

# The autocorrelation of double intermittency maps and the simulation of computer packet traffic

M. Barenco\*      D.K. Arrowsmith†

November 13, 2003

## Abstract

Since the discovery in the early 1990's that packet traffic exhibits long-range-dependence there has been considerable interest in finding system that could generate reliable pseudo-traffic traces. One such system is a one-dimensional chaotic map. We examine a particular instance of such maps. It is non-smooth, but its nature is such that analytical results are easier to obtain. In particular we derive a theorem about the asymptotics of the autocorrelation. In the subsequent discussion, it is argued that to get a relevant description of packet traffic, one needs to venture beyond first- and second-order statistics.

## 1 Introduction

In the beginning of the 1990's, Leland and co-workers showed that packet traffic in computer networks exhibited long-range dependence [10, 9, 11], when it was previously thought that data traffic, like voice traffic, would be memoryless. This has major consequences as long-range

---

\*CoMPLEX, University College London, United Kingdom, to whom all correspondence should be addressed.

E-mail: m.barenco@ucl.ac.uk. Tel.:++44 20 76795088. Fax.:++44 20 73835519

†School of Mathematical Sciences, Queen Mary, University of London, United Kingdom.

traffic is also *bursty*, which poses major engineering challenges, at times scales spanning several orders of magnitude.

In order to address these challenges it is necessary to have at one's disposal models that can mimic packet traffic. One such model has been proposed by Erramilli *et al.* [3, 4]. The appeal of Erramilli's model lies in its simplicity; it is a one-dimensional chaotic map  $f$  defined on the unit interval by

$$f(x) = \begin{cases} x + \frac{1-d}{d^{m_1}} x^{m_1} & \text{if } 0 \leq x < d \\ x - \frac{d}{(1-d)^{m_2}} (1-x)^{m_2} & \text{if } d \leq x \leq 1, \end{cases}$$

where  $d, m_1, m_2$  are parameters<sup>1</sup>; a graph of the map is shown in figure 1. To produce pseudo-traffic traces, one runs an orbit of this map. Every time the orbit visits the interval  $[d, 1)$ , then a packet (or "1") is emitted, otherwise no packet (a "0") is emitted. As one can see, binary sequences can be produced easily with this very light setup.

The chaotic nature of the map guarantees the pseudo-randomness of the sequences it produces so that no stochastic term needs to be introduced in the model. Furthermore, the map has two *non-hyperbolic* unstable fixed points, a characteristic which enables the binary sequences to exhibit long-range dependence<sup>2</sup>.

However, Erramilli encountered great difficulties in obtaining analytical results from this map. In particular, there is no closed-form formula for the natural invariant density. Consequently, there is no non-numerical way to know which expected load (or average number of packets) the map is going to produce under a particular set of parameters. The map has also some curvature and this causes the successive lengths of spells of the same symbol to be statistically dependent from one another<sup>3</sup>.

The latter problem was solved in [13], by introducing a "random wall" between the two

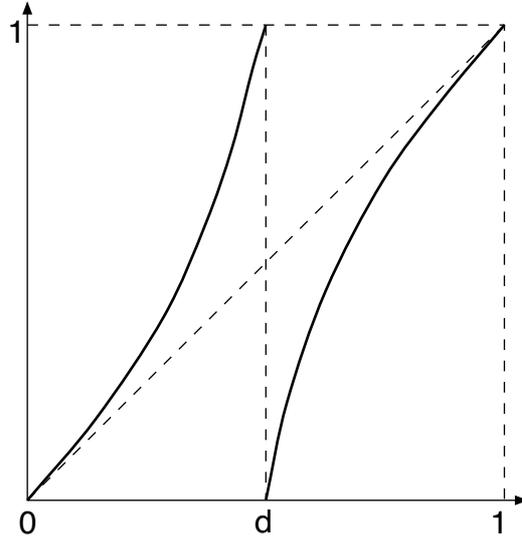
---

<sup>1</sup>Note that  $m_1, m_2 \in (3/2, 2)$  and  $d \in (0, 1)$ .

<sup>2</sup>A system with such a fixed point exhibits *intermittency*; note that in our particular case one such fixed point would be enough to produce long-range dependent binary sequences.

<sup>3</sup>The greater the length of a particular spell of a symbol, the shorter -on average- the next spell (of the other symbol) will be.

Figure 1: The Erramilli map



branches of the map, but this did not completely solve the invariant density problem. By slightly modifying the function  $f$ , it is possible to produce a system for which the invariant density is very easily determined. This approach has been proposed by the Chaos Engineering Group in Ferrara/Bologna [6, 7] and consists of rendering  $f$  piecewise linear. It is this setup which is the object of this article and is considered in the next section.

