Tools for the analysis and design of communication networks with Markovian dynamics

Arie Leizarowitz , Robert Shorten, Rade Stanojević

***Abstract*— In this paper we analyze the stochastic properties of a class of communication networks whose dynamics are Markovian. We characterize the asymptotic behavior of such network in terms of the first and second moments of a stochastic process that describes the network dynamics, and provide tools for their calculation. Specifically, we provide computation techniques for the calculation of these statistics and show that these algorithms converge exponentially fast. Finally, we suggest how our results may be used for the design of network routers to realize networks with desired statistical properties.**

I. Introduction

A. General remarks

The study of communication networks that carry TCP (the transmission control protocol) traffic has been subject of intense interest in the Computer Science, Network Engineering, and Applied Mathematics literature; see for example [5], [8], [11], [12], [14], [15], [16], [17], [23].

The principal motivation for much of this work has been to understand network behaviour, and to characterise important network properties with a view to developing analytic tools for the design of such networks. In particular, much of this work has focussed on understanding the manner in which the network allocates available bandwidth amongst competing network flows (*network fairness*) and the speed at which this bandwidth allocation takes place (*network convergence rate*). Recently, a very accurate random matrix model of TCP network dynamics was proposed in [21]. This model was shown to be capable of capturing many essential features of networks in which TCP-like network flows compete for bandwidth via a bottleneck router. By making some simplifying assumptions concerning stochastic behaviour of the network, the authors demonstrate that this model may also be used as a basis to design networks in which bandwidth can be allocated in an arbitrary manner amongst competing flows. This may be achieved by redesigning the manner in which individual sources respond to network congestion, or by redesigning the manner in which network routers respond to network congestion (or both).

The objective of this paper is to pursue further this line of research. However, rather than using the model as a basis for adjusting the behavior of the the individual flows to achieve desired network behavior, we concentrate here on using this model to redesign the manner in which the bottleneck router drops packets when the network is congested; in particular, we analyze the properties of such networks when the bottleneck router drops packets according to some Markovian rules[1]. Our principal contribution in this paper is to characterize the stochastic behavior of these networks in terms of the first and second moments of a stochastic process that describes the network dynamics, and develop computational techniques for the calculation of these statistics. We concentrate on these statistics as they provide a characterization of the average long term fairness properties of network, and some measure of the instantaneous deviation (instantaneous unfairness) from this measure. Finally, we suggest how our results may be used to design new types of communication networks.

B. Brief description of AIMD congestion control algorithms

Most traffic in communication networks is carried using the TCP protocol.[2] The standard TCP protocol (introduced by Jacobson paper [9]) is a special case of AIMD congestion control. Here we give a very brief description of the AIMD congestion control strategy; the interested reader is referred to [10], [25] for detailed description of the protocol.

A communication network consists of a number of sources and sinks connected together via links and routers. In this paper 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 *Additive-Increase Multiplicative Decrease* (AIMD) -like congestion control algorithm. AIMD congestion control operates a window based congestion control strategy. Each source maintains an internal variable $cwnd_i$ (the window size) which tracks the number of sent unacknowledged packets that can be in transit at any time, i.e. the number of packets in flight.

On safe receipt of data packets the destination sends acknowledgement (ACK) packets to inform the source. When the window size is exhausted, the source must wait for an

Email: la@techunix.technion.ac.il Department of Mathematics Technion Israel

Email: Robert.Shorten@nuim.ie Hamilton Institute NUI Maynooth Ireland

Email: Rade.Stanojevic@nuim.ie Hamilton Institute NUI Maynooth Ireland

[1]Redesigning the manner in which network routers operate to allocate bandwidth is very important for a number of reasons related to network quality of service issues. While the results in [21] are interesting from a theoretical perspective, router redesign along the lines suggested by this work would place an impossible computational burden on the network routers; on the other hand, dropping packets according to some Markovian rule could possibly by implemented using far less computational resources.

[2]$85\% - 90\%$ of all internet traffic is TCP-traffic [26].

ACK before sending a new packet. Congestion control is achieved by dynamically adapting the window size according to an additive-increase multiplicative-decrease law. Roughly speaking, the basic idea is for a source to probe the network for spare capacity by increasing the rate at which packets are inserted into the network, and to rapidly decrease the number of packets transmitted through the network when congestion is detected through the loss of data packets. In more detail, the source increments $cwnd_i(t)$ by a fixed amount α_i upon receipt of each ACK. On detecting packet loss, the variable $cwnd_i(t)$ is reduced in multiplicative fashion to $\beta_i cwnd_i(t)$. 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.

C. AIMD model and problem description

Various types of models for AIMD networks have been developed by several authors, see for example [22] or [13] and the references therein for an overview of this work. We base our discussion on a recently developed random matrix model of AIMD dynamics that was first presented in [21]. This model uses a set of stochastic matrices to characterize the behaviour of a network of AIMD flows that compete for bandwidth via a single bottleneck router (as depicted in figure 1). While other similar random matrix models have been proposed in the literature [2], [3], the model proposed in [21] has several attractive features. In particular, the authors use sets of nonnegative column stochastic matrices to model the evolution of communication networks and results from Frobenius-Perron theory to characterize the stochastic properties of such networks. We begin our discussion by reviewing the essential features of this model.

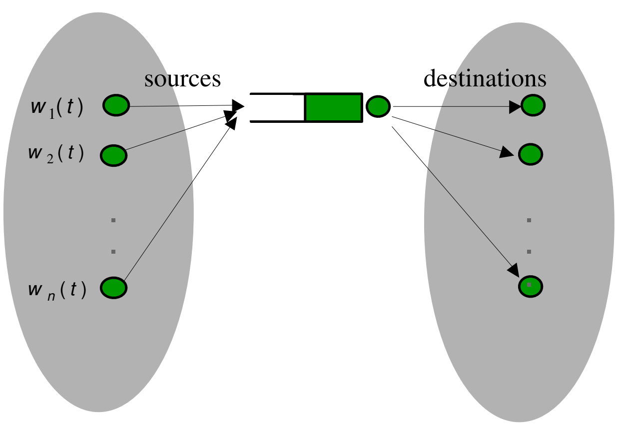

Fig. 1: Network with single bottleneck router.

Suppose that the network under consideration has n flows, all of them operating an Additive Increase Multiplicative Decrease (AIMD) congestion control algorithm, competing for bandwidth over a bottleneck link which has drop-tail queue. Then the current state of the network at times when a packet is dropped at the bottleneck router (referred to as the k'th congestion event) is given by the number of packets in flight that belong to each network source at this time. We describe the network state at the k-th congestion event by an n-dimensional vector $W(k) = \{w_i(k)\}_{i=1}^n$ where $w_i(k)$ is the i-th component of $W(k)$, which is equal to the *throughput* that belong to the i'th source when this source is informed of network congestion. It has been shown in [21] that the sequence $\{W(k)\}_{k=0}^\infty$ satisfies:

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

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

$$A(k) = \begin{bmatrix} \beta_1(k) & 0 & \cdots & 0 \\ 0 & \beta_2(k) & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \beta_n(k) \end{bmatrix} + \frac{1}{\sum_{j=1}^n \alpha_j \gamma_j} \begin{bmatrix} \alpha_1\gamma_1 \\ \alpha_2\gamma_2 \\ \cdots \\ \alpha_n\gamma_n \end{bmatrix} \begin{bmatrix} 1-\beta_1(k), & \cdots & ,1-\beta_n(k) \end{bmatrix}. \tag{I.2}$$

For every $j \in \{1, 2, \ldots, n\}$, the constant $\alpha_j > 0$ in (I.2) is the Additive Increase parameter and $\gamma_j > 0$ is the constant $1/RTT_j^2$. Here RTT_j is the round-trip time for a packet from the j-th flow just before congestion, and either $\beta_j(k) = 1$, which holds if the j-th flow didn't lose any packet during the k-th congestion event, or $\beta_j(k)$ is equal to the Multiplicative Decrease parameter $\beta_j^0 \in (0, 1)$ if the j-th flow did lose some packet in the k-th congestion event.

Comment: *We exclude the possibility that $\beta_1(k) = \beta_2(k) = \cdots = \beta_n(k) = 1$, since there is no congestion event without losing at least one packet.*

We denote by $\mathcal{M}$ the set of the possible values of the matrices $A(k)$, so that

$$\mathcal{M} = \{M_1, M_2, \ldots, M_m\}$$

for some $m \leq 2^n - 1$ and for all k. Then $A(k) \in \mathcal{M}$ for every $k \geq 0$, and we remark that strict inequality $m < 2^n - 1$ may hold; namely that in the models which we consider, certain configurations of packets' loss cannot practically occur.

Let $I(k) = \{j : \beta_j(k) = \beta_j^0\}$ be the set of labels of flows which have experienced a loss of a packet during the k-th congestion event. Note that for each k the matrix $A(k)$ has a strictly positive j-th column if and only if $j \in I(k)$, and that for $j \notin I(k)$, the j-th column of $A(k)$ is equal to e_j, the j-th column of the identity $n \times n$ matrix I_n. We denote by Σ the $(n-1)$-dimensional simplex of all the n-dimensional stochastic vectors. Recall that a vector $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ is stochastic if each one of its coordinates v_i is nonnegative and $v_1 + \cdots + v_n = 1$. It turns out that the matrices M_i $(1 \leq i \leq m)$ which compose $\mathcal{M}$, are nonnegative and column-stochastic [4]. Therefore $M_i(\Sigma) \subset \Sigma$ holds for every $1 \leq i \leq m$. By normalizing $W(0)$ to belong to Σ we may therefore assume, with no loss of generality, that $W(k) \in \Sigma$ for every $k \geq 0$.

For networks with routers employing a drop-tail queueing discipline, it is often assumed in the networking community that congestion events may in some circumstances be modelled as sequences of independent events [1], [2], [19], [21]. In terms of the model described above, this means that for networks with a single bottleneck link and with a drop-tail queue the following holds:

Assumption (i) $\{A(k)\}_{k\in N}$ is a sequence of independent and identically distributed (i.i.d) random variables $A(k)$, and for every $j \in \{1,2,\ldots,n\}$ the probability that the j-th flow detects a drop in each congestion event is positive.

As we have mentioned already, in designing rules which determine how will the network react to congestion, one can typically have two approaches. The first is the design of flow-based congestion control algorithms and the second is the design of queueing discipline. In the present paper we concentrate on the latter and propose two new queueing disciplines which we characterize by computing the stationary statistics for the vector $W(k)$ for each of these cases. We first consider queuing discipline with the property that packets are dropped from the queue in such a manner that the following assumption is valid:

Assumption (ii) $\{A(k)\}_{k\in N}$ is a *stationary* Markov chain on the finite set of matrices $\mathcal{M}$ with transition matrix $P \in \mathbb{R}^{m\times m}$. Moreover we assume that for each $j \in \{1,2,\ldots,n\}$ there exists a matrix $M \in \mathcal{M}$ with positive j-th column.

