On the higher moments of TCP

A. Schlote, F. Wirth, A. Berman, R. Shorten *

May 4, 2012

Abstract

In this paper we describe the moments of a stochastic model of the *Additive Increase Multiplicative Decrease* (AIMD) algorithm. AIMD is the algorithm that underpins the Transmission Control Protocol (TCP), which is used extensively in the internet. We prove that the Markov chain describing TCP has the remarkable property that all moments converge to their asymptotes at exactly the same rate. Further, we illustrate how a closed form solution can be obtained from the network properties, and this formula is explicitly calculated for the case of the third moment.

AMS classifications: 93A30, 15B51, 34Dxx, 39A50.
Keywords: TCP, Convergence rate of stochastic matrix model, Communication networks, Kronecker product.

1 Introduction

In studying communication networks that employ TCP (Transmission Control Protocol), one encounters network models that take the form of time-varying dynamic difference equations of the following form [1, 2, 3, 4, 5, 6, 7]:

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

where $W(k)$ is a real n-dimensional vector and where $A(k)$ is a matrix chosen randomly from the set of $m = 2^n - 1$ matrices: $\mathcal{A} = \{A_1, ..., A_m\}$. The nonnegative matrices $A_1, ..., A_m$ are defined as follows. Let $\alpha_1, .., \alpha_n$, $\beta_1, ..., \beta_n$ be positive real numbers that are smaller than 1 where the sum of the α's is equal to 1. Further, let

*A. Schlote (corresponding author, Tel.: +353 (0) 1 7086100; fax: +353 (0) 1 7086269) and R. Shorten are with the Hamilton Institute, NUI Maynooth, Ireland (e-mail: [Arieh.Schlote, Robert.Shorten]@nuim.ie). F. Wirth is with the Institut für Mathematik, Universität Würzburg , Würzburg, Germany (e-mail: wirth@mathematik.uni-wuerzburg.de). A. Berman is with the Department of Mathematics, Technion, Haifa, Israel (e-mail: berman@technion.ac.il).

$$A_1 = \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \beta_n \end{bmatrix} + \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_n \end{bmatrix} \begin{bmatrix} (1-\beta_1), \ \ldots, \ (1-\beta_n) \end{bmatrix}.$$

The matrices $A_2, .., A_m$ are constructed by taking the matrix A_1, and setting some, but not all, of the β_i to 1. We refer to A_1 as a TCP matrix and to the family of matrices $\{A_i, i = 1, \ldots, n\}$ as generated from a TCP matrix. In some real situations it is convenient to assume that the probability that $A(k) = A_i$ in (1), denoted p_i, is fixed and independent of k.

In the study of the system given by Equation (1) it is natural to consider the following convex combination of matrices:

$$M_r \;=\; \sum_{i=1}^{m} p_i A_i^{\otimes r}, \tag{2}$$

where $B^{\otimes r}$ denotes the Kronecker product $B \otimes B \otimes \cdots \otimes B$ of length r for any $B \in \mathbb{R}^{n \times n}$ and any $r \in \mathbb{N}$ and where $A \otimes B$ is defined as

$$A \otimes B = \begin{bmatrix} a_{11}B & \ldots & a_{1n}B \\ \vdots & & \vdots \\ a_{n1}B & \ldots & a_{nn}B \end{bmatrix}. \tag{3}$$

Equation (2) for $r = 1$ arises when studying the first moment of the stochastic process underlying communication networks employing the TCP algorithm, and Equation (2) for $r \geq 2$ arises when studying the higher moments of this process. The Perron eigenvectors of these matrices give the asymptotic values of the moments, and the second largest eigenvalues of M_r bounds the rate of convergence to the asymptotes. Previous work [5] considered the first and second moments of this process, and established that the second largest eigenvalues of the stochastic matrix corresponding to the first moment, and to the second moment, coincide. In this paper we prove the remarkable property that all moments converge at the same rate, and give an explicit formula for the asymptote of the third moments in two special cases.

We use the following notational conventions. $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{R}_+$ denote the sets of natural numbers, real numbers and positive real numbers respectively. For $n \in \mathbb{N}$ we denote by $\mathbb{R}^n$ the space of $n-$dimensional real vectors and by $\mathbb{R}^{n \times n}$ the space of $n \times n$ real matrices. Given $A \in \mathbb{R}^{n \times n}$, we denote its spectrum by $\sigma(A)$, and order its elements according to their absolute values: $|\lambda_1(A)| \geq |\lambda_2(A)| \geq \cdots \geq |\lambda_n(A)|$. We refer to $\lambda_2(A)$ as the second largest eigenvalue of A. This number is important because if A is a stochastic matrix describing the dynamics of a process, $\lambda_2(A)$ bounds the rate of convergence to the steady state.

This paper is structured as follows. In Section 2 we develop models for the stochastic moments of (1) and show that these models have a common bound on their convergence rates. In Section 3 we present a method in which the asymptotes for the third moments can be calculated. We also give explicit formulae in the case of three flows with distinct AIMD parameters, and in the case of a network with an arbitrary number of flows with common AIMD parameters.

2 Convergence rate of the moments

We now consider the second largest eigenvalue of the matrices given in Equation (2). The following theorem is proved in [5] and is central to our discussion.

Theorem 2.1. *Let $A_1, .., A_m$ be generated from a TCP matrix. Then there exists a non-singular matrix P such that $C_i = P^{-1} A_i P$ is of the form*

$$C_i = \left[\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline & & & \\ c_i & & S_i & \\ & & & \end{array} \right], \tag{4}$$

where all of the S_i are symmetric, positive definite and have spectral radius $\rho(S_i) \leq 1$ and the c_i are vectors of appropriate size.

In this paper we expand a result from [5, Theorem 2.2]. First recall the following properties of the Kronecker product that are proved in [8, Section 12.1 and 12.2] or follow directly from the definition.

