Interaction of TCP Flows as Billiards

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

***Abstract*— The aim of this paper is to analyze the performance of a large number of long lived TCP controlled flows sharing many routers (or links), from the knowledge of the network parameters (capacity, buffer size, topology) and of the characteristics of each TCP flow (RTT, route etc.) when taking synchronization into account. It is shown that the dynamics of such a network can be described in terms of iterate of random piecewise affine maps, or geometrically as a billiards in the Euclidean space with as many dimensions as the number of flow classes and as many reflection facets as there are routers. This class of billiards exhibits both periodic and non-periodic asymptotic oscillations, the characteristics of which are extremely sensitive to the parameters of the network. It is also shown that for large populations and in the presence of synchronization, aggregated throughputs exhibit fluctuations that are due to the network as a whole, that follow some complex fractal patterns, and that come on top of other and more classical flow or packet level fluctuations. The consequences on TCP's fairness are exemplified on a few typical cases of small dimension.**

I. Introduction

The Additive Increase, Multiplicative Decrease (AIMD) model introduced in [4] describes the joint evolution of the congestion window size of N long lived (FTP or Peer to Peer type) flows controlled by TCP and sharing a single router or link, in terms of products of random matrices. The associated large population asymptotic model which concerns the case $N \to \infty$ was studied in [12].

The present paper studies the case when the TCP flows are heterogeneous, i.e. with different Round Trip Times (RTTs) or routes, and when each flow goes through a route made of several tail-drop routers (throughout the paper, we will consider routers to be the possible bottlenecks; this could be replaced by links everywhere without altering the conclusions) in series. The corresponding model, which is introduced in §II, will be referred to as the multi-AIMD model.

Our aim is to estimate the throughput obtained by each individual flow under the competition rules imposed by TCP, and also the fluctuations of this throughput, from the sole knowledge of the route and the RTT of each flow, and the characteristics of each router and link (buffer size, link capacity etc.) in the network.

This is of course related to the classical relationships that have been obtained between the packet loss probability and TCP throughput for a given session (see e.g. [17]); in particular, it was shown in [4] that the single router AIMD dynamics resulted in a dependency between these quantities that was compatible with these formulas and was actually refining them in that it allowed one to assess the influence of synchronization.

The first models for the several router TCP network case are those of [13] and [10]. These papers analyzed the bandwidth sharing of different TCP flows over large networks in terms of optimization problems, and triggered a large number of further studies (see e.g. [16], [15]). The prediction of the throughput in the several AQM router case has also been investigated via fixed point approximation methods for mean values in [8]. The approach that is proposed in the present paper addresses the same prediction question in the tail drop router case. The main difference with these earlier approaches lies in the fact that we use a pathwise description of the dynamics of the interaction between flows, which takes into account discrete event phenomena that are of central importance for tail drop networks, such as congestion epochs, losses or synchronization of sources, as well as random phenomena which all have an impact on throughput fluctuations.

More precisely, the interaction is described by a set of evolution equations that generalize the random affine map description of the AIMD (one router) model. The basic multi-AIMD model can be seen as iterates of random *piecewise* affine maps. From this stochastic model, we define a large population asymptotic model. This asymptotic model can be seen as iterates of deterministic piecewise affine maps. These equations are shown to admit a geometrical representation in terms of a random or deterministic billiards in the Euclidean space. The dimension of this space is the number of different flow classes (typically, there is one flow class per route and RTT). This billiards has as many reflection facets as there are routers.

This new representation of the interaction between TCP flows over networks made of several links and routers and its exploitation are the main contributions of the present paper. We establish sufficient conditions for the asymptotic periodicity of the throughput obtained by the interacting flows, as well as a conservation law that relates the intensities with which routers experience congestion. Billiards are known to possibly exhibit non-periodic asymptotic behaviors. We give numerical evidence that this is possible for the class of billiards considered here. We also show that the characteristics of this asymptotic behavior are extremely sensitive w.r.t. network parameters. The implications on TCP's fairness and bandwidth sharing are exemplified on a few cases of small dimension. We show however that once the periodic regime is known, fairness can be approached analytically using a mix of linear algebra

and cycle formulas of Palm calculus [3].

The validation of this billiards representation is not addressed in the present paper; it is the object of a companion paper [6], which investigates more generally the use of this approach as a simulation tool. In addition to comparison with NS2 simulation, it is shown there that the aggregated traffic generated by this representation satisfies several empirical or statistical laws that were observed on real traces. This concerns the short time scale statistical properties observed in [20], [24], [1] and the empirical power law describing the sensitivity w.r.t. RTTs reported in [14].

II. Notation and Model Description

A. Notation

The model parameters are the following:

- Network configuration: $\mathcal{R}$ is the set of routers; C_r is the capacity of router $r \in \mathcal{R}$; all routers are assumed to be tail drop.
- Traffic configuration: $\mathcal{S}$ is the set of TCP flow classes; N_s is the number of TCP flows of class $s \in \mathcal{S}$; $\mathcal{P}_s$ is the route of class s flows; depending on the circumstances, any such route will be considered as a sequence of routers or as a sequence of pairs of routers; $RTT_s = R_s$ is the propagation delay for class s flows, which is also the minimal RTT for this class; λ_s denotes the stationary throughput of class s (one of the key variables to be determined).
- Network and traffic configuration: $\mathcal{S}_r$ is the set of classes with a route using router r; M_r is the total number of flows sharing router r: $M_r = \sum_{s\in\mathcal{S}_r} N_s$; for $s \in \mathcal{S}_r$, $a_{s,r} = N_s/M_r$, is the proportion of flows of class s within the set $\mathcal{S}_r$; $c_r = C_r/M_r$ is the throughput that one flow could get if the capacity were ideally and equally shared.

Assumption $\mathcal{A}$ (which will be used for in certain proofs) supposes that each router has at least one class with a route that contains this router only. This assumption is not essential for most properties. However, it ensures that each router reaches congestion infinitely often, which simplifies the exposition of the results.

We now give the notation of the different state variables that we will use. Most of these variables refer to the sequence $\{T_n\}$ of all *congestion epochs* in the network. As in [4], T_n is the n-th epoch at which a loss (or several simultaneous losses) occur on at least one router.

- $X^{(s,i)}(t)$ is the throughput of flow i of class s at time t;
- $X_n^{(s,i)} = X^{(s,i)}(T_n+)$ is the throughput of flow i of class s just after the n-th congestion time;
- $Y_n^{(s,i)} = X^{(s,i)}(T_n-)$ is the throughput of flow i of class s just before the n-th congestion time; by construction, it will always be true that for all $r \in \mathcal{R}$, and all time t,

$$\sum_{s\in\mathcal{S}_r}\sum_{i\in s} X^{(s,i)}(t) \le C_r. \tag{1}$$

  The congestion epoch T_n will be said of type $r \in \mathcal{R}$, if $\sum_{s\in\mathcal{S}_r}\sum_{i\in s} Y_n^{(s,i)} = C_r$. Nothing forbids to have T_n of both type r and r'.
- $\tau_{r,n+1}$ is the time between T_n and the next *virtual congestion epoch* of router r, which is defined as

$$\tau_{r,n+1} = \frac{C_r - \sum_{j,s\in\mathcal{S}_r} X_n^{(s,j)}}{\sum_{s\in\mathcal{S}_r} \frac{N_s}{R_s^2}};$$

  this is the time that would elapse between T_n and the next congestion epoch on router r, should the capacities of all other routers be infinite;
- $\gamma_n^{(s,i,r)}$ is the multiplicative variable of flow $i \in s$ on router r at the n-th congestion epoch: : $\gamma_n^{(s,i,r)} = 1/2$ if there is a loss for flow i on router r; $\gamma_n^{(s,i,r)} = 1$ otherwise; so, $\gamma_n^{(s,i,r)} \equiv 1$ if $r \notin \mathcal{P}_s$.
  Throughout this paper, we will study several types of assumptions.
  The *rate-independent* (RI) model is that where the sequences $\gamma_n^{(s,i,r)}$ are independent in r; for all fixed r, independent and identically distributed (i.i.d.) in n; for all fixed r and s, identically distributed and ergodic in $i \in s$.
  The *rate-dependent* (RD) case is that where the law of $\gamma_n^{(s,i,r)}$ is a function of s, r and $Y_n^{(s,i,r)}$ (a flow of a given class that has a large instantaneous throughput has more chances to experience a loss than another flow of the same class with a smaller throughput). Some RD cases will be studied in §III-C.
- $p_n^{(s,r)} = \mathbb{P}(\gamma_n^{(s,i,r)} = 1/2)$ is the *synchronization rate* of router r for the flows of class s at the n-th congestion epoch. In the rate-independent case, $p_n^{(s,r)} \equiv p^{(s,r)}$. The class-independent (CI) model is that where in addition, $p^{(s,r)} = p^{(r)}$, for all $s \in \mathcal{S}_r$.

