A correct version of Maurer's conjecture for ψ-mixing processes *

Miguel Abadi † Antonio Galves ‡

March 14, 2003

Abstract

We consider the number of non-overlapping blocks of n symbols occuring before the initial n-block reappears. For ψ-mixing processes on a finite alphabet, we prove that the difference between the expectation of the logarithm of this number and the entropy of n-blocks converges to the constant of Euler divided by $-\ln(2)$. This can be considered the correct version of a conjecture presented in Maurer (1992). Our theorem generalizes recent results presented in Coron and Nacache (1999), Choe and Kim (2000) and Wegenkittl (2001), in the context of Markov chains. We also prove that the difference between the variance of the logarithm of this number and the variance of the law of the n-blocks converges to an explicit constant as n diverges. The basic ingredient of the proofs is an upper-bound for the exponential approximation of the law of the number of non-overlapping n-blocks until a fixed but otherwise arbitrary n-block reappears. This is a new result which is interesting by itself.

1 Introduction.

Maurer (1992) studied the asymptotic behavior of the number of non-overlapping blocks of n symbols produced by an i.i.d. binary source until the first n-block reappears. He proved that the difference between the expectation of the logarithm of this number and n times the entropy of the source converges to the constant of Euler divided by $-\ln(2)$.

In the same paper he made the conjecture that this result should hold for any binary stationary ergodic source. However, an example showing that this conjecture was not true even for Markov chains was presented in Coron and Naccache (1999).

Running title. Hitting times in mixing processes.

AMS subject classification. 60F05, 60G10, 37A50.

Key words and phrases. Mixing processes, hitting times, repetition times, entropy, Maurer's conjecture.

Work done within the Projeto Temático "Rhythmic patterns, parameter setting and language change", supported by FAPESP (grant 98/3382-0), and as part of the activities of the Núcleo de Excelência "Critical phenomena in probability and stochastic processes" (grant 41.96.0923.00). This research was partially supported by the USP-COFECUB agreement.

†*Supported by FAPESP*

‡*Partially supported by CNPq (grant 301301/79)*

Recently Choe and Kim (2000) and Wegenkittl (2001) proved that a modified version of the conjecture was true for ergodic Markov chains on a finite alphabet. The modification consisted in replacing the entropy of the source by the per-bit entropy of blocks of length n.

In the present paper we prove that this corrected version of Maurer's conjecture holds in the much more general setting of ψ-mixing stationary processes on a finite alphabet. We also prove that the difference between the variance of the logarithm of the number of n-blocks until the initial one reappears converges and the variance of the law of the n-blocks converges to an explicit constant as n diverges. Explicit upper bounds for the rates of convergence are provided for both limits. In particular, Moreover, the convergence to the limit takes place exponentially fast when the rate of ψ-mixing is exponential.

For processes having marginals which are uniformly distributed on the alphabet, the logarithm of the number of n-blocks until the initial one reappears converges to a constant. Again this generalizes Maurer's result for independent binary symmetric random variables. We remark that this is not typically the case. For instance, for independent binary but assymmetric random variables a direct computation shows that the logarithm of the number of n-blocks until the initial one reappears is of order n.

The main tool of our proof is an upper-bound for the exponential approximation of the law of the number of blocks until a fixed but otherwise arbitrary n-block reappears. This is a new result which is interesting by itself. It's novelty is that it holds for any initial n-block, contrarily to previous theorems like, for instance, the one presented in Collet *et al.* (1999) which only holds for typical blocks and the one in Abadi (2003b) showing that the exponential approximation does not hold for atypical blocks. We refer the reader to Abadi and Galves (2001) for a recent review of the the field.

This paper is organized as follows. In section 2 we present the definitions and notation and state the results. Section 3 contains the proof of the exponential approximation for the first time after the origin the process hits any fixed n-block. Section 4 contains the proof of the exponential approximation for the process conditioned to start in the n-block. The proof of the main result is given in section 5.

2 Definitions, notation and main result.

Let $(X_m)_{m\in\mathbb{Z}}$ be a stationary stochastic process taking values on a finite *alphabet* $\mathcal{E}$ and defined on a probability space $(\Omega, \mathcal{F}, I\!P)$. We denote by $\mathcal{F}_I$ the σ-algebra generated by the cylinders with coordinates in I, $I \subseteq \mathbb{Z}$. The short notation $\mathcal{C}_n$ will denote the set of all cylinders with coordinates in $\{0, \ldots, n-1\}$.

