AIMD, Fairness and Fractal Scaling of TCP Traffic

Francois Baccelli & Dohy Hong
INRIA-ENS, ENS, 45 rue d'Ulm 75005 Paris, France, {Francois.Baccelli, Dohy.Hong@ens.fr}

Abstract—**We propose a natural and simple model for the joint throughput evolution of a set of TCP sessions sharing a common tail drop bottleneck router, via products of random matrices. This model allows one to predict the fluctuations of the throughput of each session, as a function of the synchronization rate in the bottleneck router; several other and more refined properties of the protocol are analyzed such as the instantaneous imbalance between sessions, the autocorrelation function or the performance degradation due to synchronization of losses. When aggregating traffic obtained from this model, one obtains, for certain ranges of the parameters, short time scale statistical properties that are consistent with a fractal scaling similar to what was identified on real traces using wavelets.**

Keywords—**TCP, additive increase–multiplicative decrease algorithm, IP traffic, synchronization, fairness, throughput, product of random matrices, wavelet, fractal, autocorrelation.**

I. INTRODUCTION

The present paper proposes a simple model for the joint evolution of TCP sessions sharing a common tail drop bottleneck router.

Throughout the paper, we will refer to this model as to the *AIMD* (additive increase, multiplicative decrease) model. In this model, which is described in §II, TCP is not represented at packet level, but rather via simple fluid equations that describe the joint evolutions of throughput for the set of sessions sharing the router, and the loss process in the router. This fluid vision of the problem will be discussed in §II-B and will be validated by simulation. Other fluid representations of the AIMD dynamics have also been studied e.g. in [3], [6], [7]. These earlier papers only considered the extreme cases of full synchronization or no synchronization at all. The main novelty of the present paper consists in the introduction and the analysis of the notion of *synchronization rate*, which covers the whole spectrum between these two extreme cases.

Section III which is the basis of the analysis in the rest of the paper, concentrates on the characterization of the steady state of the AIMD model. This allows us to calculate not only the mean value of the throughput obtained by each session, but also quantities that can only be obtained from a model that captures the interaction between several sessions, such as the covariance matrix or the autocorrelation matrix of the throughput obtained by all sessions. All these quantities are obtained in function of the synchronization rate in the bottleneck router. This synchronization rate, which is the basic data for the AIMD model, is defined as the proportion of sessions that experience a loss at a congestion epoch of the bottleneck router (see π and p below).

The analytical results on TCP throughput are gathered in §IV.

- Subsection IV-A shows that the throughput formulas obtained by this approach are compatible with, and actually refine, the classical relationship between the packet loss probability and TCP throughput for a given session (see e.g. [17], [18]). The main new result obtained here is the expression of the unspecified constant that shows up in this formula as a function of the synchronization rate (Formulas (23), (26) and (27)).
- Subsection IV-B concentrates on the evaluation of the performance degradation due to synchronization of losses when taking the buffer size of the bottleneck router into account. The condition under which synchronization affects mean throughput is established (see (28)), as well as functional dependency of the under-utilization w.r.t. the synchronization rate (29).
- The imbalance between the sessions is studied in §IV-C. For this, we introduce the notion of *dispersion ratio* and that of *instantaneous fairness*. We show that even under homogeneous assumptions, where the long terms averages of throughput are the same for all sessions (the so called fairness of TCP), the stationary solution of the AIMD model is such that there is a high dispersion of throughputs for different sessions at any given time. By this, we mean that under the fluid approximation of the AIMD model, the ratio of the fastest to the slowest session at any given time is a random variable with a heavy tail. In other words, what is usually called fairness (and which should be called fairness in mean) may be compatible with serious instantaneous unfairness.

Section V studies statistical properties of traffic aggregates. It concentrates on the short time scale fractal properties of TCP traffic which were first identified by Riedy and Lévy-Vehel [19] and studied using wavelet statistical tools by e.g. Feldmann, Gilbert, Willinger et al. in [12], [11]. By short time scales, we mean the same as in the above papers. In particular, this concerns time scales where the effect of the superposition of a large number of on-off sessions with Pareto *on* periods cannot be invoked to explain a non trivial scaling. In our model, there are no heavy tails in the primitive data at all: the number of active sessions *does not vary* and each session is long lived (i.e. has no *off* periods) and losses are assumed to be i.i.d. It is of course possible to add such variations to our model, and also to replace the fluid model by a packet level model taking slow start into account; however, this will not be done in the present paper for the following reason: all this can only increase variability and reinforce the properties observed in our simplified model. In other words, the main conclusions on high variability should hold true (at least on a qualitative basis) for such refined models since they hold true for the less variable basic AIMD model.

II. THE AIMD MODEL

In this section, we propose a set of fluid evolution equations allowing one to represent the key features of the AIMD mechanism for N homogeneous TCP sessions sharing one tail drop bottleneck router. We will first consider the homogeneous case, where all sessions have the same RTT. We will also give the equations for the heterogeneous case (different RTT's or synchronization rates). These equations will not be studied in the present paper.

By definition, the n-th *congestion time* is the n-th epoch at which a loss or several simultaneous losses occur on this shared

router. We use the following notation:
• N is the number of TCP sessions, which we assume to be constant with time;
• C is the capacity of the bottleneck router;
• $X_n^{(i)}$ is the throughput of session i just after the n-th congestion time;
• $W_n^{(i)}$ is the window size of session i just after the n-th congestion time;
• T_n is n-th congestion time, $\tau_{n+1} = T_{n+1} - T_n$;
• $R^{(i)}$ is the mean RTT of session i (in the homogeneous situation, $R^{(i)} = R$);
• $a_n^{(i)}$ is a 0,1 valued random variable with value 1 if session i experiences a loss at the n-th congestion time, 0 otherwise;
• $p^{(i)}$ is the synchronization rate for session i: $p^{(i)} = \mathbb{E}a_n^{(i)}$ (in the homogeneous situation, $p^{(i)} = p$);
• $\alpha^{(i)}$ is the linear growth rate of the window size of session i with time; it makes sense to take $\alpha^{(i)} = 1/R^{(i)}$;
• β is the multiplicative reduction of window size in case of loss (1/2 by default).

Most of the mathematical derivations of the present paper focus on the case of independent losses which is that where
1. the random variables $\{a_n^{(i)}, i = 1, \ldots, N\}$ are independent in n;
2. for all n, the random variables $a_n^{(i)}, i = 1, \ldots, N$ are generated independently, with $\mathbb{P}(a_0^{(i)} = 1) = \pi^{(i)}$, but only the samples such that at least one of them is 1 are kept (by definition, there is at least one loss at a congestion epoch). The resulting conditional law is easy to compute. For instance, in the homogeneous case,

$$\mathbb{P}(a_0^{(i)} = 1) \quad = \quad p = \frac{\pi}{1-(1-\pi)^N}$$

and for $i \neq j$,

$$\begin{aligned}
\mathbb{P}(a_0^{(i)} = 1, a_0^{(j)} = 1) &= \frac{\pi^2}{1-(1-\pi)^N} \\
\mathbb{P}(a_0^{(i)} = 0, a_0^{(j)} = 1) &= \frac{\pi(1-\pi)}{1-(1-\pi)^N} \\
\mathbb{P}(a_0^{(i)} = 0, a_0^{(j)} = 0) &= \frac{(1-\pi)^2\left(1-(1-\pi)^{N-2}\right)}{1-(1-\pi)^N}.
\end{aligned}$$

The parameter p will be referred to as the synchronization rate in what follows.

Other and more general loss models (including the case of throughput dependent losses) will also be considered by simulation in §V.