The *synchronization rate* should not be confused with the *packet loss rate*. Since the synchronization rate represents the *proportion* of flows that experience a loss during a congestion epoch, it is possible to simultaneously have a high synchronization rate and a low packet loss rate (e.g. when rarely all sources loose at the same time) or the converse.
We show in [6] how this synchronization rate can be estimated from the network parameters using simple queueing theoretic arguments that take into account the delay with which sources react to losses. However, we will not use the specific form of the function proposed there, and for what follows, other estimates could be used as well.

B. Dynamics in the Simplest Case

It will be assumed that routers have small buffer capacity so that it makes sense to assume that the different RTTs are constant over time and equal to R_s for class s.

Since, due to the Additive Increase (AI) rule, each flow of class s increases its send rate with slope $\frac{1}{R_s^2}$ (this is the slope obtained when assuming that the window size and the RTT are linked at any time by a Little like formula: $W = XR$), we get that the sum of the throughputs of all flows using router/link

r increases with slope $\sum_{u\in\mathcal{S}_r}\frac{N_u}{R_u^2}$. This lasts until the the next congestion epoch T_{n+1}, which is the first epoch after T_n when the sum of the instantaneous throughputs through one of the routers/links exceeds the capacity of this router/link. Assume one knows $X_n^{(s,i)}$ for all i and s. Then we get

$$Y_{n+1}^{(s,i)} = X_n^{(s,i)} + \frac{1}{R_s^2}\min_{r\in\mathcal{R}}\tau_{r,n} \quad (2)$$

$$= X_n^{(s,i)} + \frac{1}{R_s^2}\min_{r\in\mathcal{R}}\frac{C_r - \sum_{j,u\in\mathcal{S}_r}X_n^{(u,j)}}{\sum_{u\in\mathcal{S}_r}\frac{N_u}{R_u^2}}. \quad (3)$$

Let $r_n = \mathrm{argmin}_{r\in\mathcal{R}}\tau_{r,n}$. Assume that this set has one element. Then due to the Multiplicative Decrease (MD) rule,

$$X_{n+1}^{(s,i)} = \gamma_{n+1}^{(s,i,r_{n+1})}Y_{n+1}^{(s,i)}. \quad (4)$$

Should there be several elements in the last set, then one would apply the multiplicative rule for all routers of the set (the order in which the multiplicative decrease is made does not affect the result). So the global dynamical system reads: $\forall(i,s)$,

$$X_{n+1}^{(s,i)} = \gamma_{n+1}^{(s,i,r_{n+1})}\Big(X_n^{(s,i)} + \frac{1}{R_s^2}\min_{r\in\mathcal{R}}\frac{C_r - \sum_{j,u\in\mathcal{S}_r}X_n^{(u,j)}}{\sum_{u\in\mathcal{S}_r}\frac{N_u}{R_u^2}}\Big) \quad (5)$$

$$= \gamma_{n+1}^{(s,i,r_{n+1})}\Big(X_n^{(s,i)} + \frac{1}{R_s^2}\min_{r\in\mathcal{R}}\frac{c_r - \sum_{u\in\mathcal{S}_r}\frac{a_{u,r}}{N_u}\sum_{j\in u}X_n^{(u,j)}}{\sum_{u\in\mathcal{S}_r}\frac{a_{u,r}}{R_u^2}}\Big).$$

We see that the vector of throughputs at time T_{n+1} is obtained from that at time T_n via a random map which is piecewise affine.

C. Extensions

The basic model admits several extensions, which will not be discussed here. The case with non zero buffer was analyzed in [4] and [9] for the single link case, and is simulated in [6] for the general topology case; the case with non persistent traffic is also considered in [6].

D. Large Population Asymptotics

When the population grows large, this model admits a deterministic asymptotic model that generalizes that of the single router case as defined in [12]. All variables of interest then depend on a parameter N that grows large. We assume in particular that for all s, $N_s[N] = n_sN$ and that for all r, $C_r[N] = c_rM_r[N]$, so that the proportions $a_{s,r} = n_s/\sum_{u\in\mathcal{S}_r}n_u$ are kept for all s and r.

Theorem 1 *Suppose the losses are rate-independent. Assume in addition that for all s, the initial conditions $X_0^{(s,i)}[N]$ are such that for all (deterministic) sequences of subsets $\sigma[N]$ of the set of flows of class s with a cardinal $|\sigma[N]|$ that tends to ∞, the empirical mean $\frac{1}{|\sigma[N]|}\sum_{i\in\sigma[N]}X_0^{(s,i)}[N]$ converges almost surely (a.s) to a deterministic limit $x_0^{(s)}$ which does not depend on the sequence of subsets that is chosen. Then for all n,*

$$\exists\lim_{N\to\infty}\frac{1}{|\sigma[N]|}\sum_{i\in\sigma[N]}X_n^{(s,i)}[N] = x_n^{(s)} \quad a.s.$$

with $x_n^{(s)}$ deterministic, and such that the limit does not depend on the sequence of subsets that is chosen. In addition, the variables $x_n^{(s)}$, $s\in\mathcal{S}$, satisfy the evolution equation

$$x_{n+1}^{(s)} = \overline{\gamma}_{n+1}^{(s,\overline{r}_{n+1})}\left[x_n^{(s)} + \frac{1}{R_s^2}\overline{\tau}_{n+1}\right], \quad (6)$$

where $\overline{\gamma}_n^{s,r} = \mathbb{E}[\gamma_n^{(s,i,r)}]$, $\overline{\tau}_n = \min_{r\in\mathcal{R}}\overline{\tau}_{r,n}$, $\overline{r}_n = \mathrm{argmin}_{r\in\mathcal{R}}\overline{\tau}_{r,n}$,

$$\overline{\tau}_{r,n+1} = \frac{c_r - \sum_{u\in\mathcal{S}_r}a_{u,r}x_n^{(u)}}{\sum_{u\in\mathcal{S}_r}\frac{a_{u,r}}{R_u^2}}$$

If in addition, the initial condition is such that for all (s,i), the a.s. limit $\lim_{N\to\infty}X_0^{(s,i)}[N] = X_0^{(s,i)}[\infty]$ exists, then for all n, the a.s. limit $\lim_{N\to\infty}X_n^{(s,i)}[N] = X_n^{(s,i)}[\infty]$ also exists, and the sequence of random variables $X_n^{(s,i)}[\infty]$ satisfies the stochastic recurrence equation:

$$X_{n+1}^{(s,i)}[\infty] = \gamma_{n+1}^{(s,i,\overline{r}_{n+1})}\left[X_n^{(s,i)}[\infty] + \frac{1}{R_s^2}\overline{\tau}_{n+1}\right], \quad (7)$$

with $\overline{r}_{n+1}$ and $\overline{\tau}_{n+1}$ the variables defined in the last deterministic equations.

The proof can be found in [5]. Notice that empirical means correspond to what is often referred to as *aggregated traffic*, where the aggregates are here per class.

In this last model, we will denote by $y_n^{(s)}$ (resp. $x^{(s)}(t)$) the variables defined as $x_n^{(s)}$ but from the random variables $Y_n^{(s,i)}$ (resp. $X^{(s,i)}(t)$). We deduce the following inequalities from (1): for all t and r, $\sum_{s\in\mathcal{S}_r}a_{s,r}x^{(s)}(t)\le c_r$.

In what follows, (5), satisfied by the actual throughput vector, will be referred to as the *stochastic multi-AIMD model* and (6), satisfied by the the vector of empirical means, will be referred to as the associated *large population asymptotic model*.

An important question (that will be discussed numerically in §III-E) is that of the speed of convergence and of the nature of the error term when approximating the model with N large but finite by the asymptotic model. First results on the convergence of the moments are reported in [12] for the single router case. In many asymptotic models of this mean field type, a central limit theorem can be established, which allows one to prove that the fluctuations around the limit are Gaussian, and to estimate them (see e.g. [11]). Whether this type of results also holds for the general class of dynamics identified here will be the object of future research.