## 2 Description of the map and its characteristics

In this section, we describe the map in detail, then state the main theorem and give a sketch proof of it.

**Definition of the map.** The map is defined *via* two positive decreasing sequences  $\{z_i^L\}$  and  $\{z_i^R\}$ ,  $i = 1, 2, \dots$ . Both sequences are normalised, with  $z_1^L = z_1^R = 1$ , their respective sums exist and are given by

$$S_* = \sum_{i=1}^{\infty} z_i^* < \infty,$$

where the symbol “\*” is a wildcard and denotes, as in the rest of the paper, either  $L$  or  $R$ . We derive two further sequences  $\{\Delta_i^L\}$  and  $\{\Delta_i^R\}$  as

$$\Delta_i^* = z_i^* - z_{i+1}^*, \quad i = 1, 2, \dots$$

The unit interval is divided into subintervals  $J_i^*$  defined by

$$\left. \begin{aligned} J_i^L &\equiv \left( \frac{z_{i+1}^L}{2}, \frac{z_i^L}{2} \right] \\ J_i^R &\equiv \left( 1 - \frac{z_i^R}{2}, 1 - \frac{z_{i+1}^R}{2} \right] \end{aligned} \right\} i = 1, 2, \dots$$

We want our map  $f$  to be affine on each subinterval, and also require

$$\begin{aligned} f(J_i^*) &= J_{i-1}^*, \quad i = 2, 3, \dots \\ f(J_1^*) &= \bigcup_{i=1}^{\infty} J_i^{\bar{*}}, \end{aligned}$$

where  $\bar{*} = R$  if  $* = L$  and vice-versa. To satisfy these conditions, the map  $f : [0, 1] \rightarrow [0, 1]$  is given by<sup>4,5</sup>

$$f(x) = \begin{cases} 0 & \text{if } x = 0 \\ \frac{\Delta_{i-1}^L}{\Delta_i^L} \left( x - \frac{z_{i+1}^L}{2} \right) + \frac{z_i^L}{2} & \text{if } x \in J_i^L \\ \frac{\Delta_{i-1}^R}{\Delta_i^R} \left( x - 1 + \frac{z_{i+1}^R}{2} \right) + 1 - \frac{z_i^R}{2} & \text{if } x \in J_i^R \\ 1 & \text{if } x = 1. \end{cases}$$

A sketch of the map is shown in figure 2. By construction it is isomorphic to a Markov chain whose graph is pictured in figure 3.

The state  $*_i$  in the Markov chain corresponds to the interval  $J_i^*$  in the map. If  $Y_t$  is the state at time  $t$ , the non-zero transition probabilities in the Markov chain are given by

$$\begin{aligned} P(Y_{t+1} = *_i | Y_t = *_{i+1}) &= 1 \\ P(Y_{t+1} = *_i | Y_t = \bar{*}_1) &= \Delta_i^*, \end{aligned}$$

---

<sup>4</sup>We let  $\Delta_0^* = 0$  in the equation below.

<sup>5</sup>In practice, when the map is actually run on a computer, these equations are not used. The relative position of the orbit within each subinterval is invariant as long as the orbit is on the same branch. Computations have to occur only when there is a crossing from one branch to the other.

Figure 2: The piecewise linear map

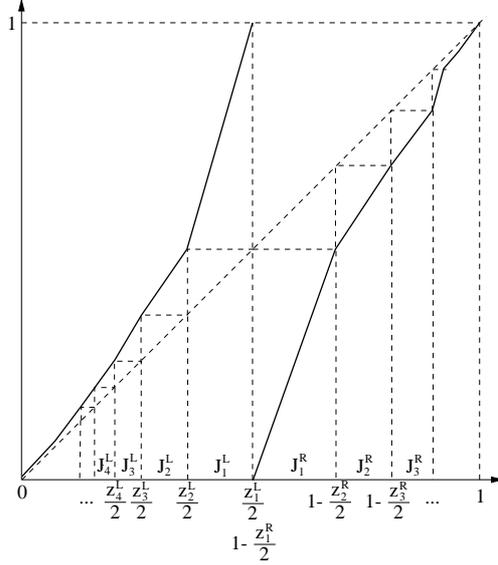
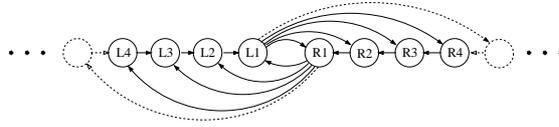


Figure 3: The equivalent Markov chain



for all  $i = 1, 2, \dots$ . This setup is meant to simulate packet traffic. This is done via the binary observable  $X_t$  defined by

$$X_t = \begin{cases} 0 & \text{if } x \in J_i^L \text{ (or if } Y_t = L_i) \\ 1 & \text{if } x \in J_i^R \text{ (or if } Y_t = R_i), \end{cases}$$

so that when  $X_t = 1$ , this means a packet is emitted at time  $t$ , while *no* packet is emitted if  $X_t = 0$ .