Let $\psi = (\psi(l))_{l\geq 0}$ be a sequence decreasing to zero. The process $(X_m)_{m\in\mathbb{Z}}$ is called ψ *-mixing* if

$$\sup_{n\in I\!N, B\in\mathcal{F}_{\{0,.,n\}}, C\in\mathcal{F}_{\{n+l+j, j\geq 1\}}} \frac{|I\!P\{B \cap C\} - I\!P\{B\}I\!P\{C\}|}{I\!P\{B\}I\!P\{C\}} = \psi(l) \ , \qquad (1) \quad \boxed{\texttt{pmix}}$$

where the supremum is taken only for sets B and C, such that $I\!P\{B\}I\!P\{C\} > 0$. The sequence $\psi(l)$ is called the mixing rate of the process.

Let n be a fixed positive integer. The n-block process $(\overline{X}_k)_{k\in\mathbb{Z}}$, taking values on the set of n-blocks $\mathcal{E}^n$, is defined by $\overline{X}_k := \Big(X_{kn}, \ldots, X_{(k+1)n-1}\Big)$.

We remark that if the original process $(X_k)_{k\in\mathbb{Z}}$ is ψ-mixing then the n-block process $(\overline{X}_k)_{k\in\mathbb{Z}}$ is also ψ-mixing, with mixing rate $\overline{\psi}(\ell) = \psi(\ell n)$.

Given a n-block $A = (a_0, \ldots, a_{n-1})$, we define $\overline{\tau_A}$, the first *hitting time* after the origin, as follows

$$\overline{\tau_A} = \inf\{k \geq 1 : \overline{X}_k = A\} \ .$$

We recall that mixing (at any rate) implies ergodicity, and ergodicity implies that τ_A is almost surely finite (see e.g. Cornfeld *et al.* 1982).

We define $\overline{R_n}$ as the first time after the origin that the first n-block reappears, namely,

$$\overline{R_n} = \inf\{k \geq 1 : \overline{X}_k = \overline{X}_0\} \ .$$

To simplify the notation most of time we will write $I\!P\{A\}$ instead of $I\!P\{\overline{X}_j = A\}$. Since the process is stationary, this quantity does not depend on j.

We define the per-bit-entropy of the blocks of length n as

$$H_n := -\sum_{C\in\mathcal{C}_n} I\!P\{C\} \log I\!P\{C\} \ . \qquad (2) \quad \boxed{\texttt{Hn}}$$

The entropy of the process is defined as the limit

$$h = \lim_{n\to+\infty} \frac{1}{n} H_n \, .$$

The following quantities will appear in the statement of our main theorem. We define

$$\Delta := \lim_{\epsilon\to 0} \Big[\widetilde{\omega}(\epsilon) - (\log\epsilon)^2 + 2\ \Xi\ \log\epsilon\Big] \ ,$$

where

$$\widetilde{\omega}(\epsilon) := \epsilon \sum_{k\geq 1} (\log k)^2\ e^{-\epsilon(k-1)}$$

and

$$\Xi = \lim_{\epsilon \to \infty} \int_{\epsilon}^{\infty} e^{-t} \log t \; dt \; . \tag{3}$$ defxi

We can now state the main result of this paper.

teo:ent **Theorem 1** *Let* $(X_m)_{m \in \mathbb{Z}}$ *be a* ψ*-mixing process. For every positive constant* $d > 0$*, there exist two positive constants* C *and* c*, such that the following inequality holds for all* $n \in \mathbb{N}$

$$\left| \; I\!E \left(\log \overline{R_n} \right) - H_n - \Xi \; \right| \leq \varepsilon_n \; , \tag{4}$$ ent

where

$$\varepsilon_n = \begin{cases} Ce^{-cn} & \text{if } \psi(\ell) \leq e^{-d\ell} > 0 \\ \sqrt{\psi(n)} & \text{otherwise} \end{cases} \quad .$$