Lemma 2.2. *Let $A, B, C, D \in \mathbb{R}^{n \times n}$, then*

1. *the following equality holds*

$$(A \otimes B)(C \otimes D) = AC \otimes BD;$$

2. *there exists a permutation matrix $T \in \mathbb{R}^{n^2 \times n^2}$ such that*

$$B \otimes A = T^\top (A \otimes B) T;$$

3. *if A and B are lower block triangular matrices then $A \otimes B$ is again lower block triangular;*

4. *$A \otimes (B + C) = A \otimes B + A \otimes C$;*

5. $(pA) \otimes B = A \otimes (pB)$ *for all* $p \in \mathbb{R}$;

6. *if* $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}$ *and* $\sigma(B) = \{\mu_1, \ldots, \mu_n\}$, *then* $\sigma(A \otimes B) = \{\lambda_i \mu_j, i, j = 1, \ldots, n\})$;

7. $(A \otimes B)^\top = A^\top \otimes B^\top$.

Now we can derive a model for the higher moments of the system given by Equation (1). We have

$$\begin{aligned} W(k+1) \otimes W(k+1) &= (A(k)W(k)) \otimes (A(k)W(k)) \\ &= (A(k) \otimes A(k))(W(k) \otimes W(k)). \end{aligned} \tag{5}$$

Along the lines of (5), for any $r \in \mathbb{N}$, we get the following relation

$$W(k+1)^{\otimes r} = A(k)^{\otimes r} W(k)^{\otimes r}. \tag{6}$$

The entries of $W(k)^{\otimes r}$ are of the form of the following product of length r: $w_i(k)w_j(k) \ldots w_l(k)$. The expectations of these products are the moments of $W(k)$. Taking expectations in Equation (6) with respect to the probabilities p_i of the events $A(k) = A_i$ for $i = 1, \ldots, n$ yields

$$E[W(k+1)^{\otimes r}] = E[A(k)^{\otimes r}] E[W(k)^{\otimes r}] \tag{7}$$

$$= M_r E[W(k)^{\otimes r}] \tag{8}$$

thus yielding a model describing the evolution of any moment of (1).

We now present the main result of this section.

Theorem 2.3. *Let* $\mathcal{A} = \{A_1, .., A_m\}$ *be generated from a TCP matrix. Then the following assertions are true.*

1. $\sigma(M_1) \subset \sigma(M_r)$ *for all* $r \in \mathbb{N}$.

2. *Let* $1 \neq \lambda \in \sigma(M_1)$ *with multiplicity* k *then for all* $r \in \mathbb{N}$ *we have* $\lambda \in \sigma(M_r)$ *with multiplicity of at least* $r \cdot k$.

3. *The second largest eigenvalue* $\lambda_2(M_r)$ *is the same for all* $r \in \mathbb{N}$.

Proof. The eigenvalues of a block triangular matrix depend only on the diagonal blocks. Also the matrix product and the Kronecker product of lower triangular matrices generate lower triangular matrices, see Lemma 2.2(3). So in dealing with this kind of matrices we can focus on the diagonal blocks. We split the proof into three parts.

(i) We show that M_r is similar to $\sum_{i=1}^m p_i C_i^{\otimes r}$. This follows from applying first Lemma 2.2(4),(5) and afterwards repeatedly Lemma 2.2(1) to

$(P^{-1})^{\otimes r} M_r P^{\otimes r}$, where P is the non-singular transformation matrix from Theorem 2.1:

$$(P^{-1})^{\otimes r} M_r P^{\otimes r} = \left(P^{-1}\right)^{\otimes r} (\sum_{i=1}^{m} p_i A_i^{\otimes r}) P^{\otimes r} = \sum_{i=1}^{m} p_i (P^{-1})^{\otimes r} A_i^{\otimes r} P^{\otimes r}$$
$$= \sum_{i=1}^{m} p_i \left((P^{-1}) A_i P\right)^{\otimes r} = \sum_{i=1}^{m} p_i C_i^{\otimes r}. \quad (9)$$

(ii) We show that $C_i^{\otimes r}$ is similar to a block triangular matrix with diagonal blocks $1, S_i, S_i^{\otimes 2}, \ldots, S_i^{\otimes r}$. Note that these blocks are not necessarily in this order and each $S_i^{\otimes q}$ appears exactly $\binom{r}{q}$ times and one block equals 1. We show this using induction. The claim is true for $r = 1$. Let it be true for $r - 1 \geq 0$. This means there is a non singular $T_{r-1} \in \mathbb{R}^{n^{r-1} \times n^{r-1}}$ where $T_{r-1}^{-1} C_i^{\otimes r-1} T_{r-1}$ is a block triangular matrix with diagonal blocks $1, S_i, S_i^{\otimes 2}, \ldots, S_i^{\otimes r-1}$. Then there is a non-singular matrix $Q_1 := (T_{r-1} \otimes I)^{-1}$, where $I \in \mathbb{R}^{n \times n}$ is the identity matrix, such that

$$C_i^{\otimes r} = Q_1^{-1} (T_{r-1} \otimes I)^{-1} (C_i^{\otimes r-1} \otimes C_i)(T_{r-1} \otimes I) Q_1$$
$$= Q_1^{-1} (T_{r-1}^{-1} C_i^{\otimes r-1} T_{r-1} \otimes I C_i I) Q_1, \quad (10)$$

where we used again Lemma 2.2(1). We then have

$$Q_1^{-1} (T_{r-1}^{-1} C_i^{\otimes r-1} T_{r-1} \otimes I C_i I) Q_1 \quad (11)$$
$$= Q_1^{-1} \left(\begin{bmatrix} 1 & 0 & \ldots & 0 \\ * & \ddots & \ddots & \vdots \\ \vdots & \ddots & S_i^{\otimes q} & 0 \\ * & \ldots & * & \ddots \end{bmatrix} \otimes C_i \right) Q_1 \quad (12)$$
$$= Q_1^{-1} \begin{bmatrix} C_i & & & 0 \\ & \ddots & & \\ & & S_i^{\otimes q} \otimes C_i & \\ * & & & \ddots \end{bmatrix} Q_1 \quad (13)$$
$$(14)$$