We note that as in the i.i.d. case, we must have the latter assumption in order to ensure that each flow will see a drop at some point. We also remark that stationarity is assumed to avoid technical difficulties and it is not essential. An additional assumption in Sections II and III (which is relaxed in IV) is that the transition matrix P has strictly positive entries. Theorems 2.9 and 3.6 give the asymptotic values of $E(W(k))$ and $E[(W(k))(W(k))^T]$ in the limit where k tends to infinity, which we denote V^* and D^* respectively. Although we do not have explicit formulas for V^* and D^*, Theorems 2.7 and 3.5 provide iterative algorithms for computing them in a geometric convergence rate. In section IV we extend the results of the previous two sections to the case where the matrix P is merely primitive and its entries are not necessarily strictly positive.

The second queueing discipline we propose here is the following: the probability that a certain set of flows will detect a drop during the k-th congestion event depends only on the vector $W(k)$. Formally we assume that the router drops packet from the queue when it is full in such fashion that the following is true:

Assumption (iii) $\{W(k)\}_{k\in N}$ is a stochastic process in the set of stochastic vectors Σ, which has the following property:
For every $i \in \{1,2,\ldots,m\}$ and $w \in \Sigma$, the conditional probability of $A(k)$ given $W(k)$ is expressed by

$$P[A(k) = M_i|\ W(k) = w] = p_i(w)$$

for some positive continuous functions $p_i : \Sigma \to \mathbb{R}^+$ which satisfy $\sum_{i=1}^m p_i(w) = 1$ for all $w \in \Sigma$. Again, for each $i \in \{1,2,\ldots,n\}$ we require the existence of a matrix $M \in \mathcal{M}$ with positive i-th column.

In view of the relation $W(k+1) = A(k)W(k)$, Assumption (iii) implies that the distribution of $W(k+1)$ is completely determined by the distribution of $W(k)$. Section V is devoted to studying the behavior of $W(k)$ under Assumption (iii). It turns out that the study of the model under Assumption (iii) can be reduced to its study under Assumption (ii). This enables to establish the analogous results concerning the asymptotic behavior of $E(W(k))$ and $Var(W(k))$ for this case. In particular, the latter can be computed by iterative methods producing schemes which converge in a geometric rate.

II. The asymptotic expectation of $W(N)$

In this section we compute the equilibrium expected value of the window size variable $W(N)$ under Assumption (ii), and supposing that the transition probabilities P_{ij} are positive:

$$P_{ij} > 0 \text{ for every } 1 \le i,j \le m. \tag{II.1}$$

Denoting by $\rho = (\rho_1,...,\rho_m)$ the unique equilibrium distribution corresponding to P, we associate with P_{ij} the *backward transition probabilities* matrix $\tilde{P}$ (see [18], Chapter 1.9) given by

$$\tilde{P}_{ij} = \frac{\rho_i}{\rho_j}P_{ij} = \frac{P[A(k-1)=M_i]}{P[A(k)=M_j]}P[A(k)=M_j|A(k-1)=M_i] = = P[A(k-1)=M_i|A(k)=M_j]. \tag{II.2}$$

We interpret $\tilde{P}_{ij}$ as the conditional probability that the system occupied the state M_i at the previous instant of time given that it is presently at state M_j, for the *stationary* Markov chain $\{A_k\}$.

Let $\Phi : (\mathbb{R}^n)^m \to (\mathbb{R}^n)^m$ be the linear mapping given by:

$$\Phi(V) = \left(\sum_{i=1}^m \tilde{P}_{i1}M_iV_i, \ldots, \sum_{i=1}^m \tilde{P}_{im}M_iV_i\right) \tag{II.3}$$

where $V = (V_1,...,V_m)$, $V_i \in \mathbb{R}^n$ and $M_i \in \mathcal{M}$ for every $1 \le i \le m$. We have the following result:

Proposition 2.1: For an arbitrary $W(0) = s \in \Sigma$ and all $i = 1,2,\ldots,m$, the following limits exist:

$$V_i = \lim_{k\to\infty} E[W(k)|A(k) = M_i],\ i = 1,2,...,m. \tag{II.4}$$

Moreover, the vector $V = (V_1,...,V_m) \in \Sigma^m$ whose components are defined in $(II.4)$ satisfies the fixed point equation

$$V = \Phi(V). \tag{II.5}$$

Proof. The proof will be given after establishing Theorem 2.7 □

Let $\mathcal{S}$ be the subspace of $\mathbb{R}^n$ defined by:

$$\mathcal{S} = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^{n} x_i = 0 \right\}.$$

It turns out that Φ has the following property:

Proposition 2.2: The mapping Φ is linear from Σ^m into itself, and from $\mathcal{S}^m$ into itself.

Proof. Both claims follow from the facts that if M is a column stochastic $n \times n$ matrix, then for every $x \in \mathbb{R}^n$ we have

$$\sum_{i=1}^{n} (Mx)_i = \sum_{i=1}^{n} x_i,$$

and that each M_i is a column stochastic matrix. □

We know that any matrix in $\mathcal{M}$ can be written in the form

$$M = \text{diag}(\beta_1, \beta_2, ..., \beta_n) + (\delta_1, \ldots, \delta_n)^T ((1-\beta_1), \ldots, (1-\beta_n))$$

where $0 < \beta_k \leq 1$ for every $1 \leq k \leq n$ and not all of them are equal to 1. We denote $\beta = (\beta_1, ..., \beta_n)^T$, $\delta = (\delta_1, ..., \delta_n)^T$, and δ is a stochastic vector with positive entries. If the first q entries of β are equal to 1, namely $\beta_1 = \beta_2 = \cdots = \beta_q = 1$ for some $q < n$, and the last $n - q = r > 0$ entries are all smaller than 1, then our matrix M has the following form:

$$M = \begin{pmatrix} I_q & M' \\ 0 & M'' \end{pmatrix}. \tag{II.6}$$

The matrix I_q in (II.6) is the $q \times q$ identity matrix, all the entries in the last r columns are positive, and the sum of the entries in each of these columns is equal to 1.

The following result is the main technical tool that we employ in studying properties of Φ. We denote by $\| \cdot \|_1$ the L_1 norm of vectors in $\mathbb{R}^n$. We will prove it as Corollary 3.3 after having established Lemma 3.2. A direct proof can be found in [24].

Lemma 2.3: Let $M \in \mathcal{M}$. Then for every $x \in \mathcal{S}$ we have

$$Mx \neq x \Rightarrow \|Mx\|_1 < \|x\|_1. \tag{II.7}$$

Proof. See the proof of Corollary 3.3. □

Remark 2.4: The property which is established in Lemma 2.3 is referred to in the literature as the *paracontracting* property; see *[7]* chapter 8 or *[6]*. We have thus showed that the matrices M_i are paracontractive in $\mathcal{S}$ in L_1 norm. We will again use the notion of paracontractivity in Section V.

Lemma 2.5: Suppose that $M \in \mathcal{M}$ is such that the columns which contain zeros are indexed by $i_1, i_2, \ldots, i_q$, and let $x \in \mathcal{S}$ be such that $Mx = x$. Then x belongs to the subspace spanned by the basic vectors $e_{i_1}, e_{i_2}, \ldots, e_{i_q}$.

Proof. We suppose without loss of generality that the first q columns of M contain zeros and the last $r = n - q$ columns are positive, i.e. that M has the form given by (II.6). We will establish that the last r coordinates of x are equal to 0. If $r = 0$ then there is nothing to prove. If $r = n$ then M is a stochastic matrix with strictly positive entries, and therefore it is a contraction on $\mathcal{S}$, implying that $x = 0$.

We now suppose that $0 < r < n$, and let P be the $r \times r$ submatrix of M which is defined by the last r rows and last r columns. If we denote by x' the r-dimensional vector composed of the last r coordinates of x, then $Px' = x'$. But the sum of each column of P is smaller than 1 and x is nonnegative, hence $\|Px'\|_1 < \|x'\|_1$ whenever $x' \neq 0$, implying that x' must vanish. This concludes the proof of the lemma. □

We define on $(\mathbb{R}^n)^m$ the norm

$$\|V\| = \|(V_1, \ldots, V_m)\| = \max_{1 \leq i \leq m} (\|V_i\|_1),$$

and consider the subspace $\mathcal{S}^m$ and subset Σ^m of $(\mathbb{R}^n)^m$ endowed with this norm. The next result establishes that Φ^2 is a contraction on the metric space Σ^m as well as on normed space $\mathcal{S}^m$.

Proposition 2.6: Let Φ be the mapping given by $(II.3)$. We assume that $(II.1)$ holds, so that in view of $(II.2)$, all the backward probabilities $\tilde{P}_{ij}$ are positive as well. Then there exists a constant $\theta < 1$ such that

$$\|\Phi^2(U) - \Phi^2(V)\| \leq \theta \|U - V\| \tag{II.8}$$

holds for all $U, V \in \Sigma$.

Proof. We will establish that for every pair $U \neq V$ in Σ, the inequality $\|\Phi^2(U) - \Phi^2(V)\| < \|U - V\|$ holds. This will imply the assertion of the proposition in view of the compactness of Σ^m.

Thus let $U = (U_1, \ldots, U_m)$ and $V = (V_1, \ldots, V_m)$ be any two different elements belonging to Σ^m. We have

$$\|\Phi(U) - \Phi(V)\| = \max_j \| \sum_{i=1}^{m} \tilde{P}_{ij} M_i (U_i - V_i) \|_1 \tag{II.9}$$

$$\leq \max_j \sum_{i=1}^{m} \tilde{P}_{ij} \|M_i (U_i - V_i)\|_1 \tag{II.10}$$

$$\leq \max_j \sum_{i=1}^{m} \tilde{P}_{ij} \|U_i - V_i\|_1 \tag{II.11}$$

$$\leq \max_j \sum_{i=1}^{m} \tilde{P}_{ij} \|(U - V\| = \|U - V\|. \tag{II.12}$$

We will next check under which conditions equality $\|\Phi^2(U) - \Phi^2(V)\| = \|U - V\|$ can hold. We thus assume that $U \neq V$ are such that $\|\Phi(\Phi(U)) - \Phi(\Phi(V))\| = \|U - V\|$. It follows from

$$\|U - V\| = \|\Phi(\Phi(U)) - \Phi(\Phi(V))\| \leq \|\Phi(U) - \Phi(V)\| \leq \|U - V\|$$

that $\|\Phi(U) - \Phi(V)\| = \|U - V\|$. Thus in this situation all the inequalities in (II.9)-(II.12) are actually equalities.