Moreover

$$\left| \mathrm{Var}(\log \overline{R_n}) - \mathrm{Var}(\log I\!P\{X_1^n\}) - \Delta + \Xi^2 \right| \leq \varepsilon_n \; . \tag{5}$$ var

Theorem 1 implies the following corollary for Markov chains.

markov **Corollary 2** *Let* $(X_m)_{m \in \mathbb{Z}}$ *be an ergodic Markov chain taking values on* $\mathcal{E}$ *with entropy* h*. Then there exist two positive constants* C, c *such that*

$$\left| \; I\!E \left(\log \overline{R_n} \right) - (n-1)h - H_1 - \Xi \; \right| \leq Ce^{-cn} \; .$$

The main ingredients of the proof of Theorem 1 are Theorems 3 and 4 below. Theorem 3 gives an exponential approximation for the first time after the origin the process hits a fixed n-block. Theorem 4 proves the same result for the process conditioned to start in the fixed n-block. We emphasize that both theorems hold for *any* fixed n-block.

teo:hit **Theorem 3** *Let the process $(X_m)_{m\in\mathbb{Z}}$ be ψ-mixing. Let $A \in \mathcal{C}_n$, with n a positive integer. Then, there exist positive constants C_1, C and $\beta \in (0,1)$ such that for all $t > 0$, the following inequality holds*

$$\left| I\!P\left\{\overline{\tau_A} > t\right\} - e^{-I\!P\{A\}t} \right| \leq C\ \varepsilon(A)\ f(A,t)\ , \tag{6}$$ hit

where

$$\varepsilon(A) := \begin{cases} I\!P\{A\}^{\beta} & \psi(\ell) \leq \frac{1}{\ell^d} \ \text{ with } \ d > 0 \\ \sqrt{\psi(n)} & \text{otherwise} \end{cases}\ , \tag{7}$$ varen

and $f(A,t) := e^{-C_1 I\!P\{A\}t}$.

For any event A with $I\!P\{A\} > 0$ we denote by $I\!P\{. \mid A\}$ the conditional probability given A.

teo:ret **Theorem 4** *Let the process $(X_m)_{m\in\mathbb{Z}}$ be ψ-mixing. Then, for all $n \in I\!N$, all $A \in \mathcal{C}_n$ and all $t > 0$, the following inequality holds*

$$\left| I\!P_A\left\{\overline{\tau_A} > t\right\} - e^{-I\!P\{A\}t} \right| \leq C\ \ [\varepsilon(A) + \psi(n)]\ \ f(A,t)\ . \tag{8}$$ ret

where f and $\varepsilon(A)$ are the same of the preceding theorem.

3 Proof of Theorem 3.

Since the n-block process $(\overline{X}_k)_{k\in\mathbb{Z}}$ is ψ-mixing, Theorem 1 in Abadi (2003a) implies that for all $t > 0$ the following inequality holds

$$\begin{aligned} \left| I\!P\left\{\overline{\tau_A} > t\right\} - e^{-\overline{\xi_A} I\!P\{A\}t} \right| &\leq C\ \delta(A)\ e^{-\overline{\xi_A} I\!P\{A\}t}\ \ (t I\!P\{A\} \vee 1) \\ &\leq C\ \delta(A)\ e^{-(\Xi_1/2) I\!P\{A\}t}\ , \end{aligned} \tag{9}$$ ab

where $\delta(A) \to 0$, when $n \to 0$, $\overline{\xi_A} \in [\Xi_1, \Xi_2]$ and $0 < \Xi_1 < 1 \leq \Xi_2 < +\infty$ are constants independent of A, n and t. The same theorem implies that $\delta(A) < \varepsilon(A)$, where $\varepsilon(A)$ is defined in (7).

On one hand, we remark that, by the Mean Value Theorem,

$$\left|e^{-I\!P\{A\}t} - e^{-\overline{\xi_A}I\!P\{A\}t}\right| \leq \left|1-\overline{\xi_A}\right| \; I\!P\{A\} \; t \; e^{-\Xi_1 I\!P\{A\}t} \; .$$

Therefore to prove Theorem 3 it is enough to show that $\left|1-\overline{\xi_A}\right| \leq \varepsilon(A)$.