**Elementary statistics.** Whenever the orbit crosses from one side to the other (from the right to the left, say), the probability to land in the  $i$ th interval  $J_i^L$  is  $\Delta_i^L$ . From that observation, one can compute the *expected* sojourn time in the left branch of the map, it is

$$\sum_{i=1}^{\infty} i \Delta_i^L = \sum_{i=1}^{\infty} z_i^L = S_L,$$

similarly the expected sojourn time on the right is given by  $S_R = \sum_{i=1}^{\infty} z_i^R$ . The expected number of packets (or average load  $\mu$ ) is

$$\mu = \frac{S_R}{S_R + S_L}.$$

The quantity  $S_R + S_L$  can be interpreted as being an “expected cycle length”, i.e. the average number of iterations it takes for an orbit to cross from one side to the other and then back again.

The interval  $J_1^L$  is visited once in each such cycles once so that the probability that the interval being visited at any time  $t$  is  $J_1^L$ , is given by the inverse of this cycle length, i.e.

$$P(x \in J_1^L) = P(Y_t = L_1) = \frac{1}{S_R + S_L} = P_1.$$

By a symmetric argument, we have

$$P(x \in J_1^R) = P(Y_t = R_1) = \frac{1}{S_R + S_L} = P_1.$$

It can be verified that the general formula for the *stationary probability* is

$$P(x \in J_i^*) = P(Y_t = *_i) = P_1 z_i^*, \quad i = 1, 2, \dots$$

The above formula is the stationary probability for the Markov chain, as well as the invariant *measure* for the map associated with the interval  $J_i^*$ . As the map is Markov affine, the invariant *density*  $\varphi$  is constant on each subinterval and can thus be simply obtained by dividing the measure associated with a given subinterval by its width, we have

$$\varphi(x) = \frac{P_1 z_i^*}{\Delta_i^*}, \quad x \in J_i^*.$$

Finally, since  $X_t$  is binary, we have  $E(X_t) = E(X_t^2)$ , hence the *variance* of  $X_t$  is

$$\text{Var}(X_t) = E(X_t^2) - E^2(X_t) = \mu(1 - \mu).$$

**An exact formula for the autocorrelation.** The autocorrelation  $\rho(k)$  with lag  $k$  computed on the binary variable  $X_t$  is

$$\rho(k) = \frac{E(X_t X_{t+k}) - E(X_t)E(X_{t+k})}{\sqrt{\text{Var}(X_t)\text{Var}(X_{t+k})}}.$$

Since  $E(X_t)$  is constant with respect to  $t$ , we have

$$\rho(k) = \frac{E(X_t X_{t+k}) - \mu^2}{\mu(1 - \mu)},$$

so that we can focus on the term  $E(X_t X_{t+k})$ , the uncentered covariance. Since  $X_t$  is binary, the only case which is going to contribute to the expectation is the one where  $X_t = X_{t+k} = 1$ . In other words, the uncentered autocovariance with lag  $k$  is equal to the probability of having a sequence of length  $k + 1$  starting and finishing with a “1”.

It is worth computing this probability for a particular sequence of binary symbols. Consider the 36-long sequence

$$s = \underbrace{11111}_{P_1 z_5^R} \underbrace{00000}_{\Delta_5^L} \underbrace{111111}_{\Delta_6^R} \underbrace{0000}_{\Delta_4^L} \underbrace{11111}_{\Delta_5^R} \underbrace{0000000}_{\Delta_7^L} \underbrace{1111}_{z_4^R}.$$

The probability of obtaining  $s$  (among the other 36-long sequences) is

$$P(s) = P_1 z_5^R \Delta_5^L \Delta_6^R \Delta_4^L \Delta_5^R \Delta_7^L z_4^R.$$

This follows from:

- The initial sequence of five “1”’s requires the orbit to be on the fifth interval (or fifth state) on the right  $J_5^R$  (or  $R_5$ ) at time  $t$ . The corresponding probability, as computed above, is  $P_1 z_5^R$ .
- This initial spell of “1”’s must be followed by exactly five “0”’s, which means that as the orbit crosses on the left, it must land in  $J_5^L$  (or  $L_5$ ). The associated probability for this event is  $\Delta_5^L$ , hence

$$P(11111000001 \dots) = P_1 z_5^R \Delta_5^L.$$

Using the same principle, we keep on multiplying transition probabilities until the last spell of “1”’s is reached.