We now denote $W = U - V \in \mathcal{S}^m$ and we note that in view of (II.12), for some j,

$$\sum_{i=1}^{m} \tilde{P}_{ij} \|W_i\|_1 = \max_i (\|W_i\|_1) = \|W\|. \tag{II.13}$$

Since we suppose that all $\tilde{P}_{ij}$ are positive, (II.13) implies

$$\|W_i\|_1 = \|W\| \text{ for every } 1 \leq i \leq m. \tag{II.14}$$

It then follows from (II.11) that

$$\max_j \sum_{i=1}^{m} \tilde{P}_{ij} \|M_i(W_i)\|_1 = \max_j \sum_{i=1}^{m} \tilde{P}_{ij} \|W_i\|_1 = \|W\|,$$

which in view of $\|M_i(W_i)\|_1 \leq \|W_i\|_1$ and the positivity of all the P_{ij}, implies

$$\|M_i(W_i)\|_1 = \|W\| \tag{II.15}$$

for all i. It follows from (II.14), (II.15) and Lemma 2.3 that

$$M_i(W_i) = W_i \tag{II.16}$$

for all i. We thus conclude from (II.10), (II.15) and (II.16) that there exist some j such that

$$\|\sum_{i=1}^{m} \tilde{P}_{ij}(W_i)\|_1 = \sum_{i=1}^{m} \tilde{P}_{ij}\|W_i\|_1 = \|W\|.$$

However, this can happen if and only if for every $r \in \{1, 2, \ldots, n\}$, there does not exist $1 \leq i, j \leq m$ such that the r-th coordinates $(W_i)_r$ and $(W_j)_r$ are of opposite signs.

By employing the above argument, and also the conclusion (II.16) to the equality $\|\Phi(\Phi(W))\| = \|\Phi(W)\|$ rather than to $\|\Phi(W)\| = \|W\|$, we have that for all $k \in \{1, 2, \ldots, m\}$

$$M_k\left(\sum_{i=1}^{m} \tilde{P}_{ik}W_i\right) = \sum_{i=1}^{m} \tilde{P}_{ik}W_i. \tag{II.17}$$

Assumption (ii) of our model is such that for every $r \in \{1, 2, ..., n\}$ there exists a matrix $M_k \in \mathcal{M}$ with positive r-th column. It follows from Lemmas (2.5) and (II.17) that the r-th coordinate of $\sum_{i=1}^{m} \tilde{P}_{ik}W_i$ must vanish. But we have that there are no two indices i_1 and i_2 such that the r-th coordinates of W_{i_1} and W_{i_2} have opposite signs. This fact implies that the r-th coordinate of the vector W_i must vanish, and this is true for every $1 \leq i \leq m$. Since r is arbitrary, we conclude that $W_i = 0$ for all i. We have thus established that if $U, V \in \Sigma^m$ are distinct then

$$\|\Phi^2(U) - \Phi^2(V)\| < \|U - V\|.$$

The proof of the proposition is thus complete. □

Theorem 2.7: The spectral radius of the restriction of Φ to $\mathcal{S}^m$ is smaller than 1. In particular there exists a unique solution $V^\star$ for equation $(II.5)$, and the iteration scheme

$$V^{(k+1)} = \Phi(V^{(k)}), k = 0, 1, 2, ...$$

with any starting point $V^{(0)} = V_0$ in Σ satisfies

$$\lim_{k\to\infty} V^{(k)} = V^\star. \tag{II.18}$$

Proof. We will first establish that the spectral radius of the restriction of Φ to $\mathcal{S}^m$ is smaller than 1. But we know that the iterations of Φ^2 in $\mathcal{S}^m$ converge to zero for every starting point. It follows that every eigenvalue of the restriction of Φ to $\mathcal{S}^m$, say λ, satisfies $|\lambda| \leq 1$. If, however, there exists an eigenvalue which is equal to $e^{i\theta}$ for some real θ, then there exists a subspace Π of $\mathcal{S}^m$, which is invariant under Φ, and is either one- or two-dimensional. The restriction of Φ to Π is then a rotation, contradicting the fact that the iterations of Φ should tend to zero.

The uniqueness of solutions of (II.5) follows from the contractive property of Φ^2. Then for every initial $V_0 \in \Sigma$ we have $\lim_{k\to\infty} V^{(2k)} = V^\star$. Hence

$$V^{(2k+1)} = \Phi(V^{(2k)}) \to \Phi(V^\star) = V^\star \text{ as } k \to \infty,$$

and (II.18) follows. □

Proof of Proposition 2.1. Let $V_i(k) = E[W(k)|A(k) = M_i]$. The sequence of vectors $V(k) = (V_1(k), \ldots, V_m(k)) \in \Sigma^m$ satisfies

$$V(k+1) = \Phi(V(k)). \tag{II.19}$$

From Theorem (2.7), $\{V(k)\}_{k=0}^{\infty}$ converge and the existence of the limits in (II.4) follows. In view of (II.19) these limits satisfy the fixed point equation (II.5). □

We have the following result which is actually Theorem 3.1 from [21]:

Corollary 2.8: Let Assumption (i) hold, so that the probability that $A(k) = M_i$ is equal to ρ_i for every $k \geq 0$ and $1 \leq i \leq m$. Then the asymptotic expected value of $W(k)$ is the unique stochastic eigenvector of $\sum_{i=1}^{m} \rho_i M_i$ which corresponds to the eigenvalue 1.

Proof. The sequence $\{A(k)\}$ of i.i.d. random matrices can be seen as a Markov chain on the set $\mathcal{M} = \{M_i \ : \ \rho_i > 0\}$ with the $m \times m$ transition matrix P given by $P_{ij} = \rho_j$. Since P_{ij} is positive for every i and j we have that $\tilde{P}_{ij} = \rho_i P_{ij}/\rho_j = \rho_i > 0$. We look for a solution of equation (2.1) for which all the components V_i are the same, say equal to $\bar{V}$. This yields the equation

$$\bar{V} = \left(\sum_{i=1}^{m} \rho_i M_i\right)\bar{V},$$

which implies the assertion of the corollary. □

Theorem 2.9: Under Assumption (ii), and assuming that the transition matrix P has strictly positive entries, then the asymptotic behavior of the expectation of the random variable $W(N)$ is given by:

$$\lim_{N\to\infty} E(W(N)) = \sum_{i=1}^{m} \rho_i V_i^\star, \tag{II.20}$$

where $V^\star = (V_1^\star, \ldots, V_m^\star) \in \Sigma^m$ is the unique solution of $(II.5)$, and $\rho = (\rho_1, \ldots, \rho_m)$ is the Perron eigenvector of the transition probability matrix (P_{ij}).

Proof. The proof is immediate:

$$\lim_{N\to\infty} E(W(N)) =$$

$$\lim_{N\to\infty} \sum_{i=1}^{m} E[W(N) \mid A(N) = M_i]P[A(N) = M_i] = \sum_{i=1}^{m} \rho_i V_i^\star.$$

□

III. The asymptotic variance of $W(N)$

The goal of this section is to compute the asymptotic value of the second order[3] moment of $W(N)$ under Assumption (ii) and assuming a positive transition matrix P.

[3]What we call variance, or second order moment is actually covariance matrix for the vector $W(N)$: $E[(W(N))_i((W(N))_j)] - E[(W(N))_i]E[(W(N))_j]$.

Define the linear mapping $\Psi : (\mathbb{R}^{n\times n})^m \to (\mathbb{R}^{n\times n})^m$ by:

$$\Psi(D_1,\dots,D_m) = \left(\sum_{i=1}^m \tilde{P}_{i1} M_i D_i M_i^T, \dots, \sum_{i=1}^m \tilde{P}_{im} M_i D_i M_i^T\right) \tag{III.1}$$

Suppose for a moment that for $W(0) = s \in \Sigma$ the following limits exist:

$$D_i = \lim_{k\to\infty} E[W(k)W(k)^T | A(k) = M_i].$$

Comment : *Note that each D_i must be a symmetric nonnegative definite matrix, that it has nonnegative entries and it satisfies*

$$D_i u = \lim_{k\to\infty} E[W(k)W(k)^T u \mid A(k) = M_i] = \lim_{k\to\infty} E[W(k) \mid A(k) = M_i] = V_i, \tag{III.2}$$

where u is the n-dimensional vector which satisfies $u_i = 1$ for every $1 \le i \le n$.

In view of (III.2) let $\mathcal{D}$ be the set:

$$\mathcal{D} = \{(D_1,\dots,D_m) \mid D_i \in \mathbb{R}^{n\times n},\ D_i = D_i^T,\ D_i u = V_i\} \tag{III.3}$$

It turns out that Ψ maps $\mathcal{D}$ into itself. Indeed, $(\Psi(D))_j$ is symmetric whenever all D_i are such. Moreover, using (II.5) we obtain

$$\sum_{i=1}^m \tilde{P}_{ij} M_i D_i M_i^T u = \sum_{i=1}^m \tilde{P}_{ij} M_i D_i u = \sum_{i=1}^m \tilde{P}_{ij} M_i V_i = V_j,$$

implying that $\Psi(D) \in \mathcal{D}$ for every $D \in \mathcal{D}$. We have the following result:

Proposition 3.1: For arbitrary $W(0) = s \in \Sigma$ and all $i \in \{1, 2, \dots, m\}$ the following limits exist:

$$D_i = \lim_{k\to\infty} E[W(k)W(k)^T | A(k) = M_i]. \tag{III.4}$$

The m-tupple $D = (D_1, ..., D_m) \in \mathcal{D}$ defined by $(III.4)$ satisfies the fixed point equation

$$D = \Psi(D). \tag{III.5}$$

Proof. The proof will be given after having established Theorem (3.5) □

Let

$$\mathcal{B} = \{C \mid C \in \mathbb{R}^{n\times n},\ C = C^T,\ Cu = 0\}. \tag{III.6}$$

Thus $\mathcal{B}$ is the vector space of all $n \times n$ symmetric matrices C such that all the columns of B belong to $\mathcal{S}$. A computation similar to the one preceding Proposition 3.1 implies that $\Psi(\mathcal{B}^m) \subset \mathcal{B}^m$. Since the difference between any two elements from $\mathcal{D}$ belongs to $\mathcal{B}^m$, then fixing any norm on $(\mathbb{R}^{n\times n})^m$, it follows that the linear mapping Ψ is a contraction on the metric space $\mathcal{D}$ if it is a contraction on the vector space $\mathcal{B}^m$. We wish to establish the existence and uniqueness of solutions $D \in \mathcal{D}$ of equation (III.5). To this end it is enough to find a norm in which the mapping Ψ^2 is a contraction on the complete metric space $\mathcal{D}$.

Let $\|\cdot\|$ be the norm on $\mathbb{R}^{n\times n}$ defined by :

$$\|A\| = \sum_{i,j=1}^n |A_{ij}| \quad \text{for } A \in \mathbb{R}^{n\times n}.$$

The next result establishes a crucial connection between this norm and the mapping $C \mapsto MCM^T$ for $C \in \mathcal{B}$ and $M \in \mathcal{M}$. It is close in spirit to Lemma 2.3.