On the other hand, we also have the following uniform bound

$$\sup_t \left| I\!P\left\{\overline{\tau_A} > t\right\} - e^{-\overline{\lambda_A} I\!P\{A\}t}\right| \leq C \; \varepsilon(A) \; , \tag{10}$$

where $\overline{\lambda_A} \in [\Lambda_1, \Lambda_2] \subset (0,\infty)$, and $\Lambda_1 < 1 \leq \Lambda_2$ are constants, independent of A, n and t (cf. Galves and Schmitt (1997) when ψ is summable and Abadi (2001) when ψ is non-summable).

Inequalities (9) and (10) together imply that

$$\sup_t \left|e^{-\overline{\xi_A}I\!P\{A\}t} - e^{-\overline{\lambda_A}I\!P\{A\}t}\right| \leq C \; \varepsilon(A) \; .$$

Taking $t = 1/I\!P\{A\}$, we conclude that

$$\left|\overline{\xi_A} - \overline{\lambda_A}\right| \leq C \; \varepsilon(A) \; .$$

Taking $C_1 = \Xi_1/2$, it is enough to show that $\left|\overline{\lambda_A} - 1\right| \leq C I\!P\{A\}^\beta$ to conclude the proof.This is done in the next lemma.

Lemma 5 *Let the process $(X_m)_{m\in\mathbb{Z}}$ be ψ-mixing. Then, there exist positive constants C and $\beta \in (0,1)$ such that for all $n \in I\!N$ and all $A \in \mathcal{C}_n$ the following inequality holds*

$$\left| \; \overline{\lambda_A} - 1 \; \right| \leq C I\!P\{A\}^\beta \; .$$

Proof. It turns out that

$$\overline{\lambda_A} = \frac{-\log I\!P\left\{\overline{\tau_A} > f_A\right\}}{f_A I\!P\{A\}} \quad \text{with} \quad f_A = 1/I\!P\{A\}^\alpha,$$

for a suitable $\alpha \in (0,1)$.

By Taylor's expansion we have that

$$1 - e^{-\theta} \leq \theta \leq 1 - e^{-\theta} + 2\left(1 - e^{-\theta}\right)^2 \ ,$$

for all $\theta \in [0,1]$. Put $\theta = -\log I\!P\{\overline{\tau_A} > f_A\}$. Then $\theta < 1$ if and only if $I\!P\{\overline{\tau_A} \leq f_A\} < 1 - e^{-1}$. This last inequality is satisfied since

$$I\!P\{\overline{\tau_A} \leq f_A\} = \sum_{i=1}^{f_A} I\!P\{\overline{\tau_A} = i\} \leq \sum_{i=1}^{f_A} I\!P\left\{\overline{X}_i = A\right\} = f_A I\!P\{A\} \leq 1 - e^{-1} \ ,$$

for n large enough. The most right equality follows by stationarity. Therefore

$$\frac{I\!P\{\tau_A \leq f_A\}}{f_A I\!P\{A\}} \leq \frac{-\log I\!P\{\tau_A > f_A\}}{f_A I\!P\{A\}} \leq \frac{I\!P\{\tau_A \leq f_A\}}{f_A I\!P\{A\}} + 2 f_A I\!P\{A\} \ .$$

Therefore

$$\left| \frac{I\!P\{\tau_A \leq f_A\}}{f_A I\!P\{A\}} - \overline{\lambda_A} \right| \leq C f_A I\!P\{A\} \ .$$

By stationarity

$$\begin{aligned} I\!P\{\overline{\tau_A} = j+1\} &= I\!P\{\overline{\tau_A} > j\} - I\!P\{\overline{\tau_A} > j+1\} \\ &= I\!P\left\{\overline{X}_i \neq A; 2 \leq i \leq j+1\right\} - I\!P\{\overline{\tau_A} > j+1\} \\ &= I\!P\left\{\overline{X}_1 = A \ , \ \overline{X}_i \neq A; 2 \leq i \leq j+1\right\} \\ &= I\!P\left\{\overline{X}_0 = A \ , \ \overline{\tau_A} > j\right\} \ , \end{aligned} \tag{11}$$

rever

for all positive integer j. Therefore

$$\left|1 - \frac{I\!P\{\overline{\tau_A} \leq f_A\}}{f_A I\!P\{A\}}\right| = \frac{1}{f_A I\!P\{A\}} \left| \sum_{j=1}^{f_A} [I\!P\{A\} - I\!P\{\overline{\tau_A} = j\}] \right|$$