- The last spell of “1”’s must be *at least* four symbols long, which mean we can land in *any*  $J_i^R$  (or  $R_i$ ) such that  $i \geq 4$ . The transition probability for this event is  $\sum_{i=4}^{\infty} \Delta_i^R = z_4^R$ .

To compute the uncentered autocovariance with lag  $k$ , it is not necessary to generate all symbol sequences starting and finishing with a “1” and then compute the corresponding probability.

Such a method would be very time consuming ( $\mathcal{O}(2^k)$ ). Observe instead that uninterrupted spells of 1's at the extremities of a symbol sequence involve  $z_i^R$ 's while other spells -which we refer to as "transition sequences"- involve  $\Delta_i^*$ 's. Calling  $C(n)$  the uncentered covariance with lag  $n - 1$ , we can use instead the following system of recurrent equations:

$$T_{01}(n, 1) = \sum_{i=1}^{n-1} \Delta_i^L \Delta_{n-i}^R \quad (1)$$

$$T_{01}(n) = \sum_{i=2}^{n-2} T_{01}(i, 1) T_{01}(n - i) \quad (2)$$

$$T_{00}(n) = \Delta_n^L + \sum_{i=2}^{n-1} T_{01}(i) \Delta_{n-1}^L \quad (3)$$

$$Z(n) = P_1 \sum_{i=1}^{n-1} z_i^R z_{n-i}^R \quad (4)$$

$$P_{11}(n) = P_1 \sum_{i=n}^{\infty} z_i^R \quad (5)$$

$$C(n) = P_{11}(n) + \sum_{i=1}^{n-2} T_{00}(i) Z(n - i). \quad (6)$$

In the above,  $T_{01}(n, 1)$  is the total transition probability associated with all transition sequences of length  $n$ , starting with a "0" and finishing with a "1" and with a *single* (hence the second parameter) uninterrupted spell of either symbol. The quantity  $T_{01}(n)$  corresponds to all transition sequences of length  $n$  starting with a "0" and finishing with a "1", regardless of the number of spells of either symbol.  $T_{00}(n)$  is the same as  $T_{01}(n)$ , except that the corresponding sequences should finish with a "0".  $Z(n)$  captures the probability associated with the extremities of a sequence whereby the cumulative length of those two uninterrupted spell of "1"'s is equal to  $n$ . Finally,  $P_{11}(n)$  is simply the probability of having a sequence of length  $n$ , uniquely composed of "1"'s. To help understand these formulas, we reconsider below our initial example. As can be seen, the total probability for the sequence is constructed from the center outwards (the symbols  $\in$  and  $\cup$  should not be interpreted too literally).

$$\begin{array}{c}
 \underbrace{\hspace{10em}}_{\in T_{00}(27)} \\
 \underbrace{\hspace{6em}}_{\in T_{01}(11,1)} \\
 \dots \underbrace{11111}_{B} \underbrace{00000111111000011111}_{T_{01}(20)} \underbrace{000000}_{E} 1111 \dots \\
 \underbrace{\hspace{10em}}_{\in C(36)} \\
 B \cup E \in Z(9)
 \end{array}$$

Note that the set of equations (1)- (6) cuts down the number of operations involved in computing  $C(n)$ ; it is  $\mathcal{O}(n^2)$  and we require  $\mathcal{O}(n)$  storage.

**Creating intermittency.** We need now to be more specific about the two sequences  $\{z_i^L\}$  and  $\{z_i^R\}$  if we want the resulting map to create long range dependent traffic traces. We choose the extra conditions<sup>6</sup> to be the asymptotic constraints

$$\begin{aligned}\Delta_i^L &\sim \alpha_L K_L i^{-(\alpha_L+1)} \\ \Delta_i^R &\sim \alpha_R K_R i^{-(\alpha_R+1)},\end{aligned}\tag{7}$$

where  $0 < K_* < \infty$  and  $\alpha_* > 1$ . Note that the conditions on the  $\{\Delta_i^*\}$  sequences translate into the following condition on the  $\{z_i^*\}$  sequences (see Lemma 2 below):

$$\begin{aligned}z_i^L &\sim K_L i^{-\alpha_L} \\ z_i^R &\sim K_R i^{-\alpha_R},\end{aligned}$$

Note also that these latter conditions are not sufficient on their own.

It is possible to link the exponents  $\alpha_L$  and  $\alpha_R$  in the above conditions to the exponents  $m_1$  and  $m_2$  in the Erramilli setup sketched in the introduction via

$$\alpha_L = \frac{1}{m_1 - 1}, \alpha_R = \frac{1}{m_2 - 1}.$$

These equalities hold only in the neighbourhood of each fixed point. They can be obtained in various ways [1, 15].