A. Homogeneous case

In our model, we assume that the throughput and the window size are linked by a Little like law: $W_n^{(i)} = X_n^{(i)}R$, which is a first simplifying assumption (this way of linking throughput to window is heuristic; Little's law would only apply to stationary means whereas we use it for linking instantaneous values here). Then the evolution of the throughput is given by:

$$X_{n+1}^{(i)} = ((1-a_{n+1}^{(i)}) + \beta a_{n+1}^{(i)})\left(X_n^{(i)} + \frac{\alpha}{R}\tau_{n+1}\right). \qquad (1)$$

We now use the fact that losses occur as soon as the router capacity is reached to get the following relation between the $X_n^{(i)}$'s and τ_{n+1}:

$$\sum_{i=1}^{N} X_n^{(i)} + \frac{\alpha}{R}\tau_{n+1}N = C. \qquad (2)$$

At this stage we assume that the buffer capacity of the router is 0 or negligible (see §II-B for a simple way to relax this assumption). So when using the notation

$$\gamma_n^{(i)} = (1-a_n^{(i)}) + \beta a_n^{(i)},$$

we get

$$\begin{aligned}
X_{n+1}^{(i)} &= \gamma_{n+1}^{(i)}\left(X_n^{(i)} + \frac{C}{N} - \frac{1}{N}\sum_{j=1}^{N} X_n^{(j)}\right) \\
&= \gamma_{n+1}^{(i)}\left(\frac{C}{N} + (1-\frac{1}{N})X_n^{(i)} - \frac{1}{N}\sum_{j\neq i} X_n^{(j)}\right) \quad (3)
\end{aligned}$$

Let X_n be the N-dimensional vector with coordinates $X_n^{(i)}$, B_n that with coordinates $\gamma_n^{(i)}C/N$ and let A_n be the $N \times N$ random matrix with diagonal terms $(A_n)_{i,i} = \gamma_n^{(i)}(1-1/N)$ and with non-diagonal terms $(A_n)_{i,j} = -\gamma_n^{(i)}/N$. Then we have

$$X_{n+1} = A_{n+1}X_n + B_{n+1}. \qquad (4)$$

B. Some preliminary remarks on the model

A first natural question concerns the behavior of the X_n process in the neighborhood of the boundaries of its state space, and in particular when one throughput becomes very small, or when the sum of the throughputs becomes close to C, where the fluid and the discrete models seem to differ significantly at first glance.

• *Buffer Capacity.* In the AIMD model, when the sum of all throughputs reaches C, losses occur. This is of course unrealistic in that the buffer size should be taken into account. Let B be the buffer size of the shared router. We show that the above equations can be adapted to take buffering into account, and that the adaptation simply consists in replacing C by $C_n = C + x_n$, with $x_n \leq \sqrt{2BN}/R$. The justification is the following: if at time 0 the total arrival rate is C and the buffered fluid amount is 0, then at time t the total arrival rate is $C + Nt/R^2$ and the buffer contents $Nt^2/2R^2$. Thus, when the buffer size B is reached, the total arrival rate is equal to $C + \sqrt{2BN}/R$. At the time the total arrival rate reaches C, if the buffered fluid amount is larger than 0, then we get $x_n < \sqrt{2BN}/R$.

• *Small Throughputs and Timeouts.* Windows cannot be halved an unbounded number of times, be it only because they always remain larger than or equal to 1, whereas, in our model, one specific coordinate of X can actually be halved an unbounded number of times. We argue that although windows cannot be halved infinitely often indeed, throughputs (which are the actual state variables of the AIMD model) can very well become "multiplicatively small" in case of repeated timeouts. In such a case the doubling of the RTO variable (see e.g. [22]) has the very same qualitative effect on session inter-packet times and therefore throughput as the one of the AIMD model. This observation is not new as the key role played by the exponential back-off of the RTO variable in the statistical behavior of a single TCP session was already demonstrated by L.Guo, M. Crovella and I. Matta in [15]. So in spite of the fact that timeouts are apparently not present in the AIMD model, one can argue that their presence is actually taken into account via the behavior of the model in the neighborhood of the corresponding boundary.

- *Large Throughputs.* Throughputs are often limited by the presence of a maximal window size; this will not be taken into account here.

It is clear that such a fluid model cannot take into account fine packet level aspects of TCP. This is intrinsic; models that mix this high level fluid evolution of the window sizes and packet level as e.g. in [5] can of course be contemplated, but seem out of reach for the time being. Nevertheless as shown by the simulations reported in the paper, the AIMD model captures key phenomena present in the protocol, and its behavior close to the boundary of the state space is certainly not a very accurate representation of reality, but is qualitatively correct.

C. Non homogeneous case

The generalization to the heterogeneous case gives the following equations:

- Little's law (in the heuristic sense): $W_n^{(i)} = X_n^{(i)} R^{(i)}$.
- Throughput evolution: $X_{n+1}^{(i)} = \gamma_{n+1}^{(i)} \left(X_n^{(i)} + \frac{\alpha^{(i)}}{R^{(i)}} \tau_{n+1} \right)$.
- Relation between the $X_n^{(i)}$'s and τ_{n+1}:

$$\sum_{i=1}^{N} X_n^{(i)} + \tau_{n+1} \sum_{i=1}^{N} \frac{\alpha^{(i)}}{R^{(i)}} = C. \tag{5}$$

- Let $\rho^{(i)} = \frac{\frac{\alpha^{(i)}}{R^{(i)}}}{\sum_j \frac{\alpha^{(j)}}{R^{(j)}}}$. The throughputs dynamics is:

$$X_{n+1}^{(i)} = \gamma_{n+1}^{(i)} \left(\rho^{(i)} C + (1-\rho^{(i)}) X_n^{(i)} - \rho^{(i)} \sum_{j \neq i} X_n^{(j)} \right). \tag{6}$$

Let B_n be the random vector with coordinates $\gamma_n^{(i)} \rho^{(i)} C$ and let A_n be the $N \times N$ random matrix with diagonal terms $(A_n)_{i,i} = \gamma_n^{(i)}(1-\rho^{(i)})$ and with non-diagonal terms $(A_n)_{i,j} = -\gamma_n^{(i)} \rho^{(i)}$. Then we again have $X_{n+1} = A_{n+1} X_n + B_{n+1}$.

III. Steady State Solution

We are interested in the steady state solutions of (4). For certain mean values, there is no need of detailed analysis. For instance, consider the homogeneous case, and assume that there exists a stationary regime X such that $\mathbb{E}[X^{(i)}]$ is finite and the same for all i. Since each line of A_n sums up to 0, when taking expectation on both sides of (4), we get $\mathbb{E}[X^{(i)}] = \mathbb{E}[B^{(i)}]$, for all i, that is

$$\mathbb{E}[X^{(i)}] = \frac{C}{N} \left(1 - (1-\beta)p\right).$$

The term $(1-\beta)p$ measures the loss of performance due to the idleness of the router. We will evaluate this more precisely in §IV-B. In particular, if $\beta = 1/2$, then for all i,

$$\mathbb{E}[X^{(i)}] = \frac{C}{N} \left(1 - \frac{p}{2}\right). \tag{7}$$

When taking $\alpha = 1/R$ (which will be assumed to hold in what follows), immediate calculations then give the following expression for the mean stationary inter congestion time:

$$\mathbb{E}[\tau] = \frac{CR^2 p}{2N} = \frac{CR^2 \pi}{2N(1-(1-\pi)^N)}. \tag{8}$$

Further characteristics require more sophisticated tools. For all random vectors X, we will denote by $||X||_2$ its L_2 norm: $||X||_2 = \sqrt{\mathbb{E}(X^t X)}$, and $\mathbb{C}(X)$ its covariance matrix: $\mathbb{C}(X) = \mathbb{E}((X - \mathbb{E}X)(X - \mathbb{E}X)^t)$ where X^t is the transpose of X.