Lemma 3.2: Let $M \in \mathcal{M}$. Then the following relation

$$MC \neq C \Rightarrow \|MCM^T\| < \|C\| \tag{III.7}$$

holds for every $C \in \mathcal{B}$.

Proof. As in the proof of Lemma 2.5, we consider a matrix M which has the form (II.6) for some $0 \le q < n$, and where the last $r = n - q$ columns are positive. We then have

$$\|MCM^T\| = \sum_{i,j} |\sum_{k,l} m_{ik} c_{kl} m_{jl}| \le \sum_{i,j}\sum_{k,l} m_{ik} m_{jl} |c_{kl}| =$$

$$= \sum_{k,l} |c_{kl}| \sum_i m_{ik} \sum_j m_{jl} = \sum_{k,l} |c_{kl}| = \|C\| \tag{III.8}$$

since M is column stochastic. We have thus established

$$\|MCM^T\| \le \|C\| \tag{III.9}$$

for any column stochastic matrix M and any matrix C. We will next prove that equality holds in (III.9) only if $MC = C$. We remark that if $MC = C$, then in view of $C = C^T$, we have

$$MCM^T = CM^T = (MC)^T = C^T = C,$$

implying that equality holds in (III.9) if $MC = C$.

We now suppose that M and C are such that $\|MCM^T\| = \|C\|$. This is possible if and only if for each pair $1 \le i, j \le n$, the only inequality which appears in (III.8) is actually an equality. This, however, holds if and only if for every $1 \le i, j \le n$ the following holds:

Property S. *There are no two pairs of indices (k, l) and (k', l') such that $m_{ik} m_{jl}$ and $m_{ik'} m_{jl'}$ are both positive while c_{kl} and $c_{k'l'}$ have opposite signs.*

For $1 \le i \le q$ let $\Lambda_i = \{c_{il} \mid q < l \le n\} = \{c_{li} \mid q < l \le n\}$, and denote $\Lambda_0 = \{c_{kl} \mid q < k \le n,\ q < l \le n\}$. Using Property S for a pair $i, j \in \{1, 2, \dots, q\}$, and noting that for all $k, l \in \{q+1, q+2, \dots, n\}$, we have $m_{ik} m_{jj} > 0$, $m_{ii} m_{jl} > 0$ and $m_{ik} m_{jl} > 0$, and further we conclude that there are no two elements in set $\Lambda_{ij} = \Lambda_i \cup \Lambda_j \cup \Lambda_0$ with opposite sign. Since for each pair of indices (k, l) and (k', l') with $\max\{k, l\} > q$ and $\max\{k', l'\} > q$ there is pair $i, j \in \{1, 2, \dots, q\}$ such that both c_{kl} and $c_{k'l'}$ are contained in Λ_{ij}, we conclude that either

$$c_{kl} \ge 0 \text{ whenever } \max\{k, l\} > q, \tag{III.10}$$

or

$$c_{kl} \le 0 \text{ whenever } \max\{k, l\} > q. \tag{III.11}$$

For any integer $q < l \le n$ the sum of the entries in the l-th column (or row) have the same sign as the constant sign of its elements, namely nonnegative (non-positive) if (III.10) ((III.11)) holds. Since this sum is zero, we conclude that all the entries c_{kl} that are such that at least one of $k > q$ and $l > q$ holds must vanish. Thus C must have the form:

$$C = \begin{pmatrix} C' & 0 \\ 0 & 0 \end{pmatrix}$$

where $C' \in \mathbb{R}^{q\times q}$, and since $M = \begin{pmatrix} I_q & M' \\ 0 & M'' \end{pmatrix}$, it follows that $MC = C$, concluding the proof of the lemma. □

Now we are able to present the following short proof of Lemma 2.3.

Corollary 3.3: Let $M \in \mathcal{M}$. Then for every $x \in \mathcal{S}$ we have

$$Mx \neq x \Rightarrow \|Mx\|_1 < \|x\|_1. \tag{III.12}$$

Proof. Note that for $x \in \mathcal{S}$, $C = xx^T \in \mathcal{B}$ we can conclude that

$$\|MCM^T\| = \|Mxx^TM^T\| = \|Mx\|_1^2 \leq \|C\| = \|x\|_1^2$$

with equality if and only if $MC = C$. Moreover from the proof of the previous lemma we conclude that for all j such that the j-th column of M is positive, the j-th column of C must be zero, which also means that $x_j = 0$. Thus $\|Mx\|_1 = \|x\|_1$ implies that $Mx = x$. □

We wish to employ the Banach fixed point theorem to the mapping Ψ on the set $\mathcal{D}$, with the metric which is induced on $\mathcal{D}$ by the following norm

$$\|B\| = \|(B_1, \ldots, B_m)\| = \max_{1\leq i\leq m} \|B_i\| \tag{III.13}$$

on $(\mathbb{R}^{n\times n})^m$. Given this, we are now ready to establish that Ψ^2 is a contraction on the metric space $\mathcal{D}$.

Proposition 3.4: Let Ψ be the mapping given by $(III.1)$ and we assume that the transition matrix P is positive. Then there exists a constant $\eta < 1$ such that

$$\|\Psi^2(D) - \Psi^2(E)\| \leq \eta\|D - E\| \tag{III.14}$$

holds for all $D, E \in \mathcal{D}$.

Proof. We will establish that for every nonzero $B \in \mathcal{B}^m$ we have $\|\Psi^2(B)\| < \|B\|$, which implies that there exists a $0 < \eta < 1$ such that

$$\|\Psi^2(B)\| \leq \eta\|B\|, \tag{III.15}$$

since $\mathcal{B}$ is a normed linear space. Clearly (III.14) follows from (III.15) since $D - E \in \mathcal{D}$. We thus consider any $B = (B_1, \ldots, B_m) \neq 0$ such that $B_i \in \mathcal{B}$ and compute

$$\|\Psi(B)\| = \max_j \|\sum_{i=1}^m \tilde{P}_{ij} M_i B_i M_i^T\| \tag{III.16}$$

$$\leq \max_j \sum_{i=1}^m \tilde{P}_{ij} \|M_i B_i M_i^T\| \tag{III.17}$$

$$\leq \max_j \sum_{i=1}^m \tilde{P}_{ij} \|B_i\| \tag{III.18}$$

$$\leq \max_j \sum_{i=1}^m \tilde{P}_{ij} \|B\| = \|B\|. \tag{III.19}$$

We will next check under which conditions the equality $\|\Psi^2(B)\| = \|B\|$ can hold. We thus assume that $B \neq 0$ is such that $\|\Psi(\Psi(B))\| = \|B\|$. It follows from

$$\|B\| = \|\Psi(\Phi(B))\| \leq \|\Psi(B)\| \leq \|B\|$$

that $\|\Psi(B)\| = \|B\|$. Thus it follows in this situation that all the inequalities in (III.16)-(III.19) are actually equalities.

In view of (III.18) and (III.19), for some j

$$\sum_{i=1}^m \tilde{P}_{ij}\|B_i\| = \max_i(\|B_i\|) = \|B\|. \tag{III.20}$$

Since we assume that all the $\tilde{P}_{ij}$ are positive, (III.20) implies

$$\|B_i\| = \|B\| \text{ for every } 1 \leq i \leq m. \tag{III.21}$$

It then follows from (III.17) that

$$\max_j \sum_{i=1}^m \tilde{P}_{ij}\|M_i B_i M_i^T\| = \max_j \sum_{i=1}^m \tilde{P}_{ij}\|B_i\| = \|B\|,$$

which together with the positivity of all the $\tilde{P}_{ij}$ imply that

$$\|M_i B_i M_i^T\| = \|B\| \tag{III.22}$$

for all i. In view of Lemma 3.2 it therefore follows from (III.21) and (III.22) that

$$M_i B_i M_i^T = B_i M_i^T = (M_i B_i^T)^T = (M_i B_i)^T = B_i,$$

namely the equalities

$$M_i B_i = B_i \text{ and } M_i B_i M_i^T = B_i \tag{III.23}$$

hold for all i. It follows from (III.16), (III.22) and (III.23) that there exist some j such that

$$\|\sum_{i=1}^m P_{ij} B_i\| = \sum_{i=1}^m P_{ij}\|B_i\| = \|B\|.$$

However, this can happen if and only if the following property holds:

A sign condition. *For every $r, s \in \{1, 2, \ldots, n\}$, there don't exist $1 \leq i, j \leq m$ such that the (rs)-th coordinates $(B_i)_{rs}$ and $(B_j)_{rs}$ have opposite signs.*

Employing the above argument, and the conclusion (III.23), to the equality $\|\Psi(\Psi(B))\| = \|\Psi(B)\|$ it follows that for all $k \in \{1, 2, \ldots, m\}$

$$M_k \left(\sum_{i=1}^m P_{ik} B_i\right) = \sum_{i=1}^m P_{ik} B_i. \tag{III.24}$$

Now let $l \in \{1, 2, \ldots, n\}$ be arbitrary. Then by Assumption (ii) of our model, there exists some matrix $M_k \in \mathcal{M}$ with positive l-th column. From (III.24) we can conclude that the columns of $\sum_{i=1}^m P_{ik} B_i$ are eigenvectors of the matrix M_k, which correspond to the eigenvalue 1. Moreover, they are convex combination of vectors from $\mathcal{S}$, hence they also belong to $\mathcal{S}$. Using Lemma 2.5, we can therefore conclude that the l-th column of the matrix $\sum_{i=1}^m P_{ik} B_i$ must vanish, and by employing the above sign condition, it follows that corresponding entries of the various l-th columns of the matrices M_i don't have opposite signs. This implies that all the entries in the l-th columns of the matrices $B_1, \ldots, B_m$ must vanish. Since l is arbitrary, we conclude that

$$B_1 = B_2 = \cdots = B_m = 0.$$

This contradiction concludes the proof of the lemma. □

Theorem 3.5: The spectral radius of the restriction of the mapping Ψ to $\mathcal{B}^m$ is smaller than 1. In particular there exists a unique solution $D^\star$ for equation $(III.5)$, and the iteration scheme

$$D^{(k+1)} = \Psi(D^{(k)}), k = 0, 1, 2, ...$$

with the starting point $D^{(0)} = D_0$, satisfies

$$\lim_{k\to\infty} D^{(k)} = D^\star$$

for every $D_0 \in \mathcal{D}$.