Equivalent large population equations: If we take as state variables $\widetilde{x}_s = n_sx_s$, in place of x_s, $s\in\mathcal{S}$, then the large population equations can be rewritten under the equivalent form, which will also be used later:

$$\widetilde{x}_{n+1}^{(s)} = \overline{\gamma}_{n+1}^{(s,\widetilde{r}_{n+1})}\left[\widetilde{x}_n^{(s)} + \frac{1}{\widetilde{R}_s^2}\widetilde{\tau}_{n+1}\right], \quad (8)$$

where $\widetilde{R}_s = R_s/\sqrt{n_s}$ and where $\widetilde{\tau}_n = \overline{\tau}_n$, $\widetilde{r}_n = \overline{r}_n$.

E. The Three Levels

The proposed model allows one to represent and nevertheless to decouple three different levels:

1) The *network level*, which is captured by (6) or (8), and where the large population averaging takes place; this level, which we believe to be the main new paradigm identified in the present paper, will be the central object of the mathematical study of the next section.
2) The *flow level*, which is captured by the stochastic equation (7); the results obtained at the network level (e.g. the determination of the period of the sequence $\{\overline{\tau}_n, \overline{r}_n\}$, see Theorem 1) can be used to determine the effect of the network on each flow via the stochastic recurrence (7); this type of stochastic recurrences was already studied (at least in some special cases) in [4], [12] and more recently in [9], and we will not pursue the mathematical analysis of this level in the present paper.
3) The *packet level*, which is captured by the synchronization rate formula (see Formula (7) of [6]), which takes into account the delay of reaction proper to TCP. This packet level influences both the network and the flow levels via the impact of the synchronization rate on this level.

As we will see, each level is responsible for parts of the fluctuations of the throughput obtained by flows of aggregates of flows. By decoupling, we simply mean that the fluctuations of all levels can be analyzed independently.

F. Billiards Interpretation

The dynamics of the large population asymptotic model can be seen as that of a deterministic billiards model, the geometry of which is determined by the routes and the RTTs of the various flow classes and the capacity of the routers. The stochastic multi-AIMD model can be seen as a randomized version of the billiards.

This is illustrated by the three–class, two–router network of Fig. 1. Here $c_1 = c_2 = \frac{C}{2}$, $\mathcal{S}_1 = \{1, 3\}$, $\mathcal{S}_2 = \{2, 3\}$,

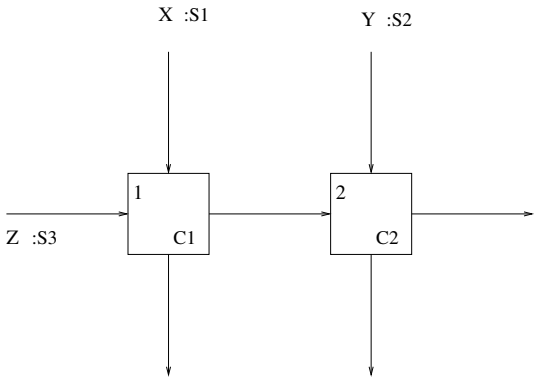

Fig. 1. 2 Router, 3 Class Network Topology

and $a_{s,r} = \frac{1}{2}$ (or equivalently $n_s = 1$) for all s and r. As for RTT's, we take $R_s = 1$ for all s. The synchronization rates are all assumed to be equal to 1, so that $\overline{\gamma}_n^{s,r} = 1/2$ for all s and r. Let us look at the evolution of the large population asymptotic vector $(x^{(1)}(t), x^{(2)}(t), x^{(3)}(t))$ in the three dimensional Euclidean space (X, Y, Z). Notice that in this particular case (6) and (8) are the same. This vector lives in the polyhedron: $X \geq 0, \quad Y \geq 0, \quad Z \geq 0, \quad X + Z \leq C, \quad Y + Z \leq C$, which is depicted on Fig. 2 and is the domain of the billiards. The plane H_1 $(X + Z = C)$ represents the

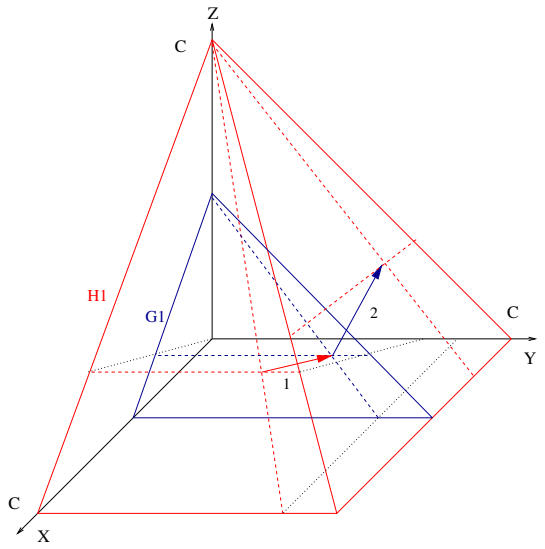

Fig. 2. Billiards Domain

capacity constraint of router 1, with a similar interpretation for the plane H_2 $(Y + Z = C)$. From any point in the domain, the ball (i.e. the throughput process) moves linearly with time along the main diagonal with a constant velocity, as a consequence of the AI rule and the fact that all RTT's are the same. If the ball reaches the plane H_1, then it instantaneously jumps (red arrow, or arrow 1 for black and white reading) to the plane G_1 $(X + Z = C/2)$, which describes the occurrence of losses on router 1. After this jump, the process (X, Y, Z) grows along the main diagonal again (blue arrow or arrow 2) until it hits one of the planes H_1 or H_2 (the last one is met first for this trajectory) and so on. Notice that due the these jumps, the process is actually closer to a pinball than to a billiards. Fig. 3 gives a more complete view of the parts of all planes H_1, H_2 and G_1, G_2, where similar phenomena take place, namely jumps from H_2 to G_2 and growth along the main diagonal from G_2 to either H_1 or H_2. Fig. 3 depicts a (projective) view along the main diagonal. In this projective view, any linear increase is just a point. An instance of sequence of additive increases and multiplicative decreases (which appear as arrows) is illustrated there where the ball departs from H_1 and then successively hits G_1, H_2, G_2 and H_1.

In the stochastic model, the multiplicative dynamics is a randomized version of the last one: facets H_1 and H_2 still exist, but reflection on say H_1 sends the ball in a random neighborhood of the point of G_1 where the deterministic billiards jumps. The neighborhood in question is approximately

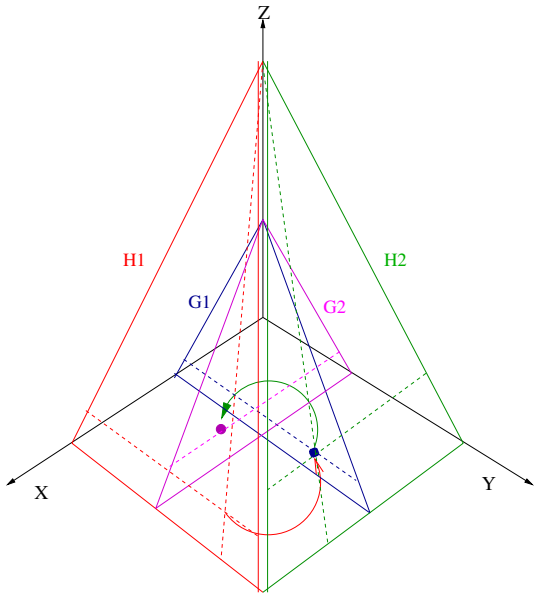

Fig. 3. Billiards Facets and Trajectories.

Gaussian in the three dimensional space.

III. Analysis of the Network Level Equations

Billiards models have been extensively studied using ergodic theory (e.g. Sinai's billiards [22]). Billiards reducible to iterates of piecewise affine maps (our TCP billiards belong to this class) have also been studied (see e.g. [19]). Within this piecewise affine class, even in the case where all maps are contracting, there are unfortunately no general results holding for all dimensions. In particular, there are simple examples of small dimension where some situations lead to a periodic behavior, whereas others lead to a non-periodic one. The subclass of TCP billiards (piecewise affine billiards that stem from TCP dynamics) has some specific properties that could make it amenable to a more specific analysis: the domains where the map is affine are intersections of half-spaces; each map is the composition of the multiplication by a diagonal matrix and of a projection on along some direction etc.

A. Periodic Regime

The aim of this subsection is to give a general sufficient condition for having only periodic behaviors; this sufficient condition is in term of a sequence of linear problems; as we will see, this also allows one to determine the period and the orbit. In this subsection, we will assume $\mathcal{A}$ to hold.