Theorem 1 *In the homogeneous case, there exists a unique non negative and finite stationary solution $\{\mathcal{X}_n\}$, $n \in \mathbb{Z}$, to (4), with*

$$\begin{aligned} \mathcal{X}_n &= B_n + A_n B_{n-1} + A_n A_{n-1} B_{n-2} \\ &\quad + A_n A_{n-1} A_{n-2} B_{n-3} + \cdots, \qquad n \in \mathbb{Z}, \end{aligned} \tag{9}$$

where we have used the fact that the sequences of i.i.d. matrices and vectors A_n and B_n could be continued to negative values of n.

The covariance matrix of the stationary state vector $\mathcal{X} = \mathcal{X}_0$ is given by the following formula:

$$\begin{aligned} \mathbb{C}(\mathcal{X}) &= \frac{C^2}{N^2} \left(\left(\frac{a-b}{1-a+\rho(a-b)} \right) \mathbb{I} \right. \\ &\quad \left. + \left((b-c) - \frac{b\rho(a-b)}{1-a+\rho(a-b)} \right) \mathbb{J} \right), \end{aligned} \tag{10}$$

where $\mathbb{I}$ is the identity matrix and $\mathbb{J}$ the matrix with all its entries equal to 1; here $\rho = 1/N$ and

$$a = \mathbb{E}[(\gamma_0^{(1)})^2] = \frac{(1-3\pi/4) - (1-\pi)^N}{1-(1-\pi)^N} \tag{11}$$

$$b = \mathbb{E}[\gamma_0^{(1)} \gamma_0^{(2)}] = \frac{(1-\pi/2)^2 - (1-\pi)^N}{1-(1-\pi)^N} \tag{12}$$

$$c = \overline{\gamma}^2 = \mathbb{E}[\gamma^{(1)}]^2 = \left(1 - \frac{\pi/2}{1-(1-\pi)^N}\right)^2. \tag{13}$$

When N grows large and $\pi < 1$,

$$\mathbb{C}(\mathcal{X}) \sim \frac{C^2}{N^2} \left(\frac{1-\pi}{3} \mathbb{I} - \frac{(1-\pi)(1-\pi/2)^2}{3N} \mathbb{J} \right). \tag{14}$$

The proof is given in Appendix VI-A.

Remark 1 *Notice that $a - c = var(\gamma_0^{(i)})$ and that $b - c = cov(\gamma_0^{(i)}, \gamma_0^{(j)})$, $i \neq j$. Simple calculations show that $b - c < 0$ and that $0 < b < a < 1$. We can rephrase the results of Theorem 1 as follows: in the homogeneous case,*

$$var(\mathcal{X}^{(i)}) = \frac{C^2}{N^2} K_1, \tag{15}$$

with

$$K_1 = b - c + \frac{(a-b)(1-\rho b)}{1-a+\rho(a-b)} \quad (16)$$

and

$$cov(\mathcal{X}^{(i)}, \mathcal{X}^{(j)}) = \frac{C^2}{N^2} K_2, \quad (17)$$

with

$$K_2 = (b-c) - \frac{b\rho(a-b)}{1-a+\rho(a-b)}. \quad (18)$$

So $\mathrm{cov}(\mathcal{X}^{(i)}, \mathcal{X}^{(j)}) \leq 0$. *It is not so surprising to find that two different coordinates of* $\mathcal{X}$ *are negatively correlated as the corresponding sessions compete for capacity.*

Remark 2 *From (2) and Theorem 1, we immediately get the following expression for the variance of the stationary inter-congestion times:*

$$\begin{aligned} \mathrm{v}ar(\tau) &= \frac{R^4}{N^2} \mathrm{v}ar(\sum_i \mathcal{X}^{(i)}) \quad (19) \\ &= \frac{R^4}{N^2} (N \mathrm{v}ar(\mathcal{X}^{(1)}) + N(N-1) \mathrm{c}ov(\mathcal{X}^{(1)}, \mathcal{X}^{(2)})). \end{aligned}$$

We now focus on the autocorrelation function. For all random vectors X, Y, we will denote by $\mathbb{C}(X,Y)$ the covariance matrix:

$$\mathbb{C}(X,Y) = \mathbb{E}\left((X - \mathbb{E}X)(Y - \mathbb{E}Y)^t\right). \quad (20)$$

Using techniques similar to the ones of Theorem 1, one easily gets the following theorem.

Theorem 2 *In the homogeneous case, the autocorrelation matrices of the* $\mathcal{X}_n$ *vectors are given by the following formula (valid for* $k \geq 1$*):*

$$\begin{aligned} \mathbb{C}(\mathcal{X}_0, \mathcal{X}_k) &= \overline{\gamma}^k \left(\mathrm{v}ar(\mathcal{X}^1) - cov(\mathcal{X}^1, \mathcal{X}^2)\right)(\mathbb{I} - \rho\mathbb{J}) \\ &= \overline{\gamma}^k \frac{C^2}{N^2} \frac{a-b}{1-a+\rho(a-b)} (\mathbb{I} - \rho\mathbb{J}). \quad (21) \end{aligned}$$

So there is no LRD in this setting. This is consistent with the fact that in TCP traffic, LRD is primarily due to the on-off structure of HTTP sessions and the fact that *off* periods are heavy tailed; here all sessions are long lived and no such on-off structure is present.

Elementary calculations based on the preceding result show that the autocorrelation of the inter-congestion times is zero. We see no obvious explanation for this as inter-congestion times are not independent.

Notice that other autocorrelation functions of interest such as the inter-loss times seen by a given source (see [2] for a nice application) could also be approached via this method.

IV. Analytical Properties of TCP Throughput

A. Refinement of the square root throughput formula

Let us first show how the throughput formula obtained by this approach is actually a refinement to the classical throughput formulas for TCP throughput [17], [18]. We will concentrate on the homogeneous case, although the non-homogeneous case can be handled in the same way.

From the cycle formula (see [4], Chapter 1), the stationary, continuous time, mean value of the throughput of a session, which we will denote by x in what follows, is given by the formula :

$$x = \frac{\mathbb{E}\left[\int_0^\tau (\mathcal{X} + s/R^2) ds\right]}{\mathbb{E}[\tau]} = \frac{\mathbb{E}[\tau\mathcal{X}] + \mathbb{E}\left[\tau^2\right]/(2R^2)}{\mathbb{E}[\tau]}, \quad (22)$$

where $\mathcal{X}$ is a coordinate of the stationary throughput vector at a loss time and τ the next inter-congestion time. Simple calculations give

$$\mathbb{E}[\tau\mathcal{X}] = \frac{R^2}{N}\Big(C\mathbb{E}[\mathcal{X}] - (N-1)\mathbb{E}[\mathcal{X}\mathcal{X}'] - \mathbb{E}[\mathcal{X}^2]\Big),$$

where $\mathcal{X}$ and $\mathcal{X}'$ are two different coordinates of the stationary throughput vector at congestion epochs. Similarly

$$\begin{aligned} \frac{1}{2R^2}\mathbb{E}\left[\tau^2\right] &= \frac{R^2}{2N^2}\Big(C^2 - 2CN\mathbb{E}[\mathcal{X}] + N\mathbb{E}[\mathcal{X}^2] \\ &\qquad + N(N-1)\mathbb{E}[\mathcal{X}\mathcal{X}']\Big). \end{aligned}$$

So the numerator of the RHS of (22) is equal to

$$\begin{aligned} &\frac{R^2}{2N^2}\Big(C^2 - N^2\mathbb{E}[\mathcal{X}\mathcal{X}'] + N(\mathbb{E}[\mathcal{X}\mathcal{X}'] - \mathbb{E}[\mathcal{X}^2])\Big) \\ &= \frac{C^2R^2}{2N^2} p f(p,N), \end{aligned}$$

where the function

$$f(p,N) = 1 - \frac{p}{4} - K_2/p + \frac{1}{Np}(K_2 - K_1). \quad (23)$$

is obtained from (15) and (17).