Our main result states that the autocorrelation decays are polynomial with respect to the lag and that the rate of decay (the exponent) depends *exclusively* on the rate of decay of the *slowest* of the two sequences. Before stating this result more precisely we outline in the next paragraph some technical results which we will use in the main proof.

**Convolution lemma.** This important lemma describes the asymptotics of a convolution when the two sequences involved decay polynomially.

---

<sup>6</sup>Note that these extra conditions cannot override the more general conditions on the sequences  $\{z_i^*\}$  we already introduced; both sequences still have to be decreasing, etc...

**Lemma 1** Let  $\{a_i\}$  and  $\{b_i\}$  be two sequences with polynomial asymptotics,

$$a_i \sim K_a i^{-\alpha}$$

$$b_i \sim K_b i^{-\beta},$$

where  $K_a, K_b > 0$ ,  $\text{Max}(\alpha, \beta) > 1$  and  $\text{Min}(\alpha, \beta) > 0$ . Moreover, define the convolution  $C_m$  to be

$$C_m = \sum_{k=1}^{m-1} a_k b_{m-k}.$$

Define also, if  $\alpha > 1$ ,  $S_a = \sum_{i=1}^{\infty} a_i$  and, if  $\beta > 1$ ,  $S_b = \sum_{i=1}^{\infty} b_i$ . Then,  $\{C_m\}$  is a sequence with polynomial asymptotics given by

$$C_m \sim \begin{cases} K_a S_b m^{-\alpha} & \text{if } \alpha < \beta \\ K_b S_a m^{-\beta} & \text{if } \alpha > \beta \\ (K_b S_a + K_a S_b) m^{-\alpha} & \text{if } \alpha = \beta. \end{cases}$$

We do not include the proof of this lemma, as it is quite lengthy, but it can be found in [1]. The lemma simply states that the rate of decay of the convolution depends on the slowest of the two sequences involved. The other sequence is present in the form of a coefficient.

**Other results.** In restricting the definition of the sequences  $\{z_i^*\}$  used to define the map  $f$ , we applied the following lemma

**Lemma 2** . Let  $\{z_i\}$  be a bounded sequence with polynomial asymptotics,

$$z_i \sim K i^{-\alpha},$$

where  $K > 0$  and  $\alpha > 1$ . Define  $S_m = \sum_{i=m+1}^{\infty} z_i$ , then

$$S_m \sim \frac{K}{\alpha - 1} m^{1-\alpha}.$$

Proving the above lemma involves elementary analysis which can be found in [1]. The following is more specific to our particular problem and again its proof can be found in [1].

**Lemma 3** Let  $\{z_i^L\}$  and  $\{z_i^R\}$  be decreasing non-negative sequences such that  $z_1^* = 1$  and let

$$S_* = \sum_{i=1}^{\infty} z_i^*,$$

where  $*$  can be either  $R$  or  $L$ . Define  $\forall i \geq 1$ ,

$$\begin{aligned} \Delta_i^* &= z_i^* - z_{i+1}^* \\ T_{01}(i, 1) &= \sum_{k=1}^{i-1} \Delta_k^L \Delta_{i-k}^R. \end{aligned}$$

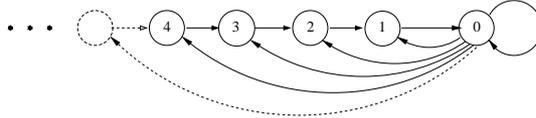
Let

$$S = \sum_{i=2}^{\infty} iT_{01}(i, 1),$$

then  $S = S_R + S_L$ .

**One-sided intermittency** Our case deals with double-sided intermittency, but something is already known about *one*-sided intermittency. This case is studied in [17, 15, 14] and concerns also a Markov affine map. Here, we briefly describe here Wang’s setup, [17]. It is essentially defined by the Markov chain whose graph is given in figure 4 below<sup>7</sup>. The non-zero transition probabilities are indicated with arrows and all are equal to 1, except for those which originate from the “0” state. The latter probabilities are given by  $p_{0N}$ , where  $N$  is the “receiving” state. Wang shows that it is the asymptotics of  $p_{0N}$  (with respect to  $N$ ) which determine the asymptotics of the autocorrelation (w.r.t. the lag).

Figure 4: Wang’s Markov chain



Even though double-sided intermittency is different from its one-sided sibling, we are going to use Wang’s result as a very similar case arises in the course of our proof.

---

<sup>7</sup>As in our case, there is an indicator function on which the autocorrelation is computed. The function takes the value 0 for all states “1”, “2”, etc. and the value 1 for the state “0”.