The reference space is the Euclidean space of dimension K, where K is the cardinality of $\mathcal{S}$. We will use (8) rather than (6). We will drop the "tilde" on the variables for the sake of easy notation. Let H_r denote the hyperplane

$$\sum_{s\in\mathcal{S}_r} x^{(s)} = \left(\sum_{s\in\mathcal{S}_r} n_s\right) c_r, \qquad (9)$$

which will be referred to as *facet* r of the billiards.

1) Discrete Time Dynamics: Rather than the continuous time dynamics, we will study the *discrete time dynamics*, $\{y_n\}$, which gives the throughput process sampled *just before* congestion epochs, that is when the ball hits one of the facets.

For all r, let $\phi_r : \mathbb{R}^K \to \mathbb{R}$ denote the affine form

$$\phi_r(y) = \frac{c_r - \sum_{u\in\mathcal{S}_r} a_{u,r} y^{(u)}}{\sum_{u\in\mathcal{S}_r} \frac{a_{u,r}}{R_r^2}}.$$

For all r, s, let $F_{r,s}$ denote the subset of H_r where when applying the discrete time dynamics once, the ball hits facet s at next step. Since the open domain of $\mathbb{R}^K$ where s is hit before any other facet is that where $\phi_s(y) < \phi_v(y)$ for all $v \neq s$, each $F_{r,s}$ is a convex polyhedron of H_r which is the intersection of H_r and of a finite family of half spaces. By definition,

- on $F_{r,s}$, the one-step discrete time dynamics is some affine map that will be denoted by $B_{r,s}$;
- the family $F_{r,s}$, $s = 1, \ldots, |\mathcal{R}|$ =card($\mathcal{R}$) is a partition of H_r (up to the boundary points).

More generally, for all sequences $r_1, \ldots, r_k$ with elements in $\{1, \ldots, |\mathcal{R}|\}$, let $F_{r,r_1,\ldots,r_k}$ be the subset of H_r where when applying the discrete time dynamics k times, one successively visits the facets $r_1, \ldots, r_k$. For all sequences $r_1, \ldots, r_k$,

- on $F_{r,r_1,\ldots,r_k}$, the k-step discrete time dynamics is the affine map $B_{r,r_1,\ldots,r_k} = B_{r_{k-1},r_k} \circ \cdots \circ B_{r,r_1}$;
- the domain $F_{r,r_1,\ldots,r_k}$ is the (possibly empty) intersection of H_r and of the finite family of half spaces:

$$\begin{aligned}
\phi_{r_1}(y) &< \phi_v(y), \ \forall v \neq r_1;\\
\phi_{r_2} \circ B_{r,r_1}(y) &< \phi_v \circ B_{r,r_1}(y), \ \forall v \neq r_2;\\
&\cdots \qquad \cdots\\
\phi_{r_k} \circ B_{r,r_1,\ldots,r_{k-1}}(y) &< \phi_v \circ B_{r,r_1,\ldots,r_{k-1}}(y), \ \forall v \neq r_k;
\end{aligned}$$

- the family $F_{r,r_1,\ldots,r_k}$ $r_i = 1, \ldots, |\mathcal{R}|$, $i = 1, \ldots, k$, forms a partition of H_r up to the boundary points.

The discrete time dynamics features a sequence $\{r_i\}$, $i \geq 0$ of faces that are successively hit, which depends on the initial condition for the throughput vector. Using $\mathcal{A}$, one proves:

Lemma 1 (facet hiting) *For all r, for all initial conditions in H_r, the sequence $\{r_i\}$, with $r_0 = r$, exits r in a number of steps bounded by a constant e_r. For all r, for all initial conditions in $F_r = \cup_{s\neq r} F_{r,s}$, the sequence $\{r_i\}$, with $r_0 = r$, returns to r in a number of steps bounded by a constant f_r.*

We will say that step i is an *exit step from* r if $r_i = r$ and $r_{i+1} = s \neq r$. Fix $r_0 = r$, some initial condition in F_r (so that $i = 0$ is an exit step from r) and consider the discrete time dynamics until the next exit step from r. This next exit step is finite as a corollary of Lemma 1.

The set of possible facet sequences $r_0 = r, r_1, \ldots, r_k$, $k \in \mathbb{N}$, r_1 in $\{1, \ldots, |\mathcal{R}|\}$, between two exit steps from r is that with $k \leq e_r + f_r$, $r_{k-1} = r$ and $r_l \neq r$, for all $l = 1, \ldots p$ with $p < k-1$ and $p \leq f_r$. The cardinality of this set, denoted by q_r, is finite.

Denote by θ_r the mapping that associates to each initial condition in F_r the point where the ball is located at the next exit step from r, or equivalently when it first returns to F_r. From what precedes, θ_r satisfies:

Lemma 2 (facet partition) *The domain F_r can be partitioned into a finite number of convex polyhedrons $E_{r,1}, E_{r,2}, \ldots, E_{r,q_r}$, each of which is the intersection of H_r and of a finite family of half spaces. These domains, which we will refer to as the* linearity domains *of facet r, are such that for all n, for all initial conditions in the interior of $E_{r,n}$,*

- *the sequence of facets that are successively hit until the first return to F_r is exactly the same;*
- *θ_r is some affine mapping $A_{r,n}$ from F_r to itself.*

2) Sufficient Conditions for Facet Periodicity: The simplest sufficient condition for the periodicity of the facet sequence, which will be referred to as the *inclusion test* in what follows, is given in the following lemma.

Lemma 3 (facet periodicity) *If for some r, for all $n = 1, \ldots, q_r$, the set $A_{r,n}(E_{r,n})$ is completely included in one of the linearity domains, say $E_{r,g(n)}$, of facet r, then, for all initial conditions of the throughput process, the sequence of facets is ultimately periodic.*

Proof: Iterates of g, which is a function from a finite set to itself are ultimately periodic. ∎

In case this condition is not satisfied, one ought to check whether $\theta_r^2 = \theta_r \circ \theta_r$ satisfies the appropriate inclusion property: denote by $E^{(2)}_{r,1}, \ldots, E^{(2)}_{r,q_r^{(2)}}$ the linearity domains of this map (again defined as the intersection of H_r and of certain half spaces) and by $A^{(2)}_{r,1}, \ldots, A^{(2)}_{r,q_r^{(2)}}$ the affine maps on these domains. Then if $A^{(2)}_{r,n}(E^{(2)}_{r,n}) \subset E^{(2)}_{r,g(n)}$ for all n, for some function $g : \{1, \ldots, |\mathcal{R}|\} \to \{1, \ldots, |\mathcal{R}|\}$, then for all initial conditions of the throughput process, the sequence of successively hit facets is ultimately periodic.

A similar sufficient condition (which will be referred to as the inclusion test of order k) can be obtained from $\theta_r^k = \theta_r \circ \theta_r^{k-1}$ for any $k \geq 2$.

3) Sufficient Conditions for Billiards Periodicity:

Lemma 4 (billiards periodicity) *Let $r_1, r_2, \ldots, r_n$ be a fixed periodic sequence of facets. Then in the class independent (CI) case, for all dynamics with a sequence of facets which is ultimately periodic, with period $r_1, r_2, \ldots, r_n$, the associated billiards is asymptotically periodic and with a uniquely defined period that is independent of the chosen initial conditions.*

The proof of the lemma, which is based on a contraction argument, can be found in [5]. Combining this last lemma and the sufficient conditions for the periodicity of the facet process provides a sufficient condition for the periodicity of the throughput process. As it will be exemplified in the following examples, this also provides a way of computing the value of the throughput process over a period.

An interesting question which is still open at this stage is that of the irreducibility. When $\mathcal{A}$ is not assumed to hold, it is quite easy to find networks (e.g. with 3 routers) where two or more different periodic regimes can be reached depending on the initial condition. When $\mathcal{A}$ holds, we did not find situations with multiple non-degenerate periodic regimes yet (i.e. regimes where the periodic regime is such that ball bounces on the intersection of more than one facet during the period).

Notice that in the general case, the number of linearity domains grows in a non-polynomial way with $K = \text{Card}\mathcal{S}$ and the order k of the inclusion test. This clearly indicates that this method, when employed as an analytical modeling tool, can unfortunately not be used to assess the properties of large networks. As we will see below, it is however an efficient tool for analyzing small networks.