Using now the expression (8) for $\mathbb{E}[\tau]$, we get

$$x = \frac{C}{N} f(p,N). \quad (24)$$

The mean number of packets sent by a session over a cycle is

$$\mathbb{E}\left[\int_0^\tau (\mathcal{X} + s/R^2) ds\right] = x\mathbb{E}[\tau].$$

So the packet loss probability p_{loss} is given by

$$p_{loss} = \frac{p}{x\mathbb{E}[\tau]} = \frac{2N^2}{C^2R^2 f(p,N)}. \quad (25)$$

Therefore, from (24) and (25),

$$x = \frac{\sqrt{2}\sqrt{f(p,N)}}{R\sqrt{p_{loss}}}. \quad (26)$$

Notice that when $N \to \infty$, $f(p,N) \sim 1 - p/4$, so that

$$x \sim \frac{\sqrt{2-\frac{p}{2}}}{R\sqrt{p_{loss}}}. \tag{27}$$

We conclude from (26) that our framework leads to formulas compatible with the classical estimates for TCP throughput ([17], [18]), and also that the unspecified constant that shows up in these formulas can actually be estimated from the synchronization rate via Formulas (23) or (27). The AIMD model predicts that this constant should be between $\sqrt{3/2}$ and $\sqrt{2}$, at least when the congestion avoidance phase is dominant. This confirms experimental results in [17].

B. Performance loss by synchronization

We will say that losses are synchronized when $a_n^{(i)} = 1$ with probability 1 for all n and i. We will say that losses are p-synchronized, with p a real close to 1 but less than 1, when $a_n^{(i)} = 1$ with probability p for all n and i.

Assume first that losses are synchronized, and all RTT's are the same. At T_n, all sessions halve their window. Let $S(t)$ denote the sum of the send rates of all sessions at time t. Let B be the buffer size of the shared router. Let $Q(t)$ be the total number of customers in the buffer at time t. In this model with non-zero buffer, we have (cf. §II-B)

$$S(T_n^-) = C + x \leq C + \frac{\sqrt{2BN}}{R}$$

where $0 \leq x < C$ (if $x \geq C$, some sessions will experience multiple window decreases).
In view of the AI rule, while $S(t) \leq C$,

$$S(T_n + t) = \frac{C+x}{2} + \frac{N}{R^2}t$$

and (with the notation $a^+ = \max(a, 0)$),

$$\begin{aligned} Q(T_n + t) &= \left(B + \int_0^t \left(\frac{-C+x}{2} + \frac{N}{R^2}v\right)dv\right)^+ \\ &= \left(B + \frac{-C+x}{2}t + \frac{N}{2R^2}t^2\right)^+. \end{aligned}$$

The minimum of the function

$$t \to B + \frac{-C+x}{2}t + \frac{N}{2R^2}t^2$$

is reached for $t^* = \frac{(C-x)R^2}{2N}$, and the value of the function at this minimum is

$$B - \frac{(C-x)^2R^2}{4N} + \frac{(C-x)^2R^2}{8N} = B - \frac{(C-x)^2R^2}{8N}.$$

So we have under-utilization of the link capacity iff

$$B < \frac{(C-x)^2R^2}{8N} \tag{28}$$

or equivalently $\frac{3\sqrt{2BN}}{CR} < 1$.

If this condition is satisfied, the under-utilization can be evaluated as one minus the ratio between the total number of packets sent and the maximum number of packets that can be sent during a period $T = T_n - T_{n-1}$ which is here constant. Let us call $\mathbb{U}$ this ratio. If ${}^*t$ is the first time at which $Q(T_n + t) = 0$, we have

$${}^*t = t^* - \frac{R^2}{N}\sqrt{\frac{(C-x)^2}{4} - \frac{2BN}{R^2}}.$$

Then

$$\mathbb{U} = \frac{1}{CT}\left(C(t^* - {}^* t) - \int_{{}^*t}^{t^*}\left(\frac{C+x}{2} + \frac{Nu}{R^2}\right)du\right).$$

After simplification, we have:

$$\mathbb{U} = 0.25\left(1 - \frac{3\sqrt{2BN}}{CR}\right). \tag{29}$$

For the case of p-synchronization, one obtains the following expression for under-utilization when N large:

$$\mathbb{U} \sim \frac{p}{4}\left(1 - \frac{4-p}{p}\frac{\sqrt{2BN}}{CR}\right)^+. \tag{30}$$

So, if $\frac{4-p}{p}\frac{\sqrt{2BN}}{CR} < 1$, then there is some underutilization.

C. Instantaneous fairness

Roughly speaking, the *instantaneous unfairness* of TCP should be a measure of the dispersion of the stationary throughputs at a given time, in the stationary regime. This dispersion is completely characterized by the joint distribution of the coordinates of $\mathcal{X}_1$ defined in (9). Let us see how to capture this notion in more compact ways.

For H a fixed subset of $\{1, \ldots, N\}$, we denote

$$T^H = \sum_{i \in H} \mathcal{X}^{(i)},$$

the aggregate of stationary throughputs over set H. We define the *dispersion ratio* for aggregates of size j, where $1 \leq j \leq N$, as the random variable

$$d_j = \frac{\max_{H,|H|=j} T^H}{\min_{H,|H|=j} T^H}.$$

In case of a sharing of the bandwidth that would be fair at any instant (e.g. the processor sharing queue), this ratio should be equal to 1. In contrast with this situation we have:

Theorem 3 *For all integers $j \leq N/2$, for x large,*

$$\mathbb{P}(d_j > x) \geq D' x^{-\xi}, \tag{31}$$

with D' a positive constant and

$$\xi = \frac{\ln(\mu)}{\ln(\frac{j+N}{2N})}, \quad \mu = \frac{\pi^j(1-\pi)^{N-j}}{1-(1-\pi)^N}. \tag{32}$$

The proof is given in Appendix VI-B.

So, in spite of the *fairness in expectation* of Equation (7) (i.e. in the steady state, all sessions have the same mean throughput), we have a rather severe *instantaneous unfairness* in that at any given time, the ratio of best throughput to worst is actually a heavy tailed random variable.

Note that since the numerator is bounded by C, this ratio is heavy tailed simply due to the fact that its denominator can be close to zero with a large enough probability (for instance, if V is an exponential random variable, $1/V$ is heavy tailed).

The log-log complementary plot of the distribution function of the dispersion ratio d_1 (which is also the ratio of max throughput over min) as obtained from simulations of the AIMD model, is given in Figure 1. Here $N = 100$ and $p = 0.01$. This ratio is invariant w.r.t. the values of C and R. The slope of the decreasing part of this curve is compatible with the lower bound estimate of (32).

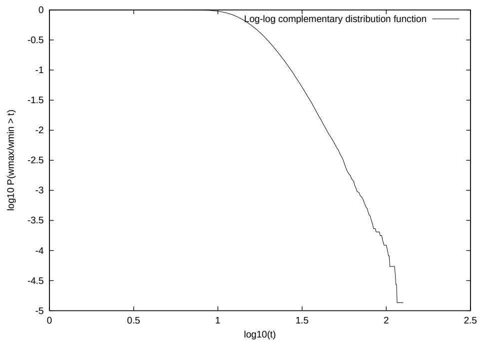

Fig. 1. Log-log complementary distribution function of $\mathcal{X}_{\max}/\mathcal{X}_{\min}$ in the stationary AIMD model.