**The main theorem.** Note that this theorem was conjectured in [6] under a slightly different form involving the Hurst parameter. A similar result was proved in [16] but involved an infinite number of sources.

**Theorem 4** *The asymptotics of the autocorrelation computed on the binary traces produced by the piecewise linear map under the conditions outlined in (7) satisfies*

$$\rho(k) \sim K k^\beta,$$

where the exponent is given by

$$\beta = 1 - \text{Min}(\alpha_L, \alpha_R)$$

and the constant is given by

$$K = \begin{cases} \frac{K_R(1-\mu)}{S(\alpha_R-1)\mu} & \text{if } \alpha_R < \alpha_L \\ \frac{K_L\mu}{S(\alpha_L-1)(1-\mu)} & \text{if } \alpha_L < \alpha_R \\ \frac{1}{S(\alpha_L-1)} \left( \frac{K_R(1-\mu)}{\mu} + \frac{K_L\mu}{(1-\mu)} \right) & \text{if } \alpha_L = \alpha_R. \end{cases}$$

The complete proof of this theorem is given in [1]. The approach is to parse through the set of equations (1)- (6), in that order and for each equation, and determine the asymptotics of the LHS. To do so, we apply the convolution lemma 1 on equations (1), (3) and (4). For the fourth equation in the system, we apply lemma 2, while equation (6) is solved through a combination of lemmas 1 and 2. Equation (2) is more demanding, hence we give below more details on how to treat it.

In (2), we are in principle not interested in  $T_{01}(n)$  and  $T_{01}(n, 1)$  if  $n$  is smaller than 2. But nothing prevents us from letting  $T_{01}(0) = 1$  and  $T_{01}(1, 1) = 0$ , which allows to rewrite (2) as

$$T_{01}(n) = \sum_{i=1}^n T_{01}(i, 1)T_{01}(n-i). \quad (8)$$

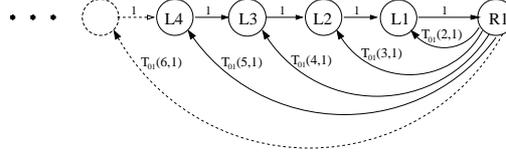
The above is a particular instance of the *renewal equation*, used in the theory of recurrent events (see for example [5]). Under certain conditions<sup>8</sup> we have, as  $n \rightarrow \infty$ ,  $T_{01}(n) \rightarrow (\sum_{i=1}^{\infty} iT_{01}(i, 1))^{-1}$ . Hence, by lemma 3,  $T_{01}(n) \rightarrow 1/(S_R + S_L) = P_1$ .

---

<sup>8</sup>These conditions, persistence and non-periodicity (of the underlying recurrent event) are fulfilled in our case.

But we want to know more about  $T_{01}(n)$ , in particular the rate at which it converges to  $P_1$ . To do so consider the following one-sided Markov chain.

Figure 5: A one-sided Markov chain



By assigning, as in the double-sided case, a value 0 to any of the  $L$  states and a value 1 to the  $R_1$  state, it is possible to show, using the method outlined above for the double-sided case, that the uncentered covariance computed on the binary sequence produced by this chain is given by  $P_1 T_{01}(n)$ .

Now, observe that this Markov chain is quite similar to the one studied by Wang (see figure 4), except for the transition probability from  $R_1$  to itself which, in the above case is zero. However, for the asymptotics of the autocorrelation, only the “far” states play a role so that we can use Wang’s result. It states that, as  $n \rightarrow \infty$ , if the transition probability  $\Delta_n$  for the  $n$ th state is such that  $\Delta_n \sim K n^{-(\alpha+1)}$ , then the autocorrelation  $\varphi(n)$  is given by

$$\varphi(n) \sim K' n^{1-\alpha}.$$

This result can be applied to our case. Firstly note that we have by (1) and lemma 1

$$T_{01}(i, 1) \sim K_1 n^{-(\alpha+1)},$$

where  $\alpha = \min(\alpha_L, \alpha_R)$  and

$$K_1 = \begin{cases} K_R \alpha & \text{if } \alpha = \alpha_R < \alpha_L \\ K_L \alpha & \text{if } \alpha = \alpha_L < \alpha_R \\ (K_L + K_R) \alpha & \text{if } \alpha = \alpha_R = \alpha_L. \end{cases}$$