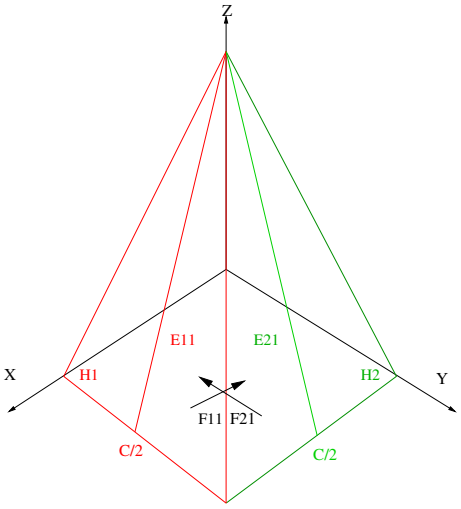

Fig. 4. Periodic Billiards Trajectories

a) Example 1: Consider the network introduced at the end of the last section. The ball lives in the polyhedron of Fig. 3. Let $y_n = (X_n, Y_n, Z_n)$ be the three dimensional vector of throughputs just before the n-th congestion epoch. At these epochs, the ball is on one of the two facets H_1 and H_2 (respectively the red or leftmost and the green or rightmost sides of this polyhedron). Let Δ be the dashed line on the red (leftmost) facet. This line partitions H_1 into two triangular domains, the rightmost of which is $F_{1,2}$: if y_0 belongs to $F_{1,2}$, y_1 belongs to H_2, and it is obtained from y_0 by the affine transformation

$$B_{1,2}(X,Y,Z) = \frac{1}{4}\begin{pmatrix} 2 & -2 & -1 \\ 0 & 2 & -1 \\ 0 & -2 & 1 \end{pmatrix}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \frac{1}{2}\begin{pmatrix} C \\ C \\ C \end{pmatrix}.$$

If y_0 belongs to the complement of $F_{1,2}$, then y_1 belongs to H_1 and it is given from y_0 via another affine transformation. The situation is similar on the facet H_2, to which one associates

two domains $F_{2,1}$ (which leads to H_1 via $B_{2,1}$) and its complement (which leads to H_2).

In this example F_1 has a single linearity domain $E_{1,1} = F_{1,2,1,2}$ and the inclusion test holds as $F_{1,2,1,2} \subset F_{1,2}$.

This implies that regardless of the initial condition in H_1, the sequence of facets is ultimately periodic with period 2 and that the sequence of affine operators that are applied is periodic too, also with period 2. Since $A_{1,1}$ is a contraction from F_1 to itself, it admits a unique fixed point.

It is not difficult to check that the last conclusions actually hold for any configuration as above but for general RTTs R_x, R_y and R_z and general synchronization rates p_x, p_y and p_z provided that $R_x = R_y$ and $p_x = p_y$.

In the special case $R_x = R_y = R_z = 1$, $p_x = p_y = p_z = 1$ that is considered above, the fixed point of $A_{1,1}$ is easily computed as being $X^* = C/2, Y^* = 2C/3, Z^* = C/3$. This fully determines the periodic behavior which is unique in this case, at least when excluding degenerate periodic regimes such as the one that oscillates from a point of the line $X = Y = 1 - Z$ to another point of the line $X = Y = 1/2 - Z$.

The stationary throughput in continuous time, which is the average of the periodic throughput process depicted on Fig. 4, is easily obtained from a cycle type formula: $\lambda_x = \lambda_y = C/2$ and $\lambda_z = C/4$. If $R_x = R_y = 1$ and $p_x = p_y = p_z = 1$, direct calculations give

$$\lambda_x = \lambda_y = \frac{3CR_z^2}{2(2R_z^2+1)}, \quad \lambda_z = \frac{3C}{4(2R_z^2+1)}. \qquad (10)$$

b) Example 2: We come back to the network of Fig. 2, still with $n_s = 1$ for all s. Here, we take $c_1 = 2c_2 = 1/2$. The billiards associated with (8) now lives in a less symmetrical polyhedron depicted in the top left part of Fig. 5. The linearity domains of H_1 are given in the top right part of this figure, which gives a view of H_1 projected on the $X = 0$ plane. From $E_{1,1} = F_{1,2,2,1,2}$, the ball hits H_2 twice before coming back to F_1, whereas in $E_{1,2} = F_{1,2,2,2,1,2}$ it hits H_2 three times before returning to F_1. The Δ line that separates the two linearity domains is here $10Y + 11Z = 8$, $Z + X = 2$. In this case, the inclusion test does not hold for $k = 1$ but it does for $k = 6$, and there is a unique periodic regime of period 19.

The projection of this periodic regime on the $Y = 0$ plane is depicted on Fig. 5. The leftmost region (up to 1.03 of the horizontal axis) is the image of $E_{1,2}$ by θ, whereas the region between 1.03 and 1.33 is the image of $E_{1,1}$. The rightmost line is the projection of H_1 on H_2. The facet period is $((H_2)^3 E_{1,2})^4 (H_2)^2 E_{1,1}$ (with a notation that should be clear) and the billiards period is easily deduced from the unique fixed point of the corresponding affine operator.

B. Non-Periodic Regimes

The aim of this section is to provide numerical evidence that non-periodic facet sequences are possible. The way for searching for such behaviors consists in choosing some topology where a single parameter like e.g. the speed of some router, or the population of some class, is varied and in plotting the period of the billiards.

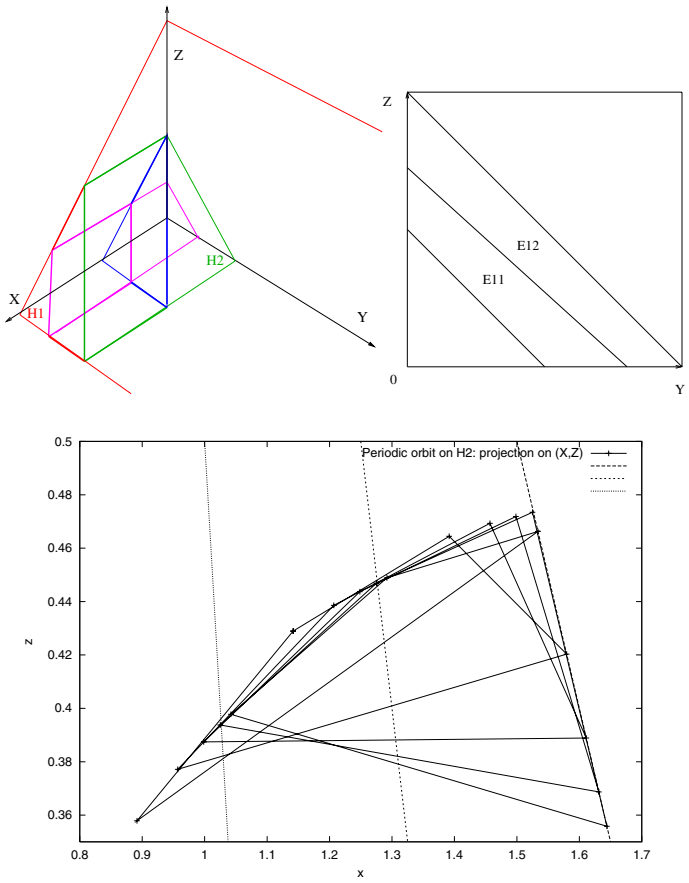

Fig. 5. Example 2. Left Top: Non symmetrical billiards. Right Top: Linearity domains. Bottom: Period

1) Example 3: Consider the three class, two router network of Fig. 2 with $R_x = R_y = R_z = 1$, $p_x = p_y = p_z = 1/2$ but this time with $C_2 = 10^5, C_1 = C$. Fig. 6 gives the period of the billiards process as a function of C.

Fig. 6 suggests that the period achieves constant integer values on a Cantor type set of the horizontal axis; in addition, this figure gives numerical evidence that there are infinitely many values of C for which either the right or the left limit of the period is infinite. For instance the period is constant and equal to 2 on the interval $[2, C_0)$ with $C_0 \sim 104295$, whereas the right limit of the period at this point seems to be $+\infty$.

The impact of this phenomenon on average throughput is exemplified on Fig. 7, where we plot the mean throughput w.r.t. C in the neighborhood of C_0. Class 1 takes advantage of the increase of C; there is no such monotonicity for class 2 nor for class 3. Notice the very irregular shape of mean throughput (which is itself a fractal as shown by the zoom) and the singularity at C_0.

2) Example 4: This is the 6-class, 3-router network of Fig. 8 also with $R = 1$, $p = 1/2$. The default value for the speed of a router is $C = 10^5$. Fig. 9, where we plot the successive values achieved by the throughput of a given class over the period as a function of C. This figure shows that the set of achieved values lives on a fractal.