In order to check that this effect is not simply an artefact stemming from the model, we have used NS to evaluate the dispersion ratio under the more accurate packet level description of TCP offered by NS. Figure 2 gives the log-log complementary plot of the distribution function of the dispersion ratio d_1 as obtained via the NS simulation of 100 TCP sessions sharing a 10 000 packets/s router. The router is FIFO and with a buffer of 70 packets; the minimal value of RTT is .1 seconds. One should not try to compare this curve with that of Figure 1 since the two simulated systems have no direct relation (in general, it is not possible to impose a given synchronization rate at congestion epochs via NS). Note however that there is also a linear part over several scales, similar to that of the AIMD model case. We conclude from this that this high dispersion is also present in NS simulations.

In the heterogeneous case, similar results can be obtained along the same lines. Of course throughput dispersion increases even more whenever RTT's themselves have more dispersion. So the heavy tailed dispersion of throughputs which is already present in the homogeneous case is reinforced in the non-homogeneous case.

Note that the proof of the heavy tailedness of dispersion extends directly to throughput-dependent synchronization rates (by this, we mean that for each session, π is a function of the current throughput of this session), provided these functions are bounded away from 0.

We would like to stress that the presence of heavy tailed distri-

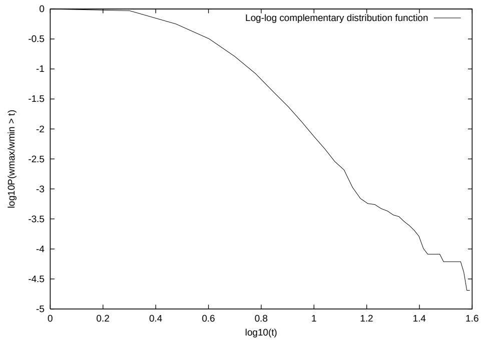

Fig. 2. Log-log complementary distribution function of $\mathcal{X}_{\max}/\mathcal{X}_{\min}$ via NS.

butions for the stationary distribution of certain affine dynamical systems based on random matrices with light tails (such as (4) is not new. The first results along these lines are due to Kesten [16]. A survey on the matter can also be found in [9])

V. Fractal Scaling

In [19] and [14], the fractal scaling of certain real TCP traces was studied. In this section we show that stationary traces generated by the AIMD model exhibit four key properties identified on real traces:

1. Time aggregation leads to trajectories with important fluctuations over several scales;
2. The wavelet energy function plot yields a linear region starting at the lowest scales (high frequencies), and with a finite upper cutoff;
3. The wavelet partition sum exhibits linearity characteristic of a fractal.
4. The multiscale diagram is non linear.

All these observations are compatible with a fractal scaling of the trajectories of the AIMD process; the last property is compatible with a multifractal scaling (see e.g. [1], p. xxii).

The methods used here are the same as those used in [14] and [1]. Some of these methods require Gaussian assumptions (see e.g. [1]), This seems justified in the following framework where we focus on aggregates of K sessions (typically we take $K = \phi N$ sessions with $0 < \phi < 1$ and N large).

A. *Some simulations*

A.1 Aggregation of AIMD traffic

Figure 3 shows throughput traces generated from the AIMD model ($N = 100$, $p = 0.01$). We observe the sum of the throughputs of $Obs = 10$ sessions on 4 different time scales: the 4 first curves are generated from the sum of $Obs = 10$ session throughputs sampled every 0.01 s and by aggregation of these over 0.1, 1 and 10 seconds respectively.

This is to be compared to Figure 5 of [21]. Fluctuations remain quite important on each of these time scales.

A.2 Wavelet analysis

Given a time series $\{X_{n,k}, k = 1, \ldots, 2^n\}$, its discrete (Haar) wavelets coefficients $d_{n,k}$ are constructed as in [1],[11]. The

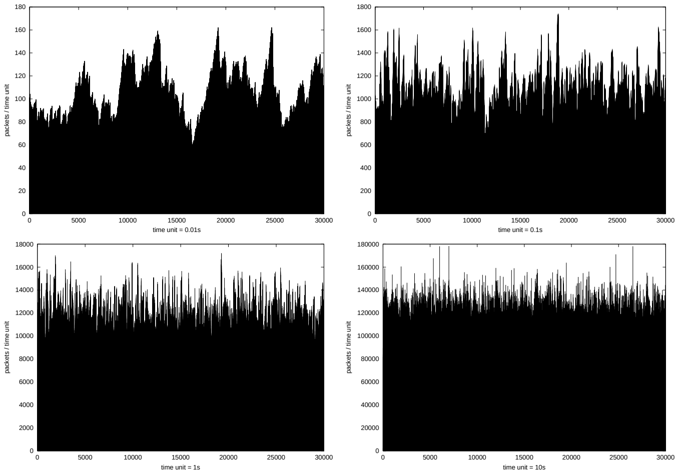

Fig. 3. Traces generated from the AIMD model.

average energy at scale j is then defined as:

$$E_j = \frac{1}{2^j} \sum_{k=1}^{2^j} |d_{j,k}|^2.$$

In Figure 4 we plot the evolution of the sum of the throughputs of Obs= 10, 100, 450 sessions in the case $N = 500$, $p = 0.1$.

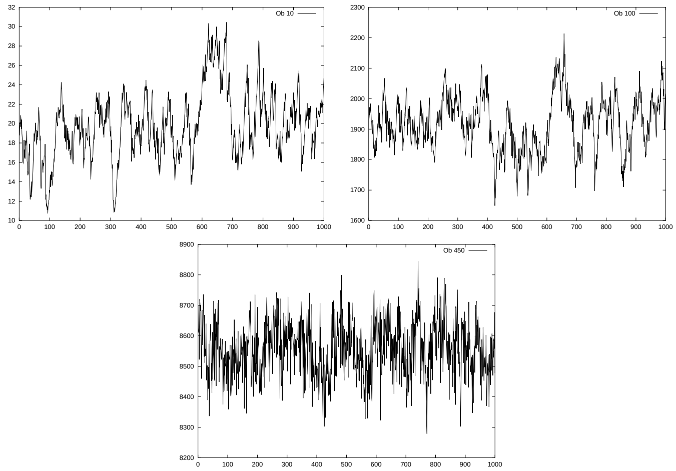

Fig. 4. Plot of TCP simulation by AIMD.

Figure 5 gives the logscale diagram (logarithm of the energy function as a function of the scale, as defined in e.g. [1]) of the traces of Figure 4 for Obs= 10, 100, 450.

These plots are compatible with a fractal behavior in view of the fact that the scaling is concentrated at the lowest scales (highest frequencies) with an upper cutoff (here around the scales 7-8, the slope varies from 1 to 2.).

Figure 6 displays the energy function for the same case as Figure 5 but for a state dependent synchronization rate: here $\gamma_n^{(i)}$ depends on the throughput just before the n-th congestion time $Y_n^{(i)}$ as follows : if $Y_n^{(i)} = y$, $\gamma_n^{(i)}[y] = 1/2$ with probability $1 - q^y$ where $0 < q < 1$ is a parameter that has been fixed such that $\mathbb{E}(\gamma[\mathcal{Y}])$ is equal to 0.1 ($\mathcal{Y}$ denotes the stationary throughput

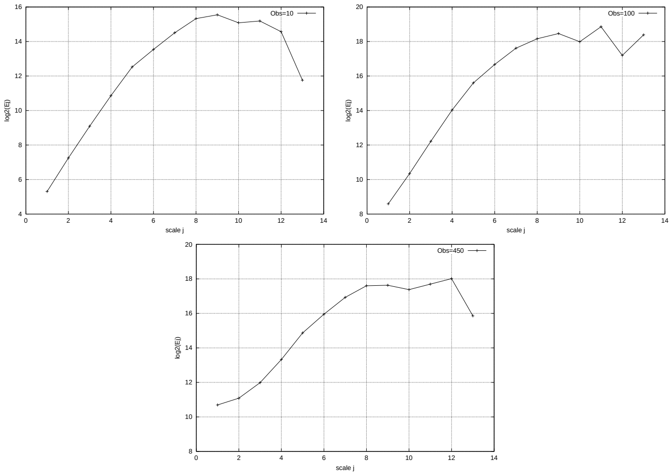

Fig. 5. Plot of energy functions: TCP simulation by AIMD.