Thus $C_i^{\otimes r}$ is similar to a block triangular matrix with diagonal blocks $1 \otimes C_i, S_i \otimes C_i, S_i^{\otimes 2} \otimes C_i, \ldots, S_i^{\otimes r-1} \otimes C_i$, where the blocks again do not have to appear in this specific order. According to Lemma 2.2(2), for each block $S_i^{\otimes q} \otimes C_i$ with $q = 1, \ldots, r - 1$ exists a permutation matrix T_q of appropriate size, such that

$$C_i \otimes S_i^{\otimes q} = T_q^{-1} \left(S_i^{\otimes q} \otimes C_i\right) T_q. \quad (15)$$

Further, we can find a permutation matrix $Q_2 \in \mathbb{R}^{n^r \times n^r}$, which is a block diagonal matrix with blocks T_q, $q = 1, \dots, r-1$, such that

$$Q_1^{-1} Q_2^\top (T_{r-1} \otimes I)^{-1} (C_i^{\otimes r-1} \otimes C_i)(T_{r-1} \otimes I) Q_2 Q_1$$

is a block triangular matrix with diagonal blocks

$$C_i \otimes 1, C_i \otimes S_i, C_i \otimes S_i^{\otimes 2}, \dots, C_i \otimes S_i^{\otimes r-1}.$$

Using $C_i = \begin{bmatrix} 1 & 0 \\ * & S_i \end{bmatrix}$ the claim follows.

We conclude that M_r is similar to a block triangular matrix with blocks $1, \sum_{i=1}^m p_i S_i, \dots, \sum_{i=1}^m p_i S_i^{\otimes r}$, where each block $\sum_{i=1}^m p_i S_i^{\otimes r}$ appears exactly $\binom{r}{q}$ times for all $q = 1, \dots, r$ and one block equals 1. The theorem's first two claims follow immediately.

(iii) To prove (3) we show that the largest eigenvalue of $\sum_{i=1}^m p_i S_i^{\otimes r}$ is less than or equal to the largest eigenvalue of $\sum_{i=1}^m p_i S_i^{\otimes(r-1)}$ for all $r \in \mathbb{N}$. To this end let $I \in \mathbb{R}^{n-1 \times n-1}$ be the identity matrix and by Lemma 2.2(6) $\sigma\left(\sum_{i=1}^m p_i S_i^{\otimes(r-1)}\right) \subset \sigma\left(I \otimes \left(\sum_{i=1}^m p_i S_i^{\otimes(r-1)}\right)\right) = \sigma\left(\sum_{i=1}^m p_i I \otimes S_i^{\otimes(r-1)}\right)$, and no new eigenvalues arise by taking the Kronecker product with the identity. According to Theorem 2.1, for every vector $z \in \mathbb{R}^{(n-1)^r}$ we have that

$$\begin{aligned} & z^\top \left(\sum_{i=1}^m p_i I \otimes S_i^{\otimes(r-1)} - \sum_{i=1}^m p_i S_i^{\otimes r}\right) z \\ =& z^\top \left(\sum_{i=1}^m p_i (I - S_i) \otimes S_i^{\otimes(r-1)}\right) z \geq 0, \end{aligned} \tag{16}$$

since $S_i^{\otimes(r-1)}$ is positive definite and $I - S_i$ is positive semi-definite for all $i = 1, \dots, m$. As the S_i are real symmetric matrices, by Lemma 2.2(7) so are $\left(\sum_{i=1}^m p_i I \otimes S_i^{\otimes(r-1)}\right)$ and $\left(\sum_{i=1}^m p_i S_i^{\otimes r}\right)$. Now we denote by μ the largest eigenvalue of $\left(\sum_{i=1}^m p_i I \otimes S_i^{\otimes(r-1)}\right)$ and by ν the largest eigenvalue of $\left(\sum_{i=1}^m p_i S_i \otimes S_i^{\otimes(r-1)}\right)$. By the Rayleigh-Ritz Theorem [9, Theorem 4.2.2], we have

$$\mu = \max_{\|z\|=1} z^\top \left(\sum_{i=1}^m p_i I \otimes S_i^{\otimes(r-1)}\right) z, \tag{17}$$

$$\nu = \max_{\|z\|=1} z^\top \left(\sum_{i=1}^m p_i S_i \otimes S_i^{\otimes(r-1)}\right) z \tag{18}$$

and $\mu \geq \nu$ by Equation (16). This concludes the proof.

□

In the following we are interested only in situations where the moments of our model actually converge. This is assured by Theorem 2.3 if we make the additional assumption that the second largest eigenvalue of M_1 is strictly less than 1. It is straightforward to see that the second largest eigenvalue of M_1 can be equal to 1 if and only if there is at least one entry of $W(k)$ that is a strictly increasing function in k. In the context of communication networks this is impossible, as it would correspond to the existence of a TCP flow that never backs off. Under this additional assumption Theorem 2.3 not only assures that all moments converge to unique asymptotes, but furthermore it gives a uniform bound for the rate of convergence of all moments. In the next section we describe a method in which all higher moments of the system given by Equation (1) can be calculated directly from the network parameters.

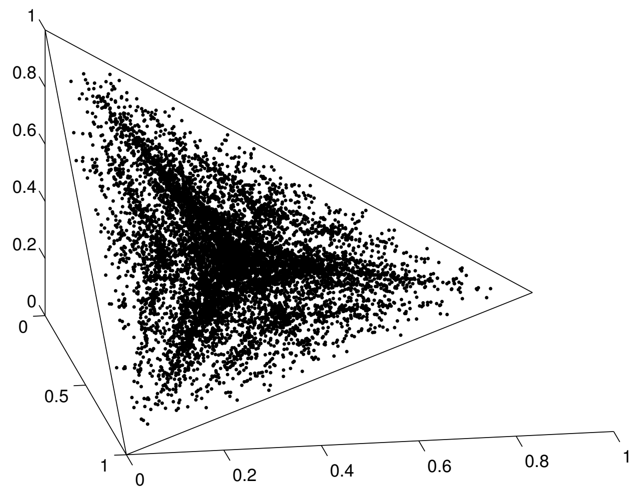
3 Formulae for the Perron vector of M_r

In the previous section we have shown that the moments of the TCP model converge to their asymptotes with a uniform bound on their rate of convergence. The asymptote to which they converge is determined by the Perron eigenvector of the M_r. These eigenvectors uniquely determine the moments of the process and are of practical interest. Furthermore, they are difficult to obtain experimentally, requiring a large number of simulations, and the probability density from which they are derived is hard to analyse.