Now, by Wang’s result we have that

$$T_{01}(i) - P_1 \sim K_2 i^{1-\alpha},$$

where it can be shown that

$$K_2 = K_1 / ((S_L + S_R)^2 \alpha (\alpha - 1)). \quad (9)$$

This last result on the coefficient  $K_2$  can be obtained by considering (8), and defining a new sequence  $z_i$  as

$$z_n = \sum_{i=n}^{\infty} T_{01}(i, 1),$$

So that  $z_1 = z_2 = 1$ , and  $T_{01}(i, 1) = z_i - z_{i+1}$ . We replace the latter in (8), which yields

$$T_{01}(n) = \sum_{i=1}^n (z_i - z_{i+1}) T_{01}(n - i),$$

which, after elementary manipulations can be rewritten as

$$\sum_{i=1}^{n+1} z_i T_{01}(n + 1 - i) = \sum_{i=1}^n z_i T_{01}(n - i).$$

Calling the RHS  $A_n$ , the above reads  $A_{n+1} = A_n$  and it can be verified that  $A_1 = 1$ . Thus for all  $n$ , we have

$$\sum_{i=1}^n z_i T_{01}(n - i) = 1.$$

We know that  $T_{01}(n)$  converges to  $1/(S_R + S_L)$  and are interested in the remainder term  $\varepsilon(n)$  such that  $T_{01}(n) = 1/S + \varepsilon(n)$ . Hence we write

$$\sum_{i=1}^n z_i \varepsilon(n - i) = 1 - \frac{1}{S} \sum_{i=1}^n z_i = \frac{1}{S} \sum_{i=n+1}^{\infty} z_i, \quad (10)$$

where the last equality holds because  $S - \sum_{i=1}^{\infty} i T_{01}(i, 1) = \sum_{i=1}^{\infty} z_i$  (see lemma 3). We use lemma 2 to evaluate the RHS in (10). The exponent of  $\varepsilon_i$  is known via Wang's result but with lemma 1 one can also determine its coefficient and thereby prove (9).

As mentioned before, the remaining equations (3)- (6) can be solved via an appropriate use of lemmas 1 and 2 and this directly leads to the proof of theorem 4.

**Non-hyperbolic dynamics.** In the next section, we discuss the engineering implications of theorem 4 but before that it is worth placing this result in its mathematical context which is non-hyperbolic dynamics. Non-hyperbolic systems are difficult to study because they do

not lend themselves to local linearisation (or first-order approximation). One powerful body of theory addresses the issue of autocorrelation decay in presence of indifferent fixed points. These techniques have been proposed by L.S. Young, they allow to compute *upper bounds* of the correlation decay.

Briefly these bounds can be obtained by the technique of *inducing* the system of interest. Inducing means choosing an appropriate subinterval and constructing the first return map on that subinterval (call this map  $f'$  and the original map  $f$ ). By considering the tail of the distribution of the return time with respect to  $f'$ 's invariant measure (the so-called “Markov towers”) one can obtain the asymptotics of the correlation decay of  $f$  (or, more precisely, its upper bound). A thorough description of these techniques is given in [18], while in [8] concrete examples of one-dimensional maps with one indifferent fixed point are studied under this formalism, including some systems with sub-polynomial decay of correlation.

It is important to note that this theory works under much more general assumptions than those under focus of this article, firstly because it admits more general observable functions<sup>9</sup> and secondly because the systems under scrutiny can exhibit some (bounded) distortion<sup>10</sup>. Let us mention here that the results obtained with our system confirm that L.S. Young’s upper bounds are obviously correct but also -and more interestingly- sharp.

### 3 Discussion

Theorem 4 states that it is the slowest decaying sequence among  $\{z_i^L\}$  and  $\{z_i^R\}$  that governs the rate of decay of the correlation profile. Unfortunately, this does not mean that we can arbitrarily choose for the simulations whether long range dependence can be produced by inactive periods or conversely, by active periods.

The distinction is quite crucial. R. Mondragon *et al.* [12] have shown that a queuing

---

<sup>9</sup>These functions have to be Hölder-continuous. Note that in our case, the binary observable is not even Hölder-continuous, but the discontinuity point at 1/2 is benign in that respect.

<sup>10</sup>Being piecewise linear, our system has zero distortion.

system exhibited a completely different behaviour depending on which kind of incoming traffic it was confronted to, even though this incoming traffic was long range dependent in every case. In the first case, the incoming traffic was characterised by heavy tailed silent periods and exponentially decaying active periods and the resulting queue behaved in the same way as it would have done under Poisson-like conditions: the queue length distribution decayed exponentially. On the contrary, when these characteristics were inverted (heavy tailed active periods/short tailed silent periods), the resulting queue length distribution was itself heavy tailed.

This example simply shows that we require a deeper level of understanding on the nature of packet traffic since two pseudo-traffic traces, both long-range dependent but with a different underlying feature, can produce such different outcomes. In the context of this model this feature concerns the distribution of active and silent periods.