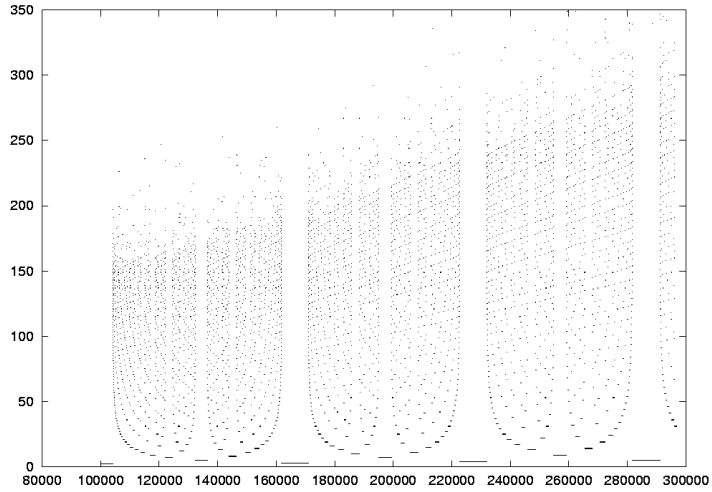

Fig. 6. Billiards Period as a Function of C

C. Fairness

1) Single Router, Several RTTs: Assume that a single router is shared by several classes that only differ through their RTT. Let R_i denote the RTT of class i and p_i its synchronization rate (that we will later take as a function of the average rate). It follows from (7) that a typical flow of class i satisfies the stochastic recurrence: $X_{n+1}^{(i)} = \gamma_{n+1}^{(i)}(X_n^{(i)} + \frac{\overline{\tau}_n}{R_i^2})$ where the sequence $\{\gamma_n^{(i)}\}$ is i.i.d. As a consequence of results in [12], $\overline{\tau}_n$ then converges to a constant $\overline{\tau}$. Taking expectations in the last equation determines the stationary throughput at congestion epochs ($\overline{X}^{(i)}$). Within this setting, the stationary throughput in continuous time is obtained from $\overline{X}^{(i)}$ via the relation $\lambda_i = \overline{X}^{(i)} + \overline{\tau}/2R_i^2$ (see Section 4.1 in [4]). Elementary manipulations give:

$$\lambda_i = \left(\frac{\overline{\gamma}^{(i)}}{1-\overline{\gamma}^{(i)}} + \frac{1}{2}\right)\frac{\overline{\tau}}{R_i^2} = \frac{1}{2}\frac{1+\overline{\gamma}^{(i)}}{1-\overline{\gamma}^{(i)}}\frac{\overline{\tau}}{R_i^2} = \frac{4-p_i}{2p_i}\frac{\overline{\tau}}{R_i^2}.$$

So for all i, j, we have:

$$\frac{\lambda_i}{\lambda_j} = \frac{R_j^2}{R_i^2}\frac{(4-p_i)p_j}{(4-p_j)p_i} \sim \frac{R_j^2}{R_i^2}\frac{p_j}{p_i}, \tag{11}$$

where the last equivalence is when the synchronization rates are small. If we assume that synchronization probabilities are proportional to the rate λ_i, i.e., $\frac{p_i}{p_j} = \frac{\lambda_i}{\lambda_j}$, then throughput is proportional to the inverse of RTT (cf. [13], [18], [7], [21]). If we assume that p_i does not depend on the throughputs, we get throughputs proportional to the square of the inverse of RTT.

Let us now concentrate on the RD case. If one takes

$$p_i = \beta(1-\exp(-\lambda_i\delta)), \tag{12}$$

as suggested in Formula (7) of [6], then the stationary throughputs should satisfy the fixed point equation:

$$\frac{\lambda_i}{\lambda_j} = \frac{R_j^2}{R_i^2}\frac{(1-\exp(-\lambda_j\delta))}{(1-\exp(-\lambda_i\delta))}. \tag{13}$$

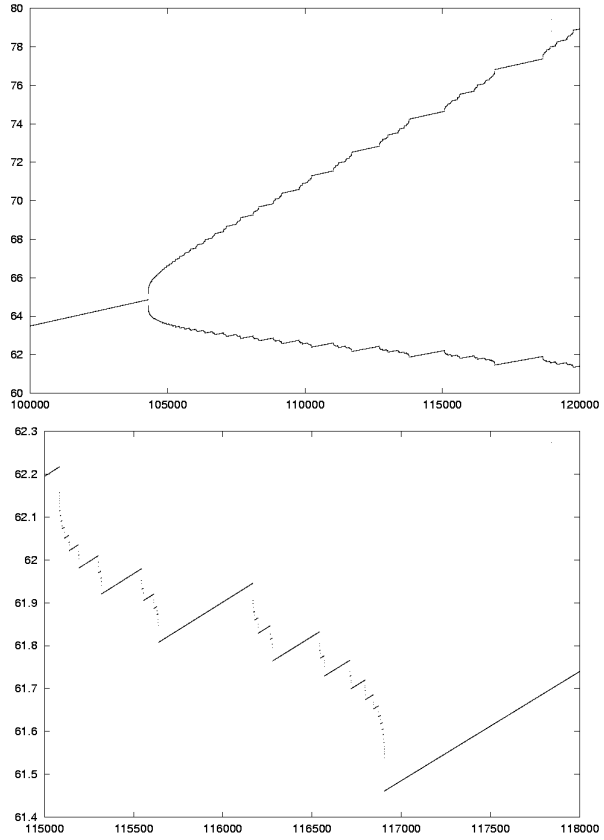

Fig. 7. Top: Mean Throughput of Class 1 (upper curve) and Class 2 (lower one) of Example 3. Bottom: Zoom for Class 2. The y axis is scaled down by a factor of 10^3

If $R_i < R_j$, then $\lambda_i > \lambda_j$, and hence $p_j < p_i$. In addition, from (12), $p_i/p_j < \lambda_i/\lambda_j$. Therefore we always have:

$$\frac{R_j}{R_i} < \frac{\lambda_i}{\lambda_j} < \left(\frac{R_j}{R_i}\right)^2. \tag{14}$$

This confirms experimental studies (cf. [14]) which suggest that the ratio λ_i/λ_j is always proportional to $(R_j/R_i)^a$ with $1 < a < 2$. Let us identify the possible values of a from our analytical framework. When $\lambda\delta$ (defined in (12)) is small, we see that (14) is valid indeed with a close to 1; since δ in (12) is common to all flows, $\lambda\delta$ will be small for the slow flows (here those with large RTTs). Similarly, for those sources with $\lambda_i\delta$ large enough (the fastest flows, or equivalently here those with small RTTs), p_i is close enough to 1, and hence a is close to 2. So, if there is a large enough range of RTT's, the logarithm of the stationary throughput should be a linear function of the logarithm of the RTT, with a slope that is close to -1 for small throughputs, and close to -2 for larger throughputs.

2) Two Routers, Several RTTs: Let us revisit Example 1 (§III-A.3.a) with some more general parameters. The RTT of class i is R_i and its synchronization rate p_i. Whenever the sequence of facets is periodic with period two, one can then identify the periodic regime from the following set of affine

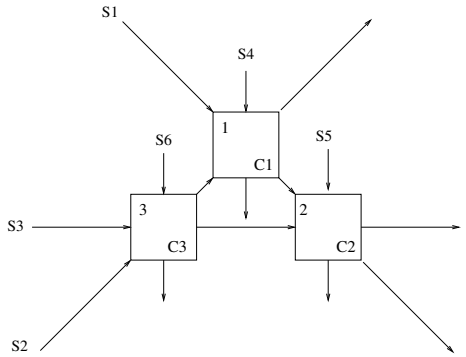

Fig. 8. Example 4: Triangle Network with 6 Classes

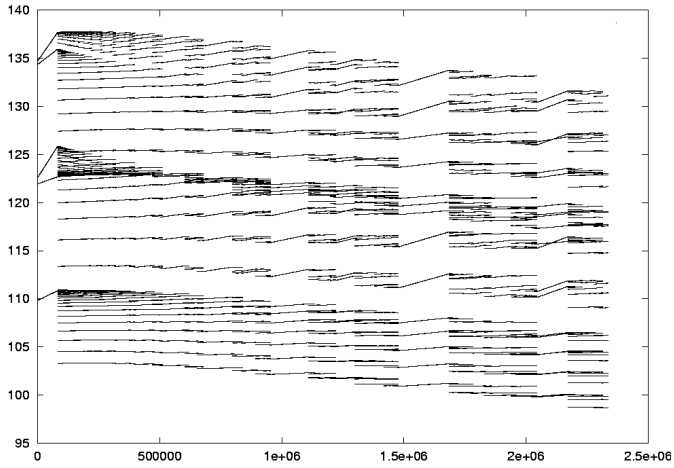

Fig. 9. Instantaneous Throughput as a Function of C for Ex. 4

equations:

$$\mu_x = (\mu'_x + \frac{\overline{\tau'}}{R_x^2}),\ \mu_y = \overline{\gamma}^y(\mu'_y + \frac{\overline{\tau'}}{R_y^2}),\ \mu_z = \overline{\gamma}^z(\mu'_z + \frac{\overline{\tau'}}{R_z^2})$$

$$\mu'_x = \overline{\gamma}^x(\mu_x + \frac{\overline{\tau}}{R_x^2}),\ \mu'_y = (\mu_y + \frac{\overline{\tau}}{R_y^2}),\ \mu'_z = \overline{\gamma}^z(\mu_z + \frac{\overline{\tau}}{R_z^2}).$$