Proof. The proof of this theorem follows the same lines and uses the same arguments as those employed in the proof of Theorem 2.7. □

Proof of Proposition 3.1. Similarly to the proof of Proposition 2.1, let $D_i(k) = E[W(k)W(k)^T|A(k) = M_i]$ and $D(k) = (D_1(k), \ldots, D_m(k) \in (\mathbb{R}^{n\times n})^m$. Then $D_i(k)u = E[W(k)W(k)^T u|A(k) = M_i] = E[W(k)|A(k) = M_i] = V_i(k) \to V_i$ as $k \to \infty$ which means that

$$\lim_{k\to\infty} \text{dist}(D(k), \mathcal{D}) = 0.$$

Let $\varepsilon > 0$ and let $k_0 \in N$ be such that $\|D(k_0) - \tilde{D}\| < \varepsilon$ for some $\tilde{D} \in \mathcal{D}$. Since the mapping Ψ is nonexpansive in the norm given by (III.13) on all $(\mathbb{R}^{n\times n})^m$ we have:

$$\|D(k_0+r) - \Psi^r(\tilde{D})\| = \|\Psi^r(D(k_0) - \tilde{D})\| \le \|D(k_0) - \tilde{D}\| \le \varepsilon.$$

On the other hand since $\tilde{D}$ is in $\mathcal{D}$,

$$\lim_{r\to\infty} \Psi^r(\tilde{D}) = D'$$

exists by the previous theorem. This means that there is k_1 such that for all $r > k_1$,

$$\|\Psi^r(\tilde{D}) - D'\| \le \varepsilon.$$

Now we conclude that for all $r > k_0 + k_1$:

$$\|D(r) - D'\| \le \|D(r) - \Psi^{r-k_0}(\tilde{D})\| + \|\Psi^{r-k_0}(\tilde{D}) - D'\| \le 2\varepsilon.$$

The last relation means that the sequence $\{D(r)\}$ is a Couchy, and therefore $\lim_{r\to\infty} D(r)$ exists and (III.4) follows. Having this, (III.5) follows from the continuity of the linear mapping Ψ. □

Theorem 3.6: Under Assumption (ii) and the positivity of the transition matrix P, the asymptotic behavior of the variance

$$Var(W(N)) = E[W(N)W(N)^T] - E[W(N)]E[W(N)]^T$$

is given by:

$$\lim_{N\to\infty} Var(W(N)) =$$

$$= \sum_{i=1}^m \rho_i D_i^\star - \left(\sum_{i=1}^m \rho_i V_i^\star\right)\left(\sum_{i=1}^m \rho_i V_i^\star\right)^T \tag{III.25}$$

where $D^\star = (D_1^\star, \ldots, D_m^\star) \in \mathcal{D}$ is the unique solution of $(III.5)$, and $\rho = (\rho_1, \ldots, \rho_m)$ is the Perron eigenvector of the transition matrix (P_{ij}).

Proof. We have the following equalities:

$$\lim_{N\to\infty} Var(W(N)) =$$

$$\lim_{N\to\infty} E[W(N)W(N)^T] - \lim_{N\to\infty} E[W(N)]E[W(N)]^T =$$

$$= \lim_{N\to\infty} \sum_{i=1}^m E[W(N)W(N)^T \mid A(N) = M_i]P[A(N) = M_i]$$

$$- \lim_{N\to\infty} E[W(N)]E[W(N)]^T =$$

$$= \sum_{i=1}^m \rho_i D_i^\star - \left(\sum_{i=1}^m \rho_i V_i^\star\right)\left(\sum_{i=1}^m \rho_i V_i^\star\right)^T.$$

□

Corollary 3.7: Let Assumption (i) hold, so that the probability that $A(k) = M_i$ is equal to ρ_i for every $k \ge 0$ and $1 \le i \le m$. Then the asymptotic behavior of $Var(W(k))$ is given by

$$\lim_{N\to\infty} Var(W(N)) = \bar{D} - \bar{V}\bar{V}^T, \tag{III.26}$$

where $\bar{V}$ is the unique stochastic eigenvector of the matrix $\sum_{i=1}^m \rho_i M_i$, and $\bar{D}$ is the unique solution of the matrix equation

$$\sum_{i=1}^m \rho_i M_i D M_i^T = D,$$

which satisfies $Du = \bar{V}$. Moreover, $\bar{D}$ is the unique eigenvector corresponding to eigenvalue 1 and satisfying $Du = \bar{V}$ of the linear mapping

$$D \mapsto \sum_{i=1}^m \rho_i M_i D M_i^T \tag{III.27}$$

defined on $\mathbb{R}^{n\times n}$.

Proof. The sequence $\{A(k)\}$ of i.i.d. random matrices can be seen as Markov chain on the set $\mathcal{M} = \{M_i : \rho_i > 0\}$ with the $m \times m$ transition matrix P given by $P_{ij} = \rho_j$. Since P is positive for every i and j, it follows that $\tilde{P}_{ij} = \rho_i P_{ij}/\rho_j = \rho_i > 0$. We look for a solution of equation (III.5) for which all the components D_i are the same, say equal to $\bar{D}$. This yields the equation

$$\bar{D} = \sum_{i=1}^m \rho_i M_i \bar{D} M_i^T, \tag{III.28}$$

which implies the first assertion of the corollary.

For the second assertion we represent the linear mapping in (III.27) by an $n^2 \times n^2$ matrix with nonnegative entries, call it T, and apply Perron Frobenius Theorem to T. By Theorem 3.5 the spectral radius of the restriction of T to $\mathcal{B}$ is smaller than 1, but by (III.28) we have that $\bar{D}$ is an eigenvector corresponding to the eigenvalue 1. Since the iterations of the mapping (III.27) converge to $\bar{D}$, it follows that 1 is the unique eigenvector satsfying $\bar{D}u = \bar{V}$. The second assertion follows. □

Example 3.8: In this example we illustrate how the previous results can be applied. We consider a network where the bottleneck router operates according to Assumption(ii). In particular, consider a network of 5 flows with additive increase parameters $\alpha = [5, 4, 3, 2, 1]$, multiplicative decrease parameters given by $\beta = [1/3, 2/4, 3/5, 4/6, 5/7]$, and with corresponding vector γ given by $\gamma = [1/60, 1/70, 1/80, 1/90, 1/100]$. We assume that at congestion events the router drops packets from only one flow. Thus the set $\mathcal{M}$ has 5 elements:

$$M_1 = \begin{bmatrix} 0.5054 & 0 & 0 & 0 & 0 \\ 0.1475 & 1.0000 & 0 & 0 & 0 \\ 0.1291 & 0 & 1.0000 & 0 & 0 \\ 0.1147 & 0 & 0 & 1.0000 & 0 \\ 0.1033 & 0 & 0 & 0 & 1.0000 \end{bmatrix}$$

$$M_2 = \begin{bmatrix} 1.0000 & 0.1291 & 0 & 0 & 0 \\ 0 & 0.6106 & 0 & 0 & 0 \\ 0 & 0.0968 & 1.0000 & 0 & 0 \\ 0 & 0.0860 & 0 & 1.0000 & 0 \\ 0 & 0.0774 & 0 & 0 & 1.0000 \end{bmatrix}$$

$$M_3 = \begin{bmatrix} 1.0000 & 0 & 0.1033 & 0 & 0 \\ 0 & 1.0000 & 0.0885 & 0 & 0 \\ 0 & 0 & 0.6774 & 0 & 0 \\ 0 & 0 & 0.0688 & 1.0000 & 0 \\ 0 & 0 & 0.0620 & 0 & 1.0000 \end{bmatrix}$$

$$M_4 = \begin{bmatrix} 1.0000 & 0 & 0 & 0.0860 & 0 \\ 0 & 1.0000 & 0 & 0.0738 & 0 \\ 0 & 0 & 1.0000 & 0.0645 & 0 \\ 0 & 0 & 0 & 0.7240 & 0 \\ 0 & 0 & 0 & 0.0516 & 1.0000 \end{bmatrix}$$

$$M_5 = \begin{bmatrix} 1.0000 & 0 & 0 & 0 & 0.0738 \\ 0 & 1.0000 & 0 & 0 & 0.0632 \\ 0 & 0 & 1.0000 & 0 & 0.0553 \\ 0 & 0 & 0 & 1.0000 & 0.0492 \\ 0 & 0 & 0 & 0 & 0.7585 \end{bmatrix}$$

Let the transition matrix P to be given by:

$$P = \begin{bmatrix} 0.2667 & 0.2467 & 0.2133 & 0.1667 & 0.1067 \\ 0.2606 & 0.2424 & 0.2121 & 0.1697 & 0.1152 \\ 0.2526 & 0.2368 & 0.2105 & 0.1737 & 0.1263 \\ 0.2444 & 0.2311 & 0.2089 & 0.1778 & 0.1378 \\ 0.2370 & 0.2259 & 0.2074 & 0.1815 & 0.1481 \end{bmatrix}$$

Then

$$lim_{k\to\infty} E(W(k)) = V^\star = \begin{bmatrix} 0.2359 \\ 0.2295 \\ 0.2124 \\ 0.1852 \\ 0.1370 \end{bmatrix}$$

The meaning of this results is that the first flow should expect to get 23.59% of bandwidth, while the fifth flow should expect to get 13.70% of the bandwidth over the bottleneck link, provided that they will last long enough. The asymptotic behavior of variance of $W(k)$ in this example is given by:

$$lim_{k\to\infty} \text{Var}(W(k)) =$$

$$\begin{bmatrix} 0.0144 & -0.0061 & -0.0042 & -0.0027 & -0.0013 \\ -0.0061 & 0.0118 & -0.0029 & -0.0019 & -0.0009 \\ -0.0042 & -0.0029 & 0.0089 & -0.0012 & -0.0005 \\ -0.0027 & -0.0019 & -0.0012 & 0.0060 & -0.0002 \\ -0.0013 & -0.0009 & -0.0005 & -0.0002 & 0.0030 \end{bmatrix}.$$

The rate at which of $V^{(k)}$ and $D^{(k)}$ converge to their equilibrium values is depicted graphically in Figures 2 and 3 respectively.

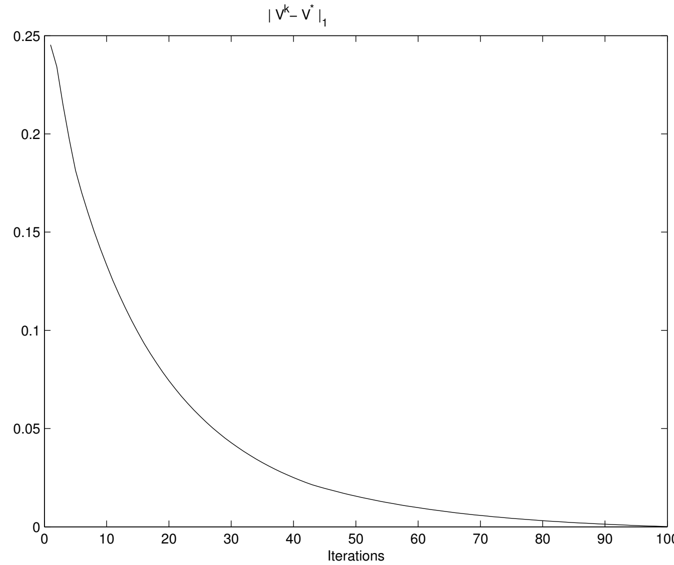

Fig. 2: Evolution of $\|V^{(k)} - V^\star\|_1$.