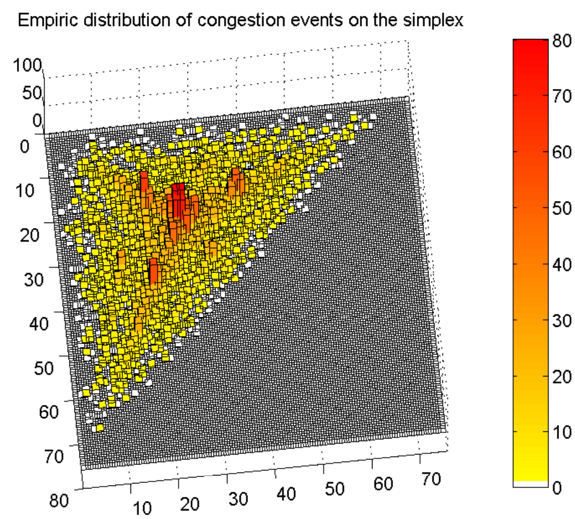
Examples are depicted in Figures 1 and 2, where the fractal-like characteristics of the density can be observed. For the simulations we used the following parameters: $\alpha_1 = \alpha_2 = \alpha_3 = \frac{1}{3}, \beta_1 = \beta_2 = \beta_3 = \frac{1}{4}, W(0) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^\top$. For the experiment depicted in Figure 1 all possible congestion events have the same probability, for the experiment depicted in Figure 2 only those congestion events were allowed, where two or more flows react to the congestion; these occur with equal probabilities. Congestion events are time instants at which the competing users (w_i) share the full available resource between them and one or more users have to perform a multiplicative decrease. Each congestion event corresponds to a time step of (1).

In view of these experiments, it is not only remarkable that information about the moments can be captured in the Perron eigenvectors of M_r, but also that closed form expressions for these vectors can be obtained. In this section we give a procedure for calculating the Perron vector of M_r. Details are given for M_3 but the same procedure follows for any M_r. This approach follows and expands upon the ideas given on an *ad-hoc* basis in [4].

Specifically, we now present a method in which the third moments of (1) can

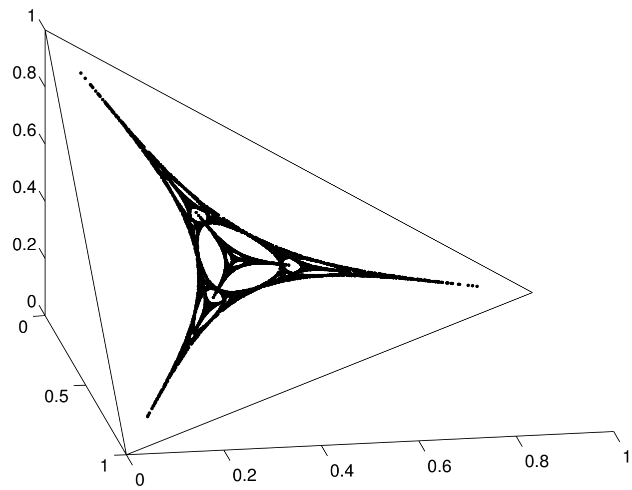


(a)

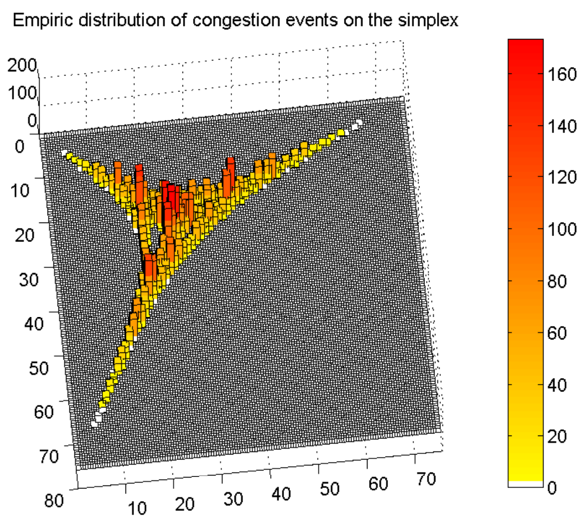


(b)

Figure 1: One realisation of a simulation of 10000 congestion events in a three flow network. (a) comes from a simulation of three flows competing for bandwidth in a single bottleneck network, where all possible congestion events have the same probability and each axis corresponds to the window size of one flow; (b) shows a corresponding histogram, i.e the number of congestion events in each segment of a partition of the simplex from (a).



(a)



(b)

Figure 2: A second realisation of a simulation of 10000 congestion events in a three flow network. (a) comes from a simulation of three flows competing for bandwidth in a single bottleneck network, where congestion events have distinct probabilities and each axis corresponds to the window size of one flow; (b) shows a corresponding histogram, i.e the number of congestion events in each segment of a partition of the simplex from (a).

be expressed in terms of the second and first moments and the network parameters. As we have mentioned, this enables us to abstain from doing extensive simulations. We now proceed to obtain an expression for this moment. We assume that the first and second moments are known [4], and we will now obtain the third moment from these vectors and the network parameters. Throughout this section we are omitting the dependence of variables on the time k whenever there are no ambiguities. We are interested in finding a method to compute the third moments of (1). If we assume (without any loss of generality) that $\sum_{i=1}^{n} w_i = 1$, then using

$$E[w_i] = E[w_i \cdot 1] = E[w_i \cdot (w_1 + w_2 + \cdots + w_n)] \tag{19}$$

and the linearity of the expectation, we obtain the following three equations

$$E[w_i^2] = E[w_i^3] + \sum_{j \neq i} E[w_i^2 w_j] \tag{20}$$

$$E[w_i] = E[w_i^3] + \sum_{j \neq i} E[w_i w_j^2] + 2 \sum_{j \neq i} E[w_i^2 w_j] + \sum_{j,l=1; j,l \neq i; j \neq l}^{n} E[w_i w_j w_l] \tag{21}$$

$$E[w_i w_j] = E[w_i^2 w_j] + E[w_i w_j^2] + \sum_{l \neq i,j} E[w_i w_j w_l]. \tag{22}$$