Direct calculations lead to:

$$\lambda_x = \frac{4-p_x}{2p_x}\frac{T}{R_x^2}, \quad \lambda_y = \frac{4-p_y}{2p_y}\frac{T}{R_y^2},$$

$$\lambda_z = \frac{2}{p_z(4-p_z)}\frac{T}{R_z^2}\left(\frac{8-4p_z+p_z^2}{4} - \frac{T'p_z^2}{2T^2}\right)$$

with $T = \overline{\tau} + \overline{\tau'}$ and $T' = \overline{\tau}\,\overline{\tau'}$. We see that for the ratio λ_x/λ_y, the result is as the single router case, whereas

$$\lambda_z/\lambda_x = \frac{R_x^2}{R_z^2}\frac{4p_x}{p_z(4-p_z)(4-p_x)}\left(\frac{8-4p_z+p_z^2}{4} - \frac{T'p_z^2}{2T^2}\right).$$

Hence

$$\frac{\lambda_z}{\lambda_x} = \frac{R_x^2}{R_z^2}\frac{p_x(4-p_z)}{p_z(4-p_x)} \times \alpha,$$

with $1/3 \le \alpha \le 1/2$. This means that even if the flow that crosses the two routers had the same RTT as the two others, in the best case (p proportional to λ) this flow is 30% slower than the two others; in the worst case ($p_x = p_y = p_z$), it could be 3 times slower than the others. These theoretical bounds are quite realistic: cf. Table 2 in [6], where NS simulation with $N_i = 10$ gives a slow down ranging from a 20% slowdown to 5.3 times slower. Now if its RTT is twice larger than that of the other flows, then the best it can get is 3 times slower compared to the others, and in the worst case its connection is 12 times slower!

3) Fairness in the Non-Periodic Case: The aim of this section is to analyze bandwidth sharing as a function of the network parameters, and in particular the speed of the routers.

The following 3 figures illustrate bandwidth sharing for Example 4. We plot the throughput obtained by certain classes against that obtained by other classes, when varying the speed C_1 of router 1 on some interval. We do this both for mean throughput and for instantaneous throughput (the set of values achieved by the throughput process over its period, sampled at the hitting times of a certain face).

In Fig. 10, we plot the sum of the mean throughput of all classes that use router 3 (classes 2,3 and 6) against the mean throughput of the 2-hop class that does not use router 3 (class 1) for C_1 ranging from 7300 to 8700 appr. We again observe a fractal and a non-monotonic behavior. Notice the similarity with the shape obtained for the same kind of functions from a packet level model of window flow control over a two router network in [2].

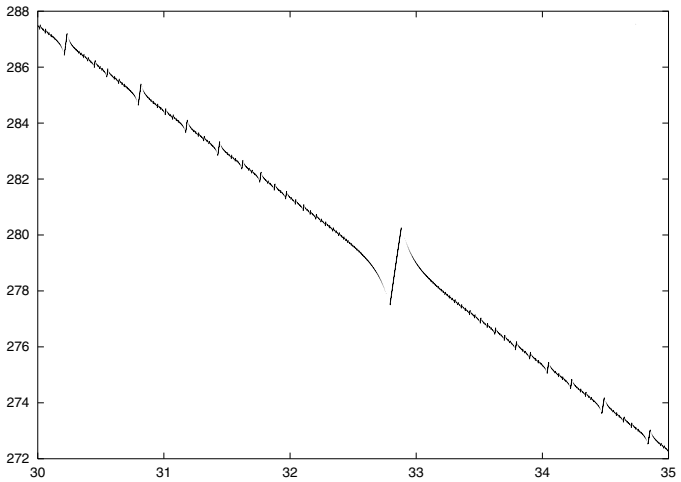

Fig. 10. Example 4: Sum of the Mean Throughputs of Classes 2,3 and 6 w.r.t. that of Class 1

The upper part Fig. 11 plots the sum of the instantaneous throughput of classes 1 and 4 w.r.t. the instantaneous throughput obtained by class 2 (these three classes are those sharing router 1). The lower part is a zoom. Here also, we find a general trend, but a quite complex fractal behavior along this trend, which leaves little hope for simple closed form formulas.

When playing with parameters, such fractals show up in all topologies (not reduced to one router) with a wide

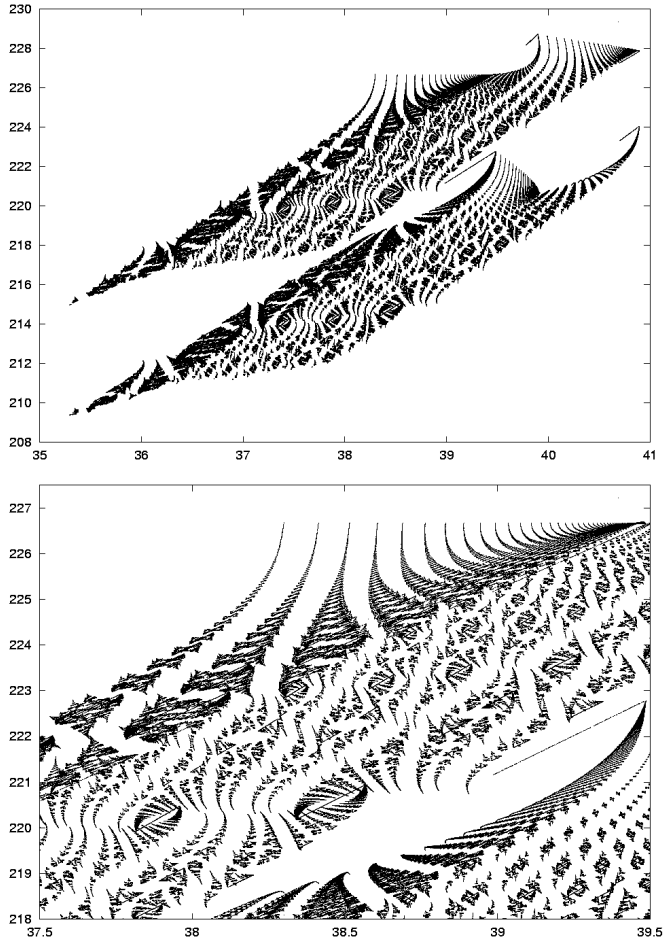

Fig. 11. Instantaneous Bandwidth Sharing for Example 4

variety of shapes. A collection of fractals generated by this class of interaction models can be downloaded at http://www.di.ens.fr/∼trec/aimd

D. A Conservation Law

Assume a stochastic TCP billiards admits a stationary regime. Assume in addition that its synchronization rate is rate and class-independent (see §II).

Let ν_r denote the (continuous time) intensity of the congestion epochs of router r. Let $S(t) = \sum_{i,s} X^{(s,i)}(t)$ be the sum of all flow throughputs in continuous time.

- But for a denumerable set of discontinuities, $S(t)$ is linearly increasing with the rate $\sum_s N_s/R_s^2$.
- Because of the class-independent assumption, each type r congestion epoch creates a jump of $S(t)$ downward of mean magnitude $C_r(1-\overline{\gamma}^r)$.

