, AL, USA), April 2001.

[22] J. Hespahha, "Stochastic hybrid modeling of on-off TCP flows." Sumitted to *Hybrid Systems, Computation & Control*, 2004.

[23] F. Bacelli and D. Hong, "AIMD, Fairness and Fractal Scaling of TCP traffic," in *Proceedings of IEEE INFOCOM*, (New York, NY, USA), June 2002.

[24] S. Floyd and V. Jacobson, "Traffic phase effects in packet-switched gateways," *Journal of Internetworking: Practice and Experience*, vol. 3, no. 3, pp. 115–156, September, 1992.

[25] W. Willinger, M. S. Taqqu, R. Sherman, and D. V. Wilson, "Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level," *IEEE/ACM Transactions on Networking*, vol. 5, no. 1, pp. 71–86, 1997.

[26] D. Leith and R. Shorten, "Throughput analysis of TCP networks." To appear in the proceedings of Networking 2004, 2004.

[27] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[28] A. Vladimirov, L. Elsner, and W.-J. Beyn, "Stability and paracontractivity of discrete linear inclusions," *Linear Algebra Appl.*, vol. 312, no. 1-3, pp. 125–134, 2000.

[29] W.-J. Beyn and L. Elsner, "Infinite products and paracontracting matrices," *Electron. J. Linear Algebra*, vol. 2, pp. 1–8, 1997.

[30] P. Bougerol and J. Lacroix, *Products of random matrices with applications to Schrödinger operators*, vol. 8 of *Progress in Probability and Statistics.* Boston, MA: Birkhäuser Boston Inc., 1985.