We obtain another set of equations relating the first three moments of (1) in the following manner. For a single flow $i \in \{1, \ldots, n\}$ we can describe its evolution by

$$w_i(k+1) = \beta_i w_i(k) + \alpha_i T, \tag{23}$$

where T is the time between the k'th and the $(k+1)'$th congestion event. This formula is the fundamental component of TCP (or AIMD). It describes a linear increase in w_i for a time interval of length T, potentially followed by an abrupt reset upon detection of congestion; see [3] for full details of this algorithm. Cubing Equation (23) yields

$$w_i^3(k+1) = \beta_i^3 w_i^3(k) + 3\alpha_i \beta_i^2 w_i^2(k) T + 3\alpha_i^2 \beta_i w_i(k) T^2 + \alpha_i^3 T^3 \tag{24}$$

and taking the expectation yields

$$E[w_i^3(k+1)] = E[\beta_i^3 w_i^3(k) + 3\alpha_i \beta_i^2 w_i^2(k) T + 3\alpha_i^2 \beta_i w_i(k) T^2 + \alpha_i^3 T^3]. \tag{25}$$

Using the linearity of the expectation, and the fact that α_i is a constant, and that probabilities are independent of $w(k)$ for all $k \in \mathbb{N}$ we obtain

$$\begin{aligned} &E[w_i^3(k+1)] = \\ &E[\beta_i^3] E[w_i^3(k)] + 3\alpha_i E[\beta_i^2 w_i^2(k) T] + 3\alpha_i^2 E[\beta_i w_i(k) T^2] + \alpha_i^3 E[T^3]. \end{aligned} \tag{26}$$

Next, we utilise the knowledge that $E[w_i^3(k)]$ converges as $k \to \infty$. Note this follows from Section 1 as the second eigenvalue of M_r is strictly less than unity for all r. Thus, for large $k_0 \in \mathbb{N}$ we can approximate $E[w_i^3(k+1)] = E[w_i^3(k)]$ for all $k \geq k_0$ and thus we obtain

$$\frac{E[1-\beta_i^3]}{\alpha_i^3} E[w_i^3(k)] = 3\frac{E[\beta_i^2 w_i^2(k) T]}{\alpha_i^2} + 3\frac{E[\beta_i w_i(k) T^2]}{\alpha_i} + E[T^3]. \tag{27}$$

For convenience, we now use the following notation in the final part of the proof:

$$p_{ijl} = \frac{E[1-\beta_i\beta_j\beta_l]}{\alpha_i\alpha_j\alpha_l}. \tag{28}$$

Equation (27) is a special case of the following equation (which is obtained in the same manner)

$$\begin{aligned} p_{ijl}E[w_iw_jw_l] =& \frac{E[\beta_i\beta_j w_i w_j T]}{\alpha_i\alpha_j} + \frac{E[\beta_j\beta_l w_j w_l T]}{\alpha_j\alpha_l} + \frac{E[\beta_i\beta_l w_i w_l T]}{\alpha_i\alpha_l} + \\ & \frac{E[\beta_i w_i T^2]}{\alpha_i} + \frac{E[\beta_j w_j T^2]}{\alpha_j} + \frac{E[\beta_l w_l T^2]}{\alpha_l} + E[T^3] \end{aligned} \tag{29}$$

These equations give rise to the following two equations by eliminating all terms containing a T. For the first equation we need the existence of three distinct flows, and for the second equation two flows suffice.

$$\begin{aligned} & 2\left(p_{iii}E[w_i^3] + p_{jjj}E[w_j^3] + p_{lll}E[w_l^3]\right) + 12p_{ijl}E[w_iw_jw_l] \\ =& 3\left(p_{iij}E[w_i^2w_j] + p_{jji}E[w_j^2w_i] + p_{jjl}E[w_j^2w_l] \right. \\ & \left. + p_{llj}E[w_l^2w_j] + p_{lli}E[w_l^2w_i] + p_{iil}E[w_i^2w_l]\right) \end{aligned} \tag{30}$$

$$3p_{iij}E[w_i^2w_j] - 3p_{jji}E[w_j^2w_i] = p_{iii}E[w_i^3] - p_{jjj}E[w_j^3] \tag{31}$$

The next proposition gives an explicit formula for the third moments in the case for an arbitrary number of flows, where all flows share the same AIMD parameters.

Proposition 3.1. *In the case of n flows, where all flows have the same network parameters, that is $\alpha_1 = \cdots = \alpha_n \in \mathbb{R}_+$ and $\beta_1 = \cdots = \beta_n \in (0,1)$ the third moments are given by:* $E[w_i^3] =$

$$\frac{\left((n-2)3E[1-\beta_1^2\beta_2] + 6E[1-\beta_1\beta_2\beta_3]\right)E[w_1^2] - 2E[1-\beta_1\beta_2\beta_3]E[w_1]}{(n-2)3E[1-\beta_1^2\beta_2] + 4E[1-\beta_1\beta_2\beta_3] + (n-1)(n-2)E[1-\beta_1^3]},$$

for all $i = 1, \ldots, n$, where we have $E[w_1] = \cdots = E[w_n] = \frac{1}{n}$.