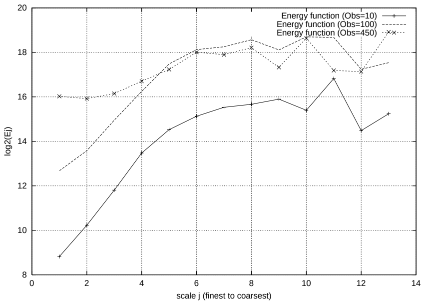

Fig. 6. Plot of energy functions: TCP simulation by AIMD at congestion times for state dependent synchronization rate.

just before congestion times). This shows that even with such a state dependent synchronization rate, the energy function curve is qualitatively the same as that in the state independent case, and that the statistical conclusions derived in this section are not an artefact of the state independent case.

Generally speaking, the slope α of the linear part is related to the local regularity parameter h of the fractal. If $\alpha > 1$, in the Gaussian case, the sample paths are continuous and non-differentiable, whereas $\alpha < 1$ corresponds to non continuous paths [10] (in fact the steepest slopes correspond to the most regular paths as well illustrated by Figure 4 and 5).

In cases $Obs \leq 300$, we have $\alpha > 1$; then the local regularity parameter h of the fractal is given by the formula $h = \frac{\alpha - 1}{2}$. For instance, for $Obs = 10$, $\alpha \sim 2$ so that $h \sim 1/2$; whereas for $Obs = 300$, $\alpha \sim 1.3$ and $h = 0.15$.

On case $Obs = 450$, we have $\alpha \sim 1$; in this and other cases where $\alpha < 1$, sample paths are discontinuous.

The partition sum at scale j is defined as:

$$S(q,j) = \sum_{k=1}^{2^j} |2^{-j/2} d_{j,k}|^q.$$

In Figure 7, we plot the log of this function of the scale for $q = 0, 2, \ldots, 18$ ($N = 500, p = 0.1, Obs = 10$).
The linearity of the log-log plot of the partition sum is also characteristic of a fractal [12].

The multiscale diagrams, which plot the slope of the partition

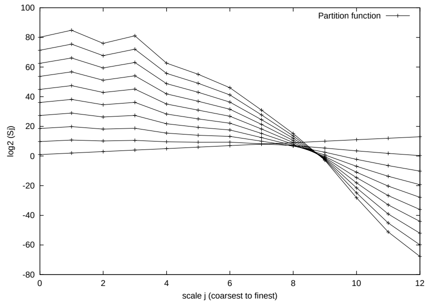

Fig. 7. Plot of partition sum: TCP simulation by AIMD.

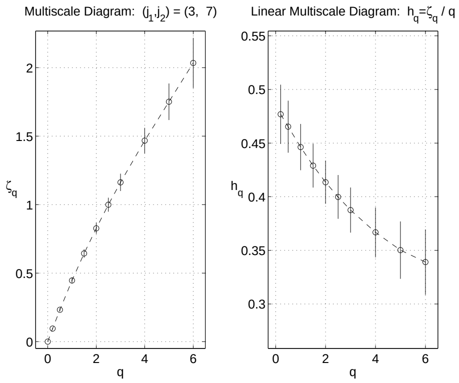

Fig. 8. Multiscale diagrams for AIMD traces.

sum as a function of q, are given in Figure 8 for the case $N = 100, p = 0.1$ and $Obs = 10$. The non-linearity of the multiscale diagram is consistent with a *multifractal* scaling.

A.3 Consistency with the value of the autocorrelation function

Under stationary Gaussian assumptions, the regularity factor h of a stochastic process $\mathcal{T}(t)$ is related to its increments by the formula:

$$\mathbb{E}[(\mathcal{T}(t+s) - \mathcal{T}(t))^2] \sim |s|^{2h}, \tag{33}$$

when s tends to 0. Let $\mathcal{T}_n = \mathcal{T}_n^H$ be the H-aggregate of the stationary throughputs for $card(H) = K = \phi N$, at the n-th congestion time. When the mean inter-congestion time is small, one can use (33) to write the following rough heuristic expressions

$$\begin{aligned} \mathbb{E}[(\mathcal{T}_1 - \mathcal{T}_0)^2] &\sim (\mathbb{E}[\tau])^{2h} \\ \mathbb{E}[(\mathcal{T}_2 - \mathcal{T}_0)^2] &\sim (2\mathbb{E}[\tau])^{2h}, \end{aligned}$$

which in turn implies

$$h \sim \frac{1}{\ln(4)} \ln \left(\frac{\mathbb{E}[(\mathcal{T}_2 - \mathcal{T}_0)^2]}{\mathbb{E}[(\mathcal{T}_1 - \mathcal{T}_0)^2]} \right). \tag{34}$$

Using the results of Theorem 2 and the relation

$$\mathbb{E}[(\mathcal{T}(t+s) - \mathcal{T}(t))^2] = 2\left(var(\mathcal{T}(t)) - cov(\mathcal{T}(t+s), \mathcal{T}(t))\right),$$

we can now compute the RHS which gives the following heuristic estimate of h when N is large:

$$h \sim \frac{1}{\ln(4)} \ln \left(\frac{2 - \frac{p}{2}}{1 + \phi\left(1 - \frac{p}{2}\right)} \right). \tag{35}$$

This estimate is consistent with the wavelet based estimates. For instance, if we take $\phi = 0.6$ and $p = 0.01$, both the wavelet estimate and (35) give $h \sim 0.16$.

VI. Appendix

A. *Proof of Theorem 1*

We have to check that the series defining $\mathcal{X}_n$, with general term $V_k = A_0 A_{-1} A_{-2} \dots A_{1-k} B_{-k}$, is convergent. For this, it is enough to show that the series $\sum_k ||V_k||_2$ is convergent,

The fact that (9) is a finite stationary solution of (4) then follows from classical arguments on iterates of random maps (see e.g. [8], [9]).

We will use the following notation

- Γ_n is the diagonal matrix with entries $(\Gamma_n)_{ii} = \gamma_n^{(i)}$;
- $\mathbb{E}(\Gamma_n) = \mathbb{G}$, the diagonal matrix with entries $\mathbb{E}\gamma_0^{(1)} = \sqrt{c}$;
- $\mathbb{R} = \rho\mathbb{J}$, with $\rho = 1/N$;

Notice that we have: $A_n = \Gamma_n(\mathbb{I} - \mathbb{R})$, and that $(\mathbb{I} - \mathbb{R})\mathbb{J}$ is the null matrix. Using this, we obtain the following expressions:

$$\mathbb{E}B_0 B_0^t = \frac{C^2}{N^2}\left((a-b)\mathbb{I} + b\mathbb{J}\right), \quad \mathbb{E}B_0 \mathbb{E}B_0^t = \frac{C^2}{N^2} c\mathbb{J},$$

so that

$$\mathbb{E}B_0 B_0^t - \mathbb{E}B_0 \mathbb{E}B_0^t = \frac{C^2}{N^2}\left((a-b)\mathbb{I} + (b-c)\mathbb{J}\right) \tag{36}$$

and

$$\mathbb{E}B_0^t B_0 = \text{Trace}\left(\mathbb{E}B_0 B_0^t\right) = \frac{C^2}{N} a.$$