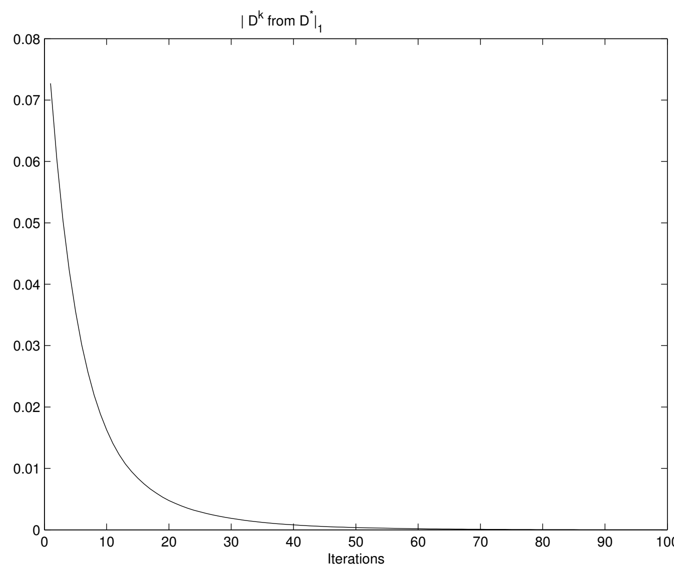

Fig. 3: Evolution of $\|D^{(k)} - D^\star\|$.

IV. A USEFUL EXTENSION

In this section we will extend the results of the previous sections in the following sense. We will consider a transition probability matrix P which does not necessarily have positive entries, but is rather primitive; namely $P^s > 0$, for some integer $1s \geq 1$. We note here that if P is primitive then $\tilde{P}$ is primitive too since they have same zero-nonzero pattern.

Lemma 4.1: If P is a primitive matrix such that $P^s > 0$ for some positive integer s, then Φ^{2s} is a contraction on Σ^m.

Proof. We first note that for all $k, j, l \in \{1, 2, \ldots, m\}$, there is sequence $(i) = (i_1, i_2, \ldots, i_{2s-1})$ of indices which contains l, such that

$$\tilde{P}_{ki_{2s-1}} \tilde{P}_{i_{2s-1}i_{2s-2}} \cdots \tilde{P}_{i_2i_1} \tilde{P}_{i_1j} > 0.$$

Indeed, since $P^s > 0$ there must exist sequences $(i') = (i'_1, \dots, i'_{s-1})$ and $(i'') = (i''_1, \dots, i''_{s-1})$ such that

$$\tilde{P}_{ki'_{s-1}} \tilde{P}_{i'_{s-1}i'_{s-2}} \cdots \tilde{P}_{i'_2 i'_1} \tilde{P}_{i'_1 l} > 0$$

and

$$\tilde{P}_{li''_{s-1}} \tilde{P}_{i''_{s-1}i''_{s-2}} \cdots \tilde{P}_{i''_2 i''_1} \tilde{P}_{i''_1 j} > 0,$$

implying the existence of a sequence (i) with the desired property.

As we noted in section II, proving that Φ^{2s} is a contraction on the metric space Σ^m is equivalent to proving that it is a contraction on the vector space $\mathcal{S}^m$, and here we establish the latter. Thus for an arbitrary $W \in \mathcal{S}^m$ the j component of $\Phi^{2s}(W)$ has the form

$$(\Phi^{2s}(W))_j = \\ = \sum_{(i)_{2s}} \tilde{P}_{i_{2s}i_{2s-1}} \tilde{P}_{i_{2s-1}i_{2s-2}} \cdots \tilde{P}_{i_1 j} M_{i_1} M_{i_2} \cdots M_{i_{2s}} W_{i_{2s}},$$

where we denote by $(i)_{2s}$ a sequence of indices which has length $2s$, and the summation is over all possible sequences $(i)_{2s}$. The sum in the last equality can be rewritten as follows:

$$(\Phi^{2s}(W))_j = \sum_{k=1}^{m} \sum_{(i)} \tilde{P}_{ki_{2s-1}} \cdots \tilde{P}_{i_1 j} M_{i_1} \cdots M_{i_{2s-1}} M_k W_k. \tag{IV.1}$$

The inner sum in (IV.1) is over all the sequences of indices (i) which have the length $2s-1$. Having this in mind we write:

$$\|\Phi^{2s}(W)\| =$$

$$= \max_j \| \sum_{k=1}^{m} \sum_{(i)} \tilde{P}_{ki_{2s-1}} \cdots \tilde{P}_{i_1 j} M_{i_1} \cdots M_{i_{2s-1}} M_k W_k \|_1 \tag{IV.2}$$

$$\leq \max_j \sum_{k=1}^{m} \sum_{(i)} \tilde{P}_{ki_{2s-1}} \cdots \tilde{P}_{i_1 j} \| M_{i_1} \cdots M_{i_{2s-1}} M_k W_k \|_1 \tag{IV.3}$$

$$\leq \max_j \sum_{k=1}^{m} \sum_{(i)} \tilde{P}_{ki_{2s-1}} \tilde{P}_{i_{2s-1}i_{2s-2}} \cdots \tilde{P}_{i_1 j} \|W_k\|_1 \tag{IV.4}$$

$$\leq \max_j \sum_{k=1}^{m} \sum_{(i)} \tilde{P}_{ki_{2s-1}} \tilde{P}_{i_{2s-1}i_{2s-2}} \cdots \tilde{P}_{i_1 j} \|W\| = \|W\| \tag{IV.5}$$

We will next establish that $\|\Phi^{2s}(W)\| = \|W\|$ implies that $W = 0$, which will conclude the proof of the lemma. If $W \in \mathcal{S}^m$ satisfies $\|\Phi^{2s}(W)\| = \|W\|$ then all the previous inequalities (IV.2), (IV.3) and (IV.4) are actually equalities. This means that there exists some $j \in \{1, 2, \dots, m\}$ for which all the above maxima are attained at this j. For a $k \in \{1, 2, \dots, m\}$ and a sequence (i) of indices we denote

$$Q_{kj}((i)) = \tilde{P}_{ki_{2s-1}} \tilde{P}_{i_{2s-1}i_{2s-2}} \cdots \tilde{P}_{i_1 j}.$$

It follows from $P^{2s} = (P^s)^2 > 0$ that for each $k \in \{1, 2, \dots, m\}$ we have $\sum_{(i)} Q_{kj}((i)) > 0$. From this property together with

$$\sum_{k=1}^{m} \|W_k\|_1 \left(\sum_{(i)} Q_{kj}((i)) \right) = \max_k \|W_k\|_1 = \|W\|$$

and

$$\sum_{k=1}^{m} \sum_{(i)} Q_{kj}((i)) = 1$$

we can conclude that for every $k \in \{1, 2, \dots, m\}$

$$\|W_k\|_1 = \|W\|.$$

It follows from the equality

$$\sum_{k=1}^{m} \sum_{(i)} Q_{kj}((i)) \|M_{i_1} M_{i_2} \cdots M_{i_{2s-1}} M_k W_k\|_1 = \\ = \sum_{k=1}^{m} \sum_{(i)} Q_{kj}((i)) \|W_k\|_1,$$

that for every sequence (i) such that $Q_{kj}((i)) > 0$ the following holds:

$$\|M_{i_1} M_{i_2} \cdots M_{i_{2s-1}} M_k W_k\|_1 = \|W_k\|_1,$$

which in turn implies

$$\|M_{i_1} \cdots M_{i_{2s-1}} M_k W_k\|_1 = \|M_{i_2} \cdots M_{i_{2s-1}} M_k W_k\|_1 \tag{IV.6}$$

$$\cdots = \|M_{i_{2s-1}} M_k W_k\|_1 = \|M_k W_k\|_1 = \|W_k\|. \tag{IV.7}$$

Employing Lemma 2.3, we conclude that for all sequences (i) with $Q_{kj}((i)) > 0$:

$$W_k = M_k W_k = M_{i_{2s-1}} W_k = \cdots = M_{i_1} W_k. \tag{IV.8}$$

Recall now that for arbitrary k and l there exists a sequence (i) which contains l such that $Q_{kj}((i)) > 0$. Using (IV.8) we conclude that

$$M_l W_k = W_k, \quad \forall k, l \in \{1, 2, \dots, m\}, \tag{IV.9}$$

and the relations (IV.9) imply

$$W_1 = \cdots = W_m = 0. \tag{IV.10}$$

Indeed, for each $h \in \{1, 2, \dots, m\}$ there exists a matrix $M_l \in \mathcal{M}$ with positive h column(by Assumption (ii)) . Thus, Lemma 2.5 implies that the h coordinate of each W_k vanishes, and since h is arbitrary, (IV.10) follows. The proof of the lemma is complete. □

The following result may be established by using the same arguments as those employed in proving the previous lemma, and we will not repeat it here.

Lemma 4.2: If P is a primitive matrix, such that $P^s > 0$ for some positive integer s, then Ψ^{2s} is a contraction on $\mathcal{D}$.

Inspecting the proof of Theorem 2.7, we realize that we didn't use any special properties of the second power in deriving the proof while using the contractive property of Φ^2. Namely for any positive integer q, if Φ^q is contractive on Σ^m, then Theorem 2.7 follows. Similarly, if Ψ^q is contractive on $\mathcal{D}$ for some positive integer q then Theorem 3.5 follows. As a consequence of the previous two Lemmas, we have the following results.

Theorem 4.3: Let Assumption (ii) hold and suppose that the transition matrix P is primitive, so that there exists an integer $s \geq 1$ such that P^s has positive entries. Then the spectral radius of the restriction of Φ to $\mathcal{S}^m$ is smaller than

1. In particular there exists a unique solution $V^\star$ for equation $(II.5)$, and the iteration scheme

$$V^{(k+1)} = \Phi(V^{(k)}), k = 0, 1, 2, ...$$

with any starting point $V^{(0)} = V_0$ in Σ

$$\lim_{k\to\infty} V^{(k)} = V^\star. \tag{IV.11}$$

Moreover, the asymptotic behavior of the expectation of the random variable $W(N)$ is given by:

$$\lim_{N\to\infty} E(W(N)) = \sum_{i=1}^{m} \rho_i V_i^\star, \tag{IV.12}$$

where $V^\star = (V_1^\star, \dots, V_m^\star) \in \Sigma^m$ is the unique solution of $(II.5)$, and $\rho = (\rho_1, \dots, \rho_m)$ is the Perron eigenvector of the transition probability matrix (P_{ij}).