Proof. Equations (20),(21) and (30) reduce to:

$$E[w_i^3] = E[w_1] - 3(n-1)E[w_1 w_2^2] - (n-1)(n-2)E[w_1 w_2 w_3] \quad (32)$$

$$E[w_i^3] = E[w_1^2] - (n-1)E[w_1 w_2^2] \quad (33)$$

$$E[1-\beta_1^3]E[w_1^3] = 3E[1-\beta_1\beta_2^2]E[w_1 w_2^2] - 2E[1-\beta_1\beta_2\beta_3]E[w_1 w_2 w_3] \quad (34)$$

and eliminating the expectations of the mixed products proves the assertion. □

In order to compute the third moments in the case of distinct parameters we write some of the above equations in a matrix notation. To this end let $X := [E[w_i^3]] \in \mathbb{R}^n, Y := [E[w_i^2 w_j]]_{j\neq i} \in \mathbb{R}^{n(n-1)}, Z := [E[w_i^2]] \in \mathbb{R}^n$ and $V := [E[w_i w_j]]_{j\neq i} \in \mathbb{R}^{\frac{n(n-1)}{2}}$, where entries are to be ordered lexicographically according to their indices. In matrix notation we can write (20) as

$$X + U_1 Y = Z, \quad (35)$$

where if we denote by I_n the identity matrix of dimension n and by s_{n-1} the all ones row vector of length $n-1$, we then have $U_1 = I_n \otimes s_{n-1} \in \mathbb{R}^{n\times n(n-1)}$. Further, Equation (31) can be written as

$$CX + U_2 Y = 0, \quad (36)$$

where $C \in \mathbb{R}^{\frac{n(n-1)}{2}\times n}$ and $U_2 \in \mathbb{R}^{\frac{n(n-1)}{2}\times n(n-1)}$. We would like to note at this point that there is more than one possibility of writing Equation (36) in the described matrix form. In particular rows of C and U_2 can be swapped or multiplied by -1 as long as it is done simultaneously. Substituting Equation (30) in Equation (22) yields

$$RX + U_3 Y = V, \quad (37)$$

where $R \in \mathbb{R}^{\frac{n(n-1)}{2}\times n}$ and $U_3 \in \mathbb{R}^{\frac{n(n-1)}{2}\times n(n-1)}$. To calculate the third moments, we can now use (35) to substitute for X in (36) and (37) and calculate Y with the following two resulting equations:

$$R(Z - U_1 Y) + U_3 Y = V \quad (38)$$

and

$$C(Z - U_1 Y) + U_2 Y = 0. \quad (39)$$

Equivalently

$$(RU_1 - U_3)Y = RZ - V \quad (40)$$

and

$$(CU_1 - U_2)Y = CZ. \quad (41)$$

Together we have

$$\begin{bmatrix} (RU_1 - U_3) \\ (CU_1 - U_2) \end{bmatrix} Y = \begin{bmatrix} RZ - V \\ CZ \end{bmatrix}. \quad (42)$$

$\begin{bmatrix} (RU_1 - U_3) \\ (CU_1 - U_2) \end{bmatrix}$ is a square matrix of order $n^2 - n$. In all practical applications this matrix is invertible or can be made invertible by a slight change of some of the parameters. Assuming we can use the inverse of $\begin{bmatrix} (RU_1 - U_3) \\ (CU_1 - U_2) \end{bmatrix}$ we obtain

$$Y = \begin{bmatrix} (RU_1 - U_3) \\ (CU_1 - U_2) \end{bmatrix}^{-1} \begin{bmatrix} RZ - V \\ CZ \end{bmatrix}. \tag{43}$$

And together with (35) we obtain

$$X = Z - U_1 \begin{bmatrix} (RU_1 - U_3) \\ (CU_1 - U_2) \end{bmatrix}^{-1} \begin{bmatrix} RZ - V \\ CZ \end{bmatrix}. \tag{44}$$

For the special case of three flows, invertibility of the matrix $\begin{bmatrix} (RU_1 - U_3) \\ (CU_1 - U_2) \end{bmatrix}$ can be shown directly. This result is summarised in the following lemma.

Lemma 3.2. *For $n = 3$ the matrix $\begin{bmatrix} (RU_1 - U_3) \\ (CU_1 - U_2) \end{bmatrix}$ in Equation (42) is invertible.*

Proof. We use the following notation:

$$a_1 := \frac{-2p_{111} - 3p_{112}}{12p_{123}} \qquad a_2 := \frac{-2p_{111} - 3p_{113}}{12p_{123}} \tag{45}$$

$$a_3 := \frac{-2p_{222} - 3p_{221}}{12p_{123}} \qquad a_4 := \frac{-2p_{222} - 3p_{223}}{12p_{123}} \tag{46}$$

$$a_5 := \frac{-2p_{333} - 3p_{331}}{12p_{123}} \qquad a_6 := \frac{-2p_{333} - 3p_{332}}{12p_{123}}. \tag{47}$$

For $n = 3$ we can write $\begin{bmatrix} (RU_1 - U_3) \\ (CU_1 - U_2) \end{bmatrix}$ as a rank-1 update of the form $A + B$ with

$$A = \tag{48}$$

$$\begin{bmatrix} -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 & -1 \\ p_{111} + 3p_{112} & p_{111} & -p_{222} - 3p_{221} & -p_{222} & 0 & 0 \\ 0 & 0 & p_{222} & p_{222} + 3p_{223} & -p_{333} & -p_{333} - 3p_{332} \\ -p_{111} & -p_{111} - 3p_{113} & 0 & 0 & p_{333} + 3p_{331} & p_{333} \end{bmatrix}$$

and

$$B = vw^\top, \tag{49}$$

with $v = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}^\top$ and $w = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \end{bmatrix}^\top$.
If A is invertible and further $1 + w^\top A^{-1} v \neq 0$, then $A + vw^\top$ is invertible and

by the Sherman-Morrison Formula, e.g [9, p. 19],

$$(A + vw^\top)^{-1} = A^{-1} - \frac{1}{1 + w^\top A^{-1} v} A^{-1} v w^\top A^{-1}. \tag{50}$$

First we show that A is invertible. The determinant of A is equal to

$$\begin{aligned} \det A = & -3\,(p_{111}p_{222} + p_{111}p_{333} + p_{222}p_{333})\,(p_{112} + p_{113} + p_{221} + p_{223} + p_{331} + p_{332}) \\ & - 9p_{111}\,(p_{223} + p_{332})\,(p_{112} + p_{113} + p_{221} + p_{331}) \\ & - 9p_{222}\,(p_{113} + p_{331})\,(p_{112} + p_{221} + p_{223} + p_{332}) \\ & - 9p_{333}\,(p_{112} + p_{221})\,(p_{113} + p_{223} + p_{331} + p_{332}) \\ & - 27\,(p_{113} + p_{331})\,(p_{112} + p_{221})\,(p_{223} + p_{332})\,. \end{aligned} \tag{51}$$

As $p_{ijl} > 0$ for all $i, j, l = 1, 2, 3$ as can be seen directly from the definition in Equation (28), it follows that $\det A < 0$ and thus A is invertible. Using symbolic manipulation one obtains that $w^\top A^{-1} v$ is a ratio of sums of positive terms. Thus $1 + w^\top A^{-1} v \neq 0$ for any set of parameters. □

In the general case the third moments can be calculated using Equation (43) and Equation (44). In the following proposition, that follows directly from the above discussion, we present an explicit formula for the third moments in the special case of three flows with distinct AIMD parameters.