Similarly,

$$\begin{aligned} \mathbb{E}A_0 B_{-1} B_{-1}^t A_0^t &= \frac{C^2}{N^2} \mathbb{E}\left(\Gamma_0(\mathbb{I} - \mathbb{R})((a-b)\mathbb{I} + b\mathbb{J})\right. \\ &\qquad \left.(\mathbb{I} - \mathbb{R})^t \Gamma_0^t\right) \\ &= \frac{C^2}{N^2} \mathbb{E}\left(\Gamma_0(\mathbb{I} - \mathbb{R})((a-b)\mathbb{I})(\mathbb{I} - \mathbb{R})^t \Gamma_0^t\right) \\ &= \frac{C^2}{N^2}(a-b)\mathbb{E}\left(\Gamma_0(\mathbb{I} - \mathbb{R})\Gamma_0^t\right) \\ &= \frac{C^2}{N^2}(a-b)((a - \rho(a-b))\mathbb{I} - b\mathbb{R}), \end{aligned}$$

where we used the relations $\mathbb{E}(\Gamma_0 \Gamma_0^t) = a\mathbb{I}$ and $\mathbb{E}(\Gamma_0 \mathbb{R} \Gamma_0^t) = \rho(a-b)\mathbb{I} + b\mathbb{R}$. Since in addition, $\mathbb{E}A_0 B_{-1} \mathbb{E}B_{-1}^t A_0^t = 0$, we get

$$\mathbb{E}V_1 V_1^t - \mathbb{E}V_1 \mathbb{E}V_1^t = \frac{C^2}{N^2}(a-b)((a - \rho(a-b))\mathbb{I} - b\mathbb{R}) \tag{37}$$

and

$$\mathbb{E}V_1^t V_1 = \text{Trace}\left(\mathbb{E}A_0 B_{-1} B_{-1}^t A_0^t\right) = \frac{C^2}{N}(a-b)a(1-\rho).$$

More generally, by the same argument, an easy induction argument proves that for all $k \geq 1$,

$$\mathbb{E}V_k V_k^t = \frac{C^2}{N^2}(a-b)(a - \rho(a-b))^{k-1}((a - \rho(a-b))\mathbb{I} - b\mathbb{R}),$$

whereas $\mathbb{E}V_k\mathbb{E}V_k^t = 0$. Hence

$$\begin{aligned}\mathbb{E}V_kV_k^t - \mathbb{E}V_k\mathbb{E}V_k^t &= \frac{C^2}{N^2}(a-b)(a-\rho(a-b))^{k-1} \\ &\quad \times((a-\rho(a-b))\mathbb{I} - b\mathbb{R}), \quad (38)\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}V_k^tV_k &= \text{Trace}\left(\mathbb{E}V_kV_k^t\right) \\ &= \frac{C^2}{N}(a-b)(a-\rho(a-b))^{k-1}a(1-\rho). \quad (39)\end{aligned}$$

Finally, the same type of arguments, for all $l \neq k$,

$$\mathbb{E}A_0A_{-1}\cdots A_{1-l}B_{-l}B_{-k}^tA_{1-k}^t\cdots A_{-1}^tA_0^t = 0$$

and

$$\mathbb{E}A_0A_{-1}\cdots A_{-k+1}B_{-k}\mathbb{E}B_{-l}^tA_{-l+1}^t\cdots A_{-1}^tA_0^t = 0.$$

Using the facts that $0 < \rho < 1$ and $0 < b < a < 1$, we get $0 < a - \rho(a-b) < 1$; this together with (39) imply that $\sum_k ||V_k||_2 < \infty$, which concludes the proof that the series converges.

In order to prove uniqueness, one assumes that there exist two bounded stationary processes $\{X_n\}$ and $\{X_n'\}$ that are solution of (4) and such that the process at time n is independent of the matrices $A_{n+1}, A_{n+2}, \ldots$. Then, $Z_n = X_n - X_n'$ is such that for all n

$$Z_0 = A_0A_{-1}\ldots A_{-n+1}Z_{-n},$$

with Z_{-n} independent of $A_0A_{-1}\ldots A_{-n+1}$. So

$$||Z_0||_2^2 = \mathbb{E}Z_{-n}^tA_{-n+1}^t\ldots A_0^tA_0\ldots A_{-n+1}Z_{-n}.$$

Taking first the conditional expectation of the random vector $Z_{-n}^tA_{-n+1}^t\ldots A_0^tA_0\ldots A_{-n+1}Z_{-n}$ w.r.t. the matrices $A_0\ldots A_{-n+1}$ leads again to a coefficient of the form $a-\rho(a-b)^{n-1}$. This and the boundedness of Z_{-n} allow one to conclude that $||Z_0||_2 = 0$.

We now prove (10). We have $\mathbb{C}(\mathcal{X}) = \mathbb{C}(\sum_{k\geq 0} V_k)$ and

$$\mathbb{C}\left(\sum_{k\geq 0} V_k\right) = \sum_{k\geq 0}\sum_{l\geq 0}\mathbb{E}\left(V_kV_l^t\right) - \mathbb{E}\left(V_k\right)\mathbb{E}\left(V_l^t\right).$$

This and (36,37,38) completes the proof of (10).

The last relation of the theorem is obtained from the following asymptotics:

$$\begin{aligned}&a = 1 - 3\pi/4 + o(1); \quad b = (1-\pi/2)^2 + o(1); \\ &a - b = \pi(1-\pi)/4 + o(1); \quad b - c = o(N^{-1}).\end{aligned}$$

B. Proof of Theorem 3

Let $I_n = \{0 \leq i \leq N : \gamma_n^{(i)} = 1/2\}$ and

$$S(n) = \sum_{i=1}^N \mathcal{X}_n^{(i)}, \quad S_-(n) = \sum_{i\in I_n}\mathcal{X}_n^{(i)}, \quad S_+(n) = \sum_{i\notin I_n}\mathcal{X}_n^{(i)}.$$

With this notation, $\tau_{n+1} = \frac{C-S(n)}{N}R^2$ and $C = S_+(n) + 2S_-(n) = S(n) + S_-(n)$, so that from (3),

$$\mathcal{X}_{n+1}^{(i)} = \gamma_n^{(i)}\left(\mathcal{X}_n^{(i)} + \frac{S_-(n)}{N}\right).$$

If $I_n = I_{n+1} = I$, this in turn implies

$$S_-(n+1) = \frac{N+|I|}{2N}S_-(n). \quad (40)$$

So, if $I_n = I_{n+1} = \cdots = I_{n+k} = I$, for some $k > 0$, then

$$\begin{aligned}S_-(n+k) &= \left(\frac{N+|I|}{2N}\right)^k S_-(n) \\ &\leq \left(\frac{N+|I|}{2N}\right)^k C. \quad (41)\end{aligned}$$

So, for all $\epsilon > 0$, if $k = k(\epsilon)$ denotes the smallest integer such that $\left(\frac{N+|I|}{2N}\right)^k C \leq \epsilon$, then

$$\mathbb{P}(S_-(n+k) \leq \epsilon) \geq \mathbb{P}(I_n = I_{n+1} = \cdots = I_{n+k} = I) = \mu^{k+1},$$

where μ denotes the probability that $I_n = I$, conditional on the fact that there is at least one loss, which is given in (32). Direct calculations based on logarithms lead to the equivalence

$$\mu^k \sim D\epsilon^\xi,$$

with D a constant and ξ defined in (32).