Theorem 4.4: Let Assumption (ii) hold, and assume that the transition probability matrix P is primitive. Then the spectral radius of the restriction of the mapping Ψ to $\mathcal{B}^m$ is smaller than 1. In particular there exists a unique solution $D^\star$ for equation $(III.5)$, and the iteration scheme

$$D^{(k+1)} = \Psi(D^{(k)}), k = 0, 1, 2, ...$$

with the starting point $D^{(0)} = D_0$ satisfies

$$\lim_{k\to\infty} D^{(k)} = D^\star$$

for every $D_0 \in \mathcal{D}$. Moreover, the asymptotic behavior of the variance

$$Var(W(N)) = E[W(N)W(N)^T] - E[W(N)]E[W(N)]^T$$

is given by:

$$\lim_{N\to\infty} Var(W(N)) =$$

$$= \sum_{i=1}^{m} \rho_i D_i^\star - \left(\sum_{i=1}^{m} \rho_i V_i^\star\right) \left(\sum_{i=1}^{m} \rho_i V_i^\star\right)^T \tag{IV.13}$$

where $D^\star = (D_1^\star, \dots, D_m^\star) \in \mathcal{D}$ is the unique solution of $(III.5)$, and $\rho = (\rho_1, \dots, \rho_m)$ is the Perron eigenvector of the transition matrix (P_{ij}).

V. R-MODEL

In the previous sections we considered the process $\{(A(k), W(k))\}_{k=0}^\infty$ under the assumption that $\{A(k)\}_{k=0}^\infty$ is a Markov process, and the distribution of $W(k+1)$ is determined by the value of $A(k)$ and the distribution of $W(k)$. In this model, the emphasis is put on the process $\{A(k)\}_{k=0}^\infty$, and $\{W(k)\}_{k=0}^\infty$ may be considered as a 'shadow' of it since the properties of $\{W(k)\}_{k=0}^\infty$ are derived from the distribution of $\{A(k)\}_{k=0}^\infty$.

However, one can construct a router such that the probability that a packet will be dropped at the k-th congestion event depends on the information provided by the vector $W(k)$, whose j-th coordinate is equal to the throughput of the j-th flow at the k-th congestion. We thus assume throughout this section that Assumption (iii) holds, and we will describe it again below.

When we consider the model under Assumption (iii), which we call *the R-model*, we assume that the value of $W(k)$ at the k-th congestion event, say $W(k) = w$, determines the distribution of $A(k)$. Namely, there exist continuous functions $w \mapsto p_i(w) \in \mathbb{R}^+$ on Σ such that

$$P[A(k) = M_i | W(k) = w] = p_i(w) (\forall 1 \le i \le m)(\forall w \in \Sigma) \tag{V.1}$$

and

$$\sum_{i=1}^{m} p_i(w) = 1 \quad \text{for every } w \in \Sigma.$$

In order to ensure that each flow have nonzero probability to detect a drop we assume that for each flow i there exist matrix in $\mathcal{M}$ with positive i-th column.

We begin this Section by proving that for any initial distribution of $W(0)$ almost all products $\{A(k)\cdots A(0)\}_{k\in N}$ are weakly ergodic. Recall that a sequence $\{Q_k\}_{k\in N}$ of column-stochastic matrices is called weakly ergodic if

$$\lim_{k\to\infty} \text{dist}(Q_k, \mathcal{R}) = 0,$$

where we denote by $\mathcal{R}$ set of rank-1 column stochastic matrices. For any column stochastic matrix Q, we know that $Q(\mathcal{S}) \subset \mathcal{S}$ (see the proof of Proposition 2.2). Thus the restriction of Q to $\mathcal{S}$ is a mapping to itself, and we denote this map by $\tilde{Q}$. It follows from the definition of weak ergodicity given above that the sequence $\{Q_k\}_{k\in N}$ is weakly ergodic if and only if

$$\lim_{k\to\infty} \tilde{Q}_k = 0, \quad \text{(see [7])}.$$

Recall also that a linear operator on a vector space V is called paracontractive with respect to norm $\|\cdot\|$ if for all $x \in V$

$$Vx \ne x \Rightarrow \|Vx\| < \|x\|.$$

The main tool in establishing almost sure weak ergodicity will be the following result which is given in [6]:

Theorem 5.1: Let $\|\cdot\|$ be a norm on $\mathbb{R}^m$ and let $\mathcal{F} \subset \mathbb{R}^{m\times m}$ be a finite set of linear operators which are paracontractive with respect to $\|\cdot\|$. Then for any sequence $\{A_k\}_{k\in N} \subset \mathcal{F}^N$, the sequence of left products $\{A_k A_{k-1} \dots A_1\}_{k\in N}$ converges.

Proposition 5.2: Let the random variable $W(0)$ have arbitrary distribution on Σ. Under Assumption (iii), the sequence of products $\{A(k)A(k-1)\cdots A(0)\}_{k\in N}$ is weakly ergodic with probability 1, i.e.

$$\lim_{k\to\infty} \tilde{A}(k)\tilde{A}(k-1)\cdots\tilde{A}(0) = 0 \text{ almost surely.} \tag{V.2}$$

Proof. From the assumption of positivity and the continuity of the mappings p_i on the compact set Σ it follows that

$$\eta = \inf\{p_i(s) |\ s \in \Sigma,\ i \in \{1, \dots, m\}\} > 0.$$

This means that $p(A(k) \ne M_i) \le 1 - \eta < 1$ for every $1 \le i \le m$. For every such i let T_i be a matrix with positive i-th column. Then with probability 1 the matrix T_i appears infinitely often in $\{A_k\}_{k=0}^\infty$. We will next establish that this implies that weak ergodicity holds for left products $\{A_k \dots A_1 A_0\}_{k\in N}$.

By Lemma (2.3), for all $M \in \mathcal{M}$, $\tilde{M}$ is paracontracting on $\mathcal{S}$ with respect to the L_1 norm. By Theorem 5.1 it follows that the sequence $\{\tilde{A}_k\tilde{A}_{k-1}\cdots\tilde{A}_0\}_{k\in N}$ is convergent, and we claim that the limit is zero. To show this let $s \in \mathcal{S}$. Then there exist a $y \in \mathcal{S}$ such that $y = \lim_{k\to\infty}\tilde{A}_k\tilde{A}_{k-1}\cdots\tilde{A}_0 s$. For any fixed i let $\{A_{n_k}\}_{k\in N}$ be a subsequence of $\{A_k\}_{k\in N}$ with $A_{n_k} = T_i$. Then

$$y = \lim_{k\to\infty}\tilde{A}_{n_k}\tilde{A}_{n_k-1}\cdots\tilde{A}_0 s = T_i \lim_{k\to\infty}\tilde{A}_{n_k-1}\cdots\tilde{A}_0 s = T_i y,$$

hence $T_i y = y$. But by Lemma 2.5 i-th coordinate of y must be zero. Since i is arbitrary, it follows that $y = 0$. We have thus established that $\lim_{k\to\infty}\tilde{A}(k)\tilde{A}(k-1)\cdots\tilde{A}(0) = 0$, which implies the assertion of the proposition. □

Comment : The previous proposition is also established in [24], under Assumption (i), i.e. when all the functions p_i are constant.

Note that in a sense, under Assumption (iii), the roles of $\{A(k)\}_{k=0}^{\infty}$ and $\{W(k)\}_{k=0}^{\infty}$ are interchanged compared to their roles in the model under Assumption (ii): the emphasis is put on $\{W(k)\}_{k=0}^{\infty}$, and $\{A(k)\}_{k=0}^{\infty}$ is considered as its shadow process.

We will henceforth restrict attention to stationary processes. The process $\{(A(k), W(k))\}_{k=0}^{\infty}$ is Markovian in the compact state space $\{1, 2, .., m\}\times\Sigma$, and we will next establish that it has a unique equilibrium distribution

$$\{\rho_1, ..., \rho_m\} \times (\lambda_1(dw), ..., \lambda_m(dw)).$$

Namely the probability that $A(k) = M_i$ and $W(k) \in U$ is equal, in the limit where $k \to \infty$, to $\rho_i\lambda_i(U)$. The equilibrium measure is defined on the set of limit points of $\{W(k)\}_{k=0}^{\infty}$, and for a prescribed $W_0 = s$ we denote by $F(s)$ the set of all limit points of sequences $\{W(k)\}_{k=0}^{\infty}$ with $W(0) = s$. We use the following terminology and say that *weak ergodicity holds for* $\{M_1, ..., M_m\} = \mathcal{M}$ if every product

$$A_k A_{k-1}\cdots A_0,\ A_k \in \mathcal{M} \text{ for every } k \ge 0$$

in which each that each M_i appears infinitely often is weakly ergodic.

Proposition 5.3: Suppose that $p_i(w) > 0$ for every $1 \le i \le m$ and $w \in \Sigma$, and assume that weak ergodicity holds for $\{M_1, .., M_m\}$. Then

$$F(s_1) = F(s_2) =: F \text{ for every } s_1, s_2 \in \Sigma. \tag{V.3}$$

Thus F is the smallest closed subset of Σ which is invariant under each $M_i \in \mathcal{M}$, so that it satisfies

$$F = \bigcup_{i=1}^{m} M_i(F),$$

and it is the support of the unique equilibrium invariant measure
$(\lambda_1(dw), ..., \lambda_m(dw))$.

Proof: For a prescribed starting point $W_0 = s$ we define the sequence of subsets $F_k(s) \subset \Sigma$ as follows:

$$F_0(s) = \{s\},\ F_{k+1}(s) = \bigcup_{i=1}^{m} M_i(F_k(s)),\ k = 0, 1, 2, ... \tag{V.4}$$

Then $F(s)$ may be expressed in the form

$$F(s) = \bigcap_{p=1}^{\infty}\left(\text{cl} \bigcup_{k=p}^{\infty} F_k(s)\right). \tag{V.5}$$

Denote by $h(\cdot,\cdot)$ the Hausdorff metric in Σ. It then follows from the weak ergodicity of $\mathcal{M}$ that

$$h(F_k(s_1), F_k(s_2)) \to 0 \text{ as } k \to \infty. \tag{V.6}$$

This follows from the fact that each point in $F_k(s)$ is of the form

$$M_{i_k}M_{i_{k-1}}...M_{i_1}s$$

for some matrices $M_{i_j} \in \mathcal{M}$, $1 \le j \le k$, and that

$$M_{i_k}M_{i_{k-1}}...M_{i_1}s_1 - M_{i_k}M_{i_{k-1}}...M_{i_1}s_2 \to 0 \text{ as } k \to \infty$$

by weak ergodicity. It follows from (V.6) that the Hausdorff distance between $\bigcup_{k=p}^{\infty} F_k(s_1)$ and $\bigcup_{k=p}^{\infty} F_k(s_2)$ is arbitrarily small provided p is sufficiently large. In view of (V.5) it follows that $F(s_1) = F(s_2)$ for every $s_1, s_2 \in \Sigma$, establishing (V.3) and concluding the proof the Proposition. □

The dynamics of $\{(A(k), W(k)\}_{k=0}^{\infty}$ can be described as follows. For a vector $W(k) = w$ at the instant of time k, choose a matrix $A(k)$ from $\mathcal{M}$ according to the distribution $\{p_i(w)\}_{i=1}^{m}$, and set $W(k+1) = A(k)w$. We then follow the steps

$$\begin{aligned} &W(0) \to A(0) \to W(1) = A(0)W(0) \to A(1) \to \cdots \\ &\to W(k) \to A(k) \to W(k+1) = A(k)W(k) \to \cdots \end{aligned} \tag{V.7}$$

We restrict attention only to the terms $W(k)$ in the chain of variables (V.7), and if $W(0) \sim (\lambda_1, ..., \lambda_m)$ then $\{W(k)\}_{k=0}^{\infty}$ turns out to be a stationary Markov process.