Proposition 3.3. *For 3 flows with distinct AIMD parameters $\alpha_1, \alpha_2, \alpha_3$ and $\beta_1, \beta_2, \beta_3$, the third moments are obtained from the second moments through the following formulae*

$$\begin{bmatrix} E[w_1^2 w_2] \\ E[w_1^2 w_3] \\ E[w_2^2 w_1] \\ E[w_2^2 w_3] \\ E[w_3^2 w_1] \\ E[w_3^2 w_2] \end{bmatrix} = (A{+}B)^{-1} \begin{bmatrix} -\frac{p_{111}E[w_1^2] + p_{222}E[w_2^2] + p_{333}E[w_3^2]}{6p_{123}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} E[w_1 w_2] \\ E[w_1 w_3] \\ E[w_2 w_3] \end{bmatrix} \\ \begin{bmatrix} p_{111}E[w_1^2] - p_{222}[w_2^2] \\ p_{222}[w_2^2] - p_{333}[w_3^2] \\ -p_{111}[w_1^2] + p_{333}[w_3^2] \end{bmatrix} \end{bmatrix}$$

and

$$\begin{bmatrix} E[w_1^3] \\ E[w_2^3] \\ E[w_3^3] \end{bmatrix} = \begin{bmatrix} E[w_1^2] \\ E[w_2^2] \\ E[w_3^2] \end{bmatrix} - \begin{bmatrix} E[w_1^2 w_2] + E[w_1^2 w_3] \\ E[w_2^2 w_1] + E[w_2^2 w_3] \\ E[w_3^2 w_1] + E[w_3^2 w_2] \end{bmatrix},$$

where A and B are given in Equations (48) *and* (49) *respectively.*

Example : To conclude our paper we now demonstrate the efficacy of our formulae by comparing their predictions against numerical simulations. For this purpose measurements were obtained from ensembles of Matlab simulations, in which three flows compete against each other for a resource using AIMD, and where the probabilities p_i are fixed *a-priori*. In particular we used $\alpha_1 =$

$\frac{1}{6}, \alpha_2 = \frac{2}{6}, \alpha_3 = \frac{3}{6}, \beta_1 = \beta_2 = \beta_3 = \frac{1}{4}, W(0) = (\frac{2}{3}, \frac{1}{3}, 0)^\top$ and the available resource is normed to 1. At each congestion event each subset of flows has the same probability of performing a multiplicative decrease reaction. From the measurements obtained in these simulations, the first, second, and third moments were derived. The evolution of these moments toward their asymptotes is depicted in Figures 3,4 and 5. The theoretical predictions are depicted in the figures as a dashed line, where the predictions in Figures 3 and 4 are based on results of [4]. The predictions in Figure 5, are based on the formulae derived in this present paper. Figure 5 illustrates the fidelity of the results in this paper. Note that all moments converge to a unique asymptote. This is consistent with the second eigenvalue of the M_r matrices being strictly less than unity. Second, note that all trajectories converge to their asymptote within 10 time-steps. This observation is consistent with the fact that the second eigenvalue is identical for all the M_r. Finally, note that the theoretical asymptotes and the measured values are in close agreement.

References

[1] F. Baccelli and D. Hong, "AIMD, fairness and fractal scaling of TCP traffic," in *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 1, pp. 229–238, 2002.

[2] F. Baccelli and D. Hong, "Interaction of TCP flows as billiards," *IEEE/ACM Transactions on Networking*, vol. 13, no. 4, pp. 841–853, 2005.

[3] R. Shorten, F. Wirth, and D. Leith, "A positive systems model of TCP-like congestion control: Asymptotic results," *IEEE/ACM Transactions on Networking*, vol. 14, pp. 616–629, 2006.

[4] R. Shorten, C. Kellett, and D. Leith, "On the dynamics of TCP's higher moments," *IEEE Communications Letters*, vol. 11, no. 2, pp. 210–212, 2007.

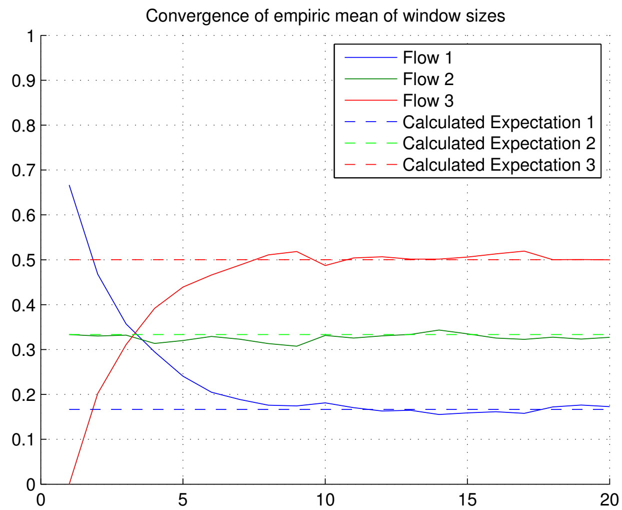
[5] A. Berman, T. Laffey, A. Leizarowitz, and R. Shorten, "On the second eigenvalue of matrices associated with TCP," *Linear Algebra Appl.*, vol. 416, no. 1, pp. 175–183, 2006.