There is strong evidence to suggest the presence heavy-tailed active periods in packet traffic. This presence may in turn be the consequence of several distinct causes. The most immediate of these is the fact that sizes of the files sent over computer networks can span several orders of magnitude- an indicator of heavy tailed distribution. Another cause might be the buffering taking place at several stages in computer networks.

In short, there is a wealth of elements pointing to active periods as an indisputable cause of long-range dependent traffic but it is not a reason to discard silent periods as a possible concomitant cause. For example, M. Crovella *et al.* discuss the physical origins of self-similarity in [2] and show that, while active periods are the dominant cause of self-similarity, there are some signs of heavy-tailed silent periods.

Unfortunately, it might not be possible to measure directly the length of silent periods in the sense intended in this article. This is because in reality time flows in a continuous manner and we postulate a discrete time frame. Consequently, the incidence of idle periods has to be determined in a more tangential manner. Early numerical experiments with the map suggest that a way forward might be to consider higher order statistics.

We designed two maps yielding both the same average load and (long-range dependent) correlation profile. They differed only by the fact that in one of them, long-range dependence was supplied by heavy-tailed silent periods with the length distribution of active periods decaying exponentially while for the other map, it was active periods which created “burstiness”<sup>11</sup>. Aggregation of traffic traces over increasing time scales showed clearly that the distribution was heavy tailed only on one side; on the right for the setup with active heavy-tailed periods and on the left for the symmetric setup. By design, this difference could not have been detected by considering first and second-order statistics while it was by measuring the skewness, a third-order statistic. We believe this strongly suggests a possible way forward in better understanding of traffic traces.

## References

- [1] M. Barenco. *Packet Traffic in Computer Networks*. PhD thesis, Queen Mary, University of London, 2002. <http://www.maths.qmul.ac.uk/msb/nm2b.pdf>.
- [2] M. E. Crovella and A. Bestavros. Self-similarity in world wide web traffic, evidence and possible causes. *Proc. ACM Sigmetrics, Philadelphia*, 1996.
- [3] A. Erramilli, R. P. Singh, and P. Pruthi. Chaotic maps as models of packet traffic. *ITC*, 14:329–338, 1994.
- [4] A. Erramilli, R. P. Singh, and P. Pruthi. An application of deterministic chaotic maps to model packet traffic. *Queueing systems*, 20:171–206, 1995.
- [5] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. John Wiley and Sons, third edition, 1950.
- [6] A. Giovanardi, G. Mazzini, and R. Rovatti. Chaos based self-similar traffic generators. *Proceedings NOLTA’00, Dresden, Germany*, pages 747–750, 2000.

---

<sup>11</sup>With an average load of 1/2, it is simply done by swapping both sequences  $\{z_i^L\}$  and  $\{z_i^R\}$ .

- [7] A. Giovanardi, G. Mazzini, R. Rovatti, and G. Setti. Features of chaotic maps with self-similar trajectories. *Proceedings NOLTA '98, Crans-Montana, Switzerland*, pages 203–206, 1998.
- [8] M. Holland. Slowly mixing systems and intermittency maps. *preprint*, 2003.
- [9] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. On the self-similar nature of ethernet traffic. *Proc. ACM Sigcomm '93, San Francisco, CA*, pages 183–193, 1993.
- [10] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Transactions on networking*, 2(1):1–15, 1994.
- [11] W. E. Leland and D. V. Wilson. High time-resolution measurement and analysis of LAN traffic: Implications for LAN interconnection. *Proc. IEEE INFOCOM '91, Bal Harbour, FL*, pages 1360–1366, 1991.
- [12] R. Mondragón, D. Arrowsmith, J. Griffith, and J. Pitts. Chaotic maps for network control: Traffic modelling and queueing performance analysis. *Proceedings, IEEE Performance and Control of Network Systems III, 19-22 September, Boston, 1999*.
- [13] R. J. Mondragón. A model of packet traffic using a random wall model. *International Journal of Bifurcation and Chaos*, 9(7):1381–1392, 1999.
- [14] M. Mori. On the intermittency of a piecewise linear map. *Tokyo Journal of Mathematics*, 16(2):411–426, 1993.
- [15] H. G. Schuster. *Deterministic Chaos*. Weinheim, third augmented edition, 1995.
- [16] M. S. Taqqu, W. Willinger, and R. Sherman. Proof of a fundamental result in self-similar traffic modeling. *ACM SIGCOMM / Computer Communication Review*, 27(2):5–23, 1997.
- [17] X.-J. Wang. Statistical physics of temporal intermittency. *Physical Review A*, 40(11):6647–6661, 1989.
- [18] L.-S. Young. Recurrence times and rates of mixing. *Israel Journal of Mathematics*, 110:153–188, 1999.