The drift upward should compensate the jumps downward, so that the following conservation law necessarily holds (see the rate conservation principle in e.g. [3]):

$$\sum_{s\in\mathcal{S}} \frac{N_s}{R_s^2} = \sum_{r\in\mathcal{R}} \nu_r C_r (1-\overline{\gamma}^r). \tag{15}$$

This implies the following relation for the associated deterministic billiards (ν_r has the same interpretation as above):

$$\sum_{r\in\mathcal{R}} \sum_{s\in\mathcal{S}_r} \frac{a_{s,r}}{R_s^2} = \sum_{r\in\mathcal{R}} \nu_r c_r (1-\overline{\gamma}^r). \tag{16}$$

E. Implications

The long periodic or non-periodic behaviors illustrated in §III-B and III-C.3 only hold for the large population asymptotic model with $N=\infty$. In order to capture the behavior of any model with finite population, one should add small Gaussian fluctuations to this, which results into a blur of the dynamics and of the limiting sets describing bandwidth sharing and throughput. Such a blur is illustrated by Fig. 12, where we plot the trajectories of the empirical means $x_n^{(s)}[N] = (\sum_{i\in s[N]} X_n^{(s,i)}[N])/(N_s[N])$, for multi-AIMD models with the same characteristics but for the population parameter N. The figure shows how the stochastic trajectories of the empirical means concentrate and converge to the trajectory of the limiting infinite population model $x_n^{(s)} = x_n^{(s)}[\infty]$, when N grows large (the period of the large population asymptotic billiards is 19). These results indicate that for

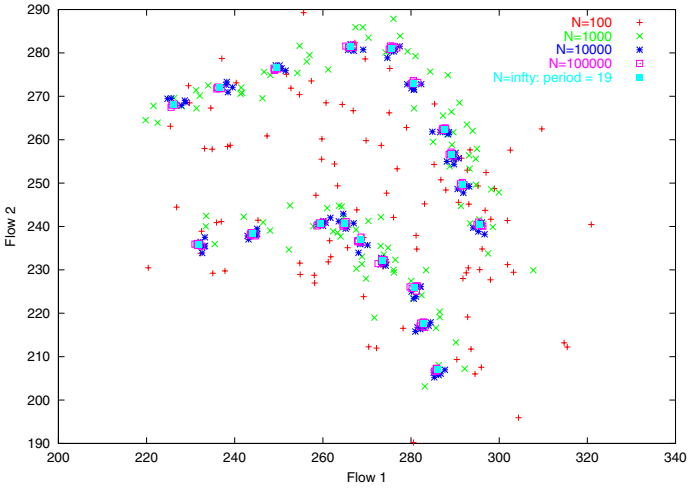

Fig. 12. Concentration toward the Large Population Asymptotic as the Population Grows Large

large populations networks, when taking synchronization into account, *aggregated throughputs* (empirical means) exhibit fluctuations that are due to the network as a whole and that follow some complex fractal pattern.

IV. Conclusion

We have introduced a model allowing one to study the bandwidth sharing operated by TCP on networks composed of several tail-drop routers or links and which takes source synchronization into account.

This model is based on the interplay between three (sub) models: a deterministic network level model (the billiards), a set of more or less independent stochastic models for individual flows, where the influence of the whole network

is taken into account via certain averages, and a packet level model that is only used for determining the delay of reaction of sources and the associated synchronization rates.

Each level creates its own type of fluctuations on throughput. The flow level and packet fluctuations have already been studied. Flow level fluctuations of throughput aggregates have for instance been studied in [4], [12] and for the throughput of individual flows in [9]. The representation of TCP controlled networks as billiards has allowed us to quantify the fluctuations at network level, that is the time/space location of the congestion epochs, these local *storms* that are induced by the AI process here and there in the network, and that are in charge of the overall control within the drop tail setting. We produced numerical evidence that both periodic and non-periodic asymptotic behaviors are possible for the empirical mean values of throughputs and that any slight changes in the model parameters, for instance trends in the population parameters $a_{s,r}$ or n_s, could result into drastic changes for the instantaneous values achieved by empirical averages.

In [20], [24], [1], it was experimentally shown that aggregated traffic of internet traces often exhibits multifractal scaling properties at short time scales. There have been many attempts to explain these properties and to link them to TCP itself (for short time scales, TCP is the only likely explanation, whereas long time scale fluctuations have been shown to be HTTP induced). Such a multifractal scaling means that aggregated traffic trajectories are extremely irregular and have fluctuations at all short time scales. The identification of network level fluctuations with complex patterns is one more step in the direction of the classification of all types of TCP induced fluctuations that contribute to the extreme irregularity of aggregated traffic. Flow level variability seems to be enough to provide a multifractal short time scale behavior [4]. However the fact that aggregated traffic could have fluctuations with arbitrarily long periods is also a possible explanation for time and flow averages to have fluctuations over several time scales. The combination of fluctuations of all three levels seems required for a complete prediction of the global statistical structure of aggregated traffic.

The properties reported in the present paper are different from the simulation based observations on the chaotic behavior of TCP reported in [23]. The main difference lies in the fact that the properties of the present paper bear on the sensitivity of aggregated traffic (the empirical mean values as defined above) w.r.t. some topology parameter (e.g. the speed of a router), whereas the observations of [23] focus on the dependence of the throughput of individual flows w.r.t. initial conditions for a given topology. There might however be a link between the sensitivity w.r.t initial conditions and the fact that the facet and billiards could have a non-periodic behavior for a given topology.

We intend to continue exploring this class of models and to try to enrich it with other types of traffic than the long lived TCP sessions on which this first step is focused.

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] Baccelli,F. and Bonald, T. (1999) Window flow control in FIFO networks with cross traffic. *Queueing Systems*, 32, 195-231.

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

[4] Baccelli, F. and Hong, D. (2002) AIMD, Fairness and Fractal Scaling of TCP Traffic. *Proc. of INFOCOM'02*, New York, July.

[5] Baccelli, F. and Hong, D. (2002) Interaction of TCP Flows as Billiards, *INRIA Report*, RR-4437, INRIA Rocquencourt, April,

[6] Baccelli, F. and Hong, D. (2003) Flow Level Simulation of Large IP Networks, *INFOCOM'03*, San Francisco, April.

[7] Bonald, T. and Massoulié, L. (2001) Impact of fairness on Internet performance. *SIGMETRICS*, pp. 82-91.

[8] Bu, T. and Towsley, D. (2000) Fixed Point Approximation for TCP Behavior in an AQM Network. *Proc. ACM SIGMETRICS*, vol. 29, no. 1, pp. 216-225.

[9] Chaintreau, A. and De Vleeschauwer, D. (2002) A Closed Form Formula for TCP Traffic Performance. *Performance Evaluation*, vol. 49, pp. 57-76, September.

[10] Gibbens, R. and Kelly, F. (1999), Resource Pricing and the Evolution of Congestion Control. *Automatica*, 35, pp. 1969-1985, 1999.

[11] Graham, C. and Méléart, S. (1994) Chaos Hypothesis for a System Interacting through Shared Resources, *Probability Theory and Related Fields*, 100(2), pp.157-174.

[12] Hong, D. and Lebedev, D. (2001) Many TCP User Asymptotic Analysis of the AIMD Model. *Technical Report*, RR-4229, July.

[13] Kelly, F., Maulloo, A. and Tan, D. (1998) Rate control for communication networks: shadow price, proportional fairness and stability. *J. of the Operational Research Society*, 49, pp. 237-252.

[14] Lakshman, T.V. and Madhow, U. (1997) The performance of TCP/IP for networks with high bandwidth-delay products and random loss. *IEEE/ACM Trans. on Networking*, 5-3, pp. 336-350.

[15] Low, S.H., Paganini, F., Wang J., Adlakha S., Doyle J. (2001) Dynamics of TCP/AQM and a Scalable Control. *Proc. of the 2001 Allerton Conference*, University of Illinois, Oct.

[16] Massoulié, L. and Roberts, J. (1999) Bandwidth sharing: objectives and algorithms. *Proc. of INFOCOM*, New York.

[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.

[18] Mo, J. and Walrand, J. (2000) Fair End-to-End Window-based Congestion Control. *IEEE Tr. on Networking* 8-5, pp. 556-567.

[19] Mori, M. (1995) Zeta Functions and Perron-Frobenius Operator for Piecewise Linear Transformations on $\mathbb{R}^k$. *Tokyo Journal of Mathematics* 18, pp.401-416.

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

[21] Roberts, J. and L. Massoulié. (1998) Bandwidth sharing and admission control for elastic traffic. *ITC Specialist Seminar*, Yokohama, October.

[22] Sinai, Ya. G. (1994). Topics in Ergodic Theory, *Princeton Mathematical Series*.

[23] Veres, A. and Boda, M. (2000) The Chaotic Nature of TCP Congestion Control, *Proc. of IEEE INFOCOM*, Tel Aviv.

[24] 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.