[6] S. Liu, T. Başar, and R. Srikant, "TCP-Illinois: A loss- and delay-based congestion control algorithm for high-speed networks," *Perform. Eval.*, vol. 65, no. 6-7, pp. 417–440, 2008.

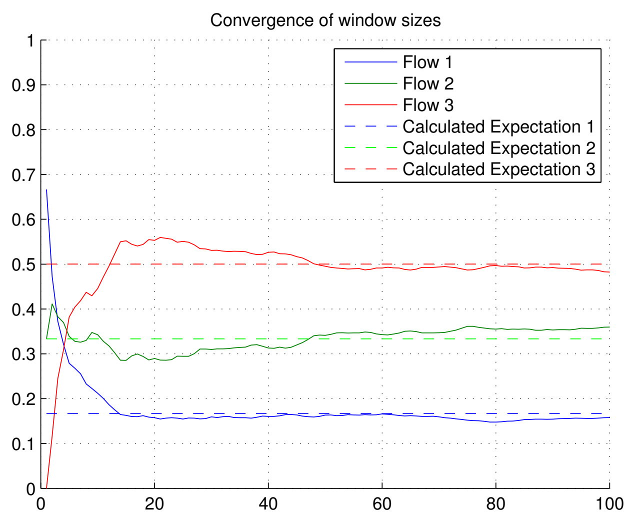
[7] R. D. Halper, E. J. Harder, B. R. Hunt, and J. A. Yorke, "Stability of TCP dynamics in large data networks.," *SIAM J. Applied Dynamical Systems*, vol. 8, no. 1, pp. 146–159, 2009.

[8] P. Lancaster and M. Tismenetsky, *The theory of matrices. 2nd ed., with applications.* Computer Science and Applied Mathematics. Orlando etc.: Academic Press (Harcourt Brace Jovanovich, Publishers). XV, 570 p., 1985.

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

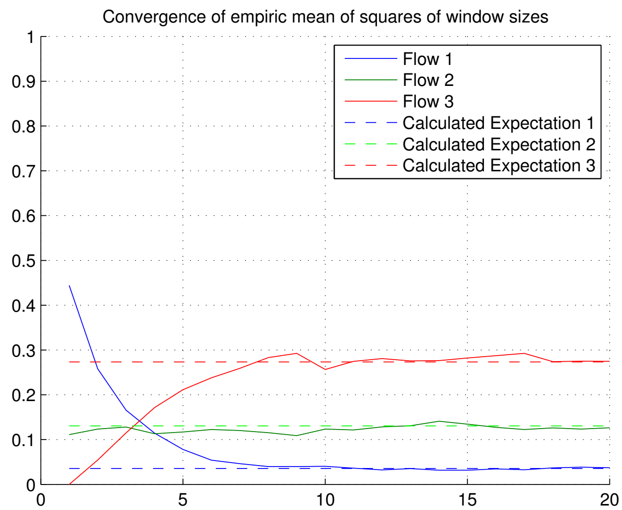


(a)



(b)

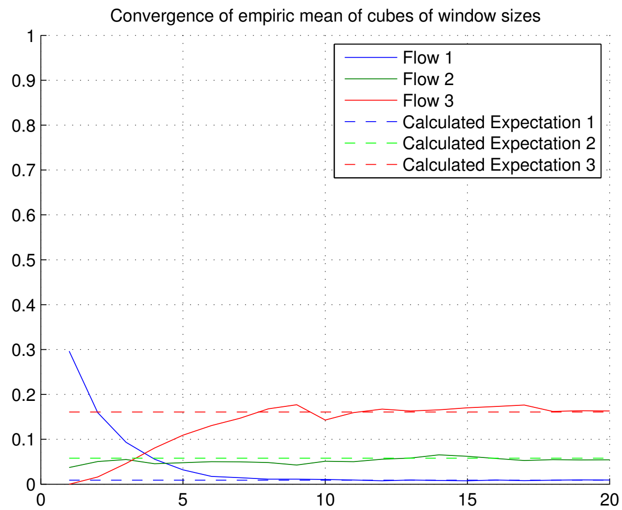
Figure 3: Graphical presentation of results of MATLAB simulations for first moments. (a) shows how the window sizes averaged over 100 trials converge to the calculated first moment; (b) shows how in one trial the temporal mean of the window sizes converges to the calculated first moment. Both horizontal axes describe time measured in congestion events.



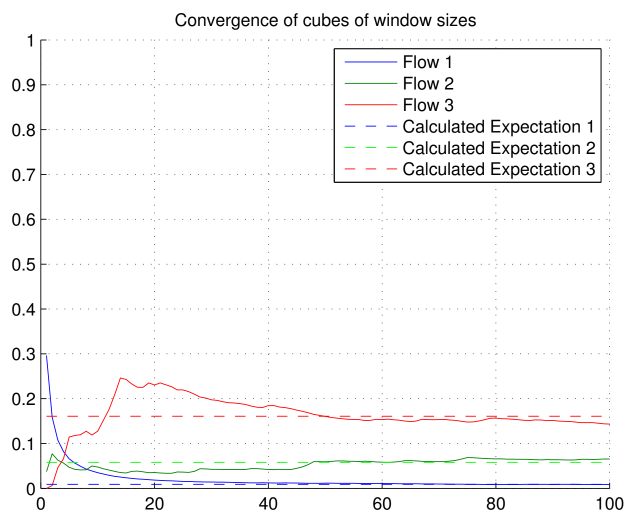
(a)

(b)

Figure 4: Graphical presentation of results of MATLAB simulations for second moments. (a) shows how the squares of the window sizes averaged over 100 trials converge to the calculated second moment; (b) shows how in one trial the temporal mean of the squares of the window sizes converges to the calculated second moment. Both horizontal axes describe time measured in congestion events.



(a)



(b)

Figure 5: Graphical presentation of results of MATLAB simulations for third moments. (a) shows how the cubes of the window sizes averaged over 100 trials converge to the calculated third moment; (b) shows how in one trial the temporal mean of the cubes of the window sizes converges to the calculated third moment. Both horizontal axes describe time measured in congestion events.