We now view the dynamics in a different manner, and this time we focus on the terms $A(k)$ in the above chain (V.7). If in the outset we restrict attention to stationary processes, then the distribution of each variable $W(k)$ is $(\lambda_1(dw), ..., \lambda_m(dw))$. Assuming this we restrict attention only to the variables $A(k)$ in (V.7), which turns out to be a stationary Markov chain in $\mathcal{M}$ provided that we take the initial distribution $A(0) \sim \rho$. We thus suppose that $W(k) \sim (\lambda_1, ..., \lambda_m)$, and that $A(k) = M_i$ for some $1 \le i \le m$. This determines the distribution of $W(k+1) = M_i W(k)$, as well as the distribution of $A(k+1)$. More explicitly we define $P_{ij} = Ep_j(M_i W(k))$ where E denotes the expectation operation with respect to the distribution $\lambda(dw)$, namely

$$P_{ij} = \int p_j(M_i w)\lambda_i(dw). \tag{V.8}$$

Although we defined $P_{ij} = Ep_j(M_i W(k))$, actually P_{ij} in (V.8) doesn't depend on k since all the variables $W(k)$ have the same distribution λ. However, we have to verify that our construction does yield this distribution to all $W(k)$. But indeed, since $\rho \times (\lambda_1, ..., \lambda_m)$ is an equilibrium distribution for the Markov process $\{(A(k), W(k))\}_{k=0}^{\infty}$, it follows that if we have $A(k) \sim \rho$, then the distribution of $W(k+1)$ is $(\lambda_1, ..., \lambda_m)$, and that of $A(k+1)$ is ρ. We summarize the above discussion as follows:

Theorem 5.4: The matrix P is a transition probability matrix of a stationary Markov chain $\{A(k)\}_{k=0}^{\infty}$ in $\mathcal{M}$ with stationary distribution $A(k) \sim \rho$. This Markov chain consist of the $A(k)$ terms in the process $\{(A(k), W(k))\}_{k=0}^{\infty}$ which describes the R-model of the process.

We are interested in the asymptotic behavior of $W(N)$ where $N \to \infty$, which in view of $W(k+1) = A(k)W(k)$ reduces to the study of the asymptotic distribution of the products

$$\left(\prod_{k=0}^{N} A(k)\right) W(0) \tag{V.9}$$

when $N \to \infty$. Since weak ergodicity holds for the matrix products $\Pi_{k=0}^{N} A(k)$ it follows that the asymptotic behavior of the expressions in (V.9) doesn't depend on $W(0)$ there, and we consider these expressions with an arbitrary choice of $W(0) \in \Sigma$. Although $\{A(k)\}_{k=0}^{\infty}$ is a stationary process, the process $\{A(N) \cdots A(0)W\}$ is not stationary, and we define

$$V_i^N = E\left[A(N-1)A(N-2)\cdots A(0)W(0)|A(N) = M_i\right], \tag{V.10}$$

where the expectation is with respect to the measure in which $\{A(k)\}_{k=0}^{\infty}$ is a stationary Markov chain with the transition probability matrix P in (V.8). Associated with this P is the matrix $\tilde{P}$ of backward probabilities, so that $\tilde{P}_{ij}$ is the probability of having $A(k) = M_i$ given that $A(k+1) = M_j$. Thus assuming $A(N+1) = M_j$, it follows from (V.10) that

$$V_j^{N+1} = E\left[A(N)A(N-1)\cdots A(0)W(0)|A(N+1) = M_j\right]$$

$$= \sum_{i=1}^{m} \tilde{P}_{ij} E\left[M_i A(N-1)\cdots A(0)W(0)|A(N) = M_i, A(N+1) = M_j\right]$$

$$= \sum_{i=1}^{m} \tilde{P}_{ij} M_i E\left[A(N-1)\cdots A(0)W(0)|A(N) = M_i,\right],$$

where in the last equality we have used the Markov property. Equating the first and last terms and using (V.10) we obtain the relations

$$V_j^{N+1} = \sum_{i=1}^{m} \tilde{P}_{ij} M_i V_i^N, \; N \geq 0. \tag{V.11}$$

But we observe now that (V.11) is an iterations scheme for the fixed point equation (II.5). Thus the results of the previous sections imply that the following limits exist

$$\lim_{N\to\infty} V_i^N = V_i^\star \text{ for every } 1 \leq i \leq m, \tag{V.12}$$

where $V^\star = (V_1^\star, ..., V_m^\star)$ is the unique solution of (II.5). As a consequence of this discussion we have the following:

Theorem 5.5: The conclusions of Theorems 4.3 and 4.4 hold true when we replace Assumption (ii) there by Assumption (iii).

VI. Conclusions

In this paper we consider the dynamics of AIMD-networks that evolve according to Markovian dynamics. We have shown that such networks have well defined stochastic equilibria and provided tools that can be used to characterise these equilibria. In particular, for routers that operate according to Assumption (ii), we have developed tools for computing $\lim_{k\to\infty} E(W(k))$ and $\lim_{k\to\infty} Var(W(k))$. We then extended these results to the R-model given by Assumption (iii).

While developing these tools represent an important first step in studying such networks, much work remains to be done. The results derived in this paper provide tools to address the problem of designing routers that achieve, in the long run, certain goals. By controlling the distribution of the random variable $A(0)$ in the i.i.d. case (i), or the transition matrix P in the Markov cases (ii) and (iii), one can guarantee that in the long run, the asymptotic expected value of $W(k)$ is close to a certain prescribed vector $V^\star$. A major objective of future work will be to investigate how this might in fact be achieved.

Another interesting designing problem that is of great practical interest, and which may be addressed in the setting provided by either Assumption (ii) or Assumption (iii), is the following. For a prescribed vector $V^\star$, consider all the transition matrices P for which

$$\lim_{k\to\infty} E(W(k)) = V^\star, \tag{VI.1}$$

and among them pick one for which

$$\lim_{k\to\infty} \text{Var}(W(k)) = T^\star \tag{VI.2}$$

is the smallest possible in a certain sense. Minimizing the variance makes it more likely that the desired long-run behavior expressed by $\lim_{k\to\infty} E(W(k)) = V^\star$ will be realized faithfully (although the cost of this choice may be a slow network convergence or some other undesirable network behaviour). This goal defines a constrained optimization problem which may be addressed either numerically or theoretically. We note that the minimization may be approached in various ways: either minimizing a certain functional, e.g. the trace of $T^\star$, or looking for a $T_0^\star$ such that

$$T_0^\star \leq T^\star$$

in the positive definite sense, where $T_0^\star$ and $T^\star$ correspond to certain matrices P_0 and P such that (VI.1) and (VI.2) hold. Finally, we note that one of the principal tools for analysing AIMD networks is the network simulator NS-2. For networks of low dimension, this tool is effective for examining the behaviour of AIMD networks. However, for networks with large number of sources this tool becomes increasingly difficult to use due to excessive simulation times. In this context, efficient methods to compute important network properties are likely of great value to network designers. The tools presented in this paper represent a first step toward the development of such tools.

REFERENCES

[1] G. Appenzeller, I. Keslassy and N. McKeown. "Sizing router buffers". Proceedings of ACM SIGCOMM '04, Portland, OR, USA, August/September 2004.

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

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

[4] A. Berman, R. Plemmons. "Nonnegative matrices in the mathematical sciences". SIAM, 1979.

[5] Y. Chait, C. V. Hollot, V. Misra, H. H., and Y. Halevi. "Dynamic analysis of congested TCP networks". Proceedings of American Control Conference, San Diego, CA, USA, June 1999.

[6] L. Elsner, I. Koltracht, M. Neumann. "On the convergence of asynchromous paracontractions with applications to tomographic reconstrucion from incomplete data". Linear Algebra Appl., vol. 130, pp. 65-82, 1990.

[7] D. Hartfiel. "Nonhomogeneous matrix products", World Scientific, 2002.

[8] C. V. Hollot and Y. Chait. "Non-linear stability analysis of a class of TCP/AQM networks". Proceedings of IEEE Conference on Decision and Control, Orlando, FL, USA, December 2001.

[9] V. Jacobson. "Congestion avoidance and control". Proceedings of ACM SIGCOMM '88, Stanford, CA, USA, August 1988.

[10] S.Jin, L. Guo, I. Matta, A. Bestavros. "A spectrum of TCP-friendly window-based congestion control algorithms". *IEEE/ACM Transactions on Networking*, vol. 11, no. 3, 341-355, 2003.

[11] 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.

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

[13] I. Khalifa, L. Trajkovic. "An overview and comparison of analytical TCP models". Proceedings of the 2004 International Symposium on Circuits and Systems, 2004. ISCAS '04, Volume: 5, pp. :469–472, 23-26 May 2004.

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

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

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

[18] J.R. Norris. "Markov Chains". Cambridge University Press, 1997.

[19] L. Qiu, Y. Zhang and S. Keshav. "Understanding the performance of many TCP flows". Computer Networks: Volume 37, Issue 3-4, pp. 277–306, 2001

[20] R.Shorten, D. Leith, J. Foy and R. Kilduff. "Analysis and design of synchronised communication networks". Automatica, vol. 41, pp. 725–730. 2005.

[21] R. Shorten, F. Wirth and D. Leith. "A positive systems model of TCP-like congestion control: Asymptotic results". Accepted by IEEE/ACM Transactions on Networking, 2004.

[22] R. Srikant. "Internet congestion control". Control theory, 14, Birkhäuser Boston Inc., Boston, MA, 2004.

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

[24] F. Wirth, R. Stanojevic, R. Shorten, D. Leith. "Stochastic equilibria of AIMD communication networks". Preprint.

[25] Y. Yang, S. Lam. "General AIMD Congestion Control". Proceedings of the IEEE International Conference on Network Protocols, Osaka, Japan, November 2000

[26] Z. Zhao, S. Darbha, A.L.N. Reddy. "A method for estimating the proportion of nonresponsive traffic at a router". *IEEE/ACM Transactions on Networking*, vol.12, no. 4, 708- 718, 2004.