Similarly, if J is a fixed subset of $\{1,\ldots,N\}$, disjoint of I and such that $|I| = |J|$ (we assume that $|I|$ is small enough for this to be possible), then

$$T^J(n+1) = T^J(n) + \frac{|I|}{N}S_-(n). \quad (42)$$

So, if $I_n = I_{n+1} = \cdots = I_{n+k} = I$, for some $k > 0$, with $|I| = j$, then

$$\begin{aligned}T^J(n+k) &= T^J(n) + \frac{j}{N}\frac{1-\left(\frac{j+N}{2N}\right)^{k+1}}{1-\frac{j+N}{2N}}S_-(n) \\ &\geq S_-(n)\frac{2j}{N-j}\left(1-\left(\frac{j+N}{2N}\right)^{k+1}\right).\end{aligned}$$

Let now $k = k(\epsilon)$ denote the smallest integer such that

$$\left(\frac{j+N}{2N}\right)^k \leq \epsilon\frac{2j}{N-j}\left(1-\left(\frac{j+N}{2N}\right)^{k+1}\right),$$

where ϵ is small. Then

$$\begin{aligned}&\mathbb{P}\left(\frac{\max_{H,|H|=j}T^H(n+k)}{\min_{H,|H|=j}T^H(n+k)} \geq \frac{1}{\epsilon}\right) \\ \geq\ &\mathbb{P}\left(I_n = \cdots = I_{n+k} = I, \frac{T^J(n+k)}{T^I(n+k)} \geq \frac{1}{\epsilon}\right) \\ =\ &\mathbb{P}(I_n = I_{n+1} = \cdots = I_{n+k} = I) \\ =\ &\mu^{k+1} \\ \sim\ &D'\left(\frac{1}{\epsilon}\right)^{-\xi}, \quad (43)\end{aligned}$$

with D' a constant and ξ defined as in (32)

VII. Conclusion and Future Work

We have proposed a new model, called the AIMD model, allowing one to describe the interaction of many TCP sessions sharing a common tail drop bottleneck router, and to take synchronization into account. We have shown that this model, which is based on mere random matrix products, captures the key features of the additive increase, multiplicative decrease protocol, and provides a qualitatively correct fluid approximation of TCP.

The model leads to closed form expressions for the steady state covariances (or the autocorrelation functions) of various quantities of interest. This was exploited in several directions.

We first studied how synchronization affects throughput by giving new formulas refining the classical square root type formula and by quantifying the under-utilization when losses are synchronized in the case with a finite buffer.

A mathematical characterization of the instantaneous unfairness of the protocol was also derived, and it was shown that even in the case where TCP is supposed to be fair, the dispersion of throughputs is very high.

Finally, we used wavelet tools similar to those of Feldmann, Gilbert, Willinger et al. to show that what is interpreted as a fractal scaling at short time scales on real TCP traffic traces is already present in the AIMD model. Note that if the fractal properties of traffic at short time scales stem in part from the AIMD principle of TCP, that is from the adaptive nature of the protocol, then open loop models of TCP resource sharing such as those based on a queue fed by some fractal input process are not adequate for assessing the impact of this scaling.

The basic model can be enriched in several ways. A first natural aim would be to have a finer interplay between such high level models of competition of several TCP sessions and detailed one-session packet level models such as the one proposed in [5]. Note that in such a refined model, the additional aggregation of packets due to slow start, and the additional variability due to the packet scale phenomena, can apparently only reinforce the dispersion and the local irregularity already present in the fluid AIMD model.

Acknowledgment

We would like to thank both R. Riedi and D. Veitch for their comments which allowed us to improve the wavelet analysis part of the paper. In particular, the multiscale diagram of Figure 8 which leads to the conclusion that the AIMD traces exhibit a multifractal scaling, were obtained using the statistical tool of P. Abry and D. Veitch: http://www.emulab.ee.mu.oz.au/∼darryl .

References

[1] Abry, P., Flandrin, P., Taqqu, M.S. and Veitch, D. (2000) Wavelet for the analysis, estimation and synthesis on scaling data. *Self Similar Traffic Analysis and Performance Evaluation*, Park, K. and Willinger, W. Eds, Wiley.

[2] Altman, E., Avrachenkov, K. and Barakat, C. (2000) A Stochastic Model of TCP/IP with Stationary Random Losses. *Proc. of the ACM-SIGCOMM'00*, Stockholm, Vol. 30, No. 4, pp. 231-242.

[3] Altman, A., Barakat, C. Laborde E., Brown P., and Collange D. (2000) Fairness Analysis of TCP/IP. *Proc. of IEEE Conference on Decision and Control*, Sydney, December.

[4] Baccelli, F. and Brémaud, P. (1994) Elements of Queueing Theory. *Springer Verlag*.

[5] Baccelli, F. and Hong, D. (2000) TCP is Max-Plus Linear. *Proc. of ACM-SIGCOMM'00*, Stockholm, Vol. 30, No. 4, pp. 219-230.

[6] Barakat C. and Altman E. (2000) A Markovian Model for TCP Analysis in a Differentiated Services Network. *Proc. of the first international workshop on Quality of future Internet Services (QofIS)*, Berlin, September.

[7] Bohacek, S. Hespanha, J. P., Lee, J. and Obraczka K. (2001) A Hybrid Systems Framework for TCP Congestion Control, *Technical Report*, USC, Los Angeles, CA, July.

[8] Bougerol, P., Picard, N. (1992) Strict stationarity on generalized autoregressive processes. *Ann. Probab.*, 20, pp. 1714-1730.

[9] Diaconis, P. and Freedman, D. (1998) Iterated Random Functions. *Preprint*, http://www.stat.Berkeley.EDU/users/freedman/511.ps.Z.

[10] Falconer, K. (1990) Fractal Geometry. *Wiley and Sons*, New York.

[11] Feldmann, A., Gilbert, A.C., Huang, P. and Willinger, W. (1999) Dynamics of IP traffic: A study of the role of variability and the impact of control. *Proc. of the ACM-SIGCOMM'99*, August-September, Cambridge, MA, pp. 301-313.

[12] Feldmann, A., Gilbert, A.C. and Willinger, W. (1998) Data networks as cascades: Investigating the multifractal nature of Internet WAN traffic. *Proc. of the ACM-SIGCOMM'98*, Vancouver, Canada, Vol. 28, No. 4, pp. 42-55.

[13] Feldmann, A., Gilbert, A.C., Willinger, W. and Kurtz, T.G. (1998) The Changing Nature of Network Traffic: Scaling phenomena. *Computer Communication Review*, Vol. 28, No. 2, pp. 5-29.

[14] Gilbert, A.C., Willinger, W. and Feldmann, A. (1999) Scaling analysis of conservative cascades, with applications to network traffic. *IEEE Transaction on Information Theory*, Vol. 45, pp. 971-991.

[15] Guo, M., Crovella, M. and Matta, I. (2000) TCP Congestion Control and Heavy Tails. *Technical Report* BUCS-TR-2000-017, Boston University.

[16] Kesten, H. (1973) Random difference equations and renewal theory for products of random matrices. *Acta Math.*, Vol. 131, pp. 207-248.

[17] Mathis, M., Semske, J., Mahdavi, J. and Ott T. (1997) The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm. *Computer Communication Review*, 27(3), July 1997.

[18] Padhye, J., Firiou V., Towsley, D. and Kurose, J. (1998) Modeling TCP throughput: a simple model and its empirical validation. *Proc. of ACM SIGCOMM*.

[19] Riedy R. and Levy-Vehel, J. (1996) Multifractal Properties of TCP Traffic. *Technical Report*, RR-3129, INRIA Rocquencourt.

[20] Willinger, W., Paxson, V. and Taqqu, M.S. (1998) Self-Similarity and Heavy Tails: Structural Modeling of Network Traffic. *A Practical Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy Tailed Distributions*, R. Adler, R. Feldman and M.S. Taqqu (Eds.), Birkhauser Verlag, Boston, MA, pp. 27-53.

[21] Willinger, W., Taqqu, M.S., Sherman, R. and Wilson, D.V. (1997) Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level. *IEEE ACM Transactions on Networking*, Vol. 5, No. 1, pp. 71-86.

[22] Wright, G. and Stevens R. (1995) TCP Illustrated Vol. 2, *Addison Wesley*.