$$
\begin{aligned}
&= \frac{1}{f_A I\!P\{A\}} \sum_{j=1}^{f_A} I\!P\left\{\overline{X}_0 = A \ , \ \overline{\tau_A} \leq j-1\right\} \\
&\leq \frac{1}{f_A I\!P\{A\}} \sum_{j=1}^{f_A} \sum_{i=1}^{j-1} I\!P\left\{\overline{X}_0 = A \ , \ \overline{X}_i = A\right\} \\
&\leq \frac{1}{f_A I\!P\{A\}} \sum_{j=1}^{f_A} \sum_{i=0}^{j-1} I\!P\{A\}^2 \left(1+\psi(in)\right) \\
&\leq C f_A I\!P\{A\} \ .
\end{aligned}
$$

The second equality follows by (11). The second inequality follows from the ψ-mixing property of the process. This concludes the proof of the lemma.

4 Proof of Theorem 4.

By the triangle inequality

$$
\begin{aligned}
& \left| I\!P\{\overline{X}_0 = A \ , \ \overline{\tau_A} > t\} - I\!P\{\overline{X}_0 = A\} e^{-I\!P\{\overline{X}_0 = A\}t} \right| \\
\leq \ & \left| I\!P\{\overline{X}_0 = A \ , \ \overline{\tau_A} > t\} - I\!P\{\overline{X}_0 = A\} I\!P\{\overline{\tau_A} > t\} \right| \\
+ \ & I\!P\{\overline{X}_0 = A\} \left| I\!P\{\overline{\tau_A} > t\} - e^{-I\!P\{\overline{X}_0 = A\}t} \right| \qquad (12)
\end{aligned}
$$

To obtain an upperbound for the first term of the right hand side of the inequality, we observe that for any $t \geq 2$ we have

$$
\begin{aligned}
& \left| I\!P\{\overline{X}_0 = A \ , \ \overline{\tau_A} > t\} - I\!P\{\overline{X}_0 = A\} I\!P\{\overline{\tau_A} > t\} \right| \\
= \ & \left| I\!P\{\overline{X}_0 = A, \overline{X}_k \neq A \ ; \ 1 \leq k \leq t\} - I\!P\{\overline{X}_0 = A\} I\!P\{\overline{X}_k \neq A \ ; \ 1 \leq k \leq t\} \right| \\
\leq \ & \left| I\!P\{\overline{X}_0 = A, \overline{X}_k \neq A \ ; \ 1 \leq k \leq t\} - I\!P\{\overline{X}_0 = A, \overline{X}_k \neq A \ ; \ 2 \leq k \leq t\} \right| \qquad (13) \\
+ \ & \left| I\!P\{\overline{X}_0 = A, \overline{X}_k \neq A \ ; \ 2 \leq k \leq t\} - I\!P\{\overline{X}_0 = A\} I\!P\{\overline{X}_k \neq A \ ; \ 2 \leq k \leq t\} \right| \\
+ \ & I\!P\{\overline{X}_0 = A\} \left| I\!P\{\overline{X}_k \neq A \ ; \ 2 \leq k \leq t\} - I\!P\{\overline{X}_k \neq A \ ; \ 1 \leq k \leq t\} \right| \ .
\end{aligned}
$$

The first term in the right hand side of inequality (13) can be bounded above as follows

$$\begin{aligned} & I\!P\{\overline{X}_0 = \overline{X}_1 = A \ , \ \overline{X}_k \neq A \ ; \ 2 \leq k \leq t\} \\ \leq \ & I\!P\{\overline{X}_0 = A\} I\!P\{\overline{X}_1 = A\} \left(1 + \psi(0)\right)^2 I\!P\{\overline{X}_k \neq A \ ; \ 2 \leq k \leq t\} \\ = \ & C \ I\!P\{A\}^2 \ I\!P\{\overline{\tau_A} > t - 1\} \ . \end{aligned} \tag{14}$$

up1

Using the mixing property, the second term in the right hand side of (13) is bounded above by

$$I\!P\{\overline{X}_0 = A\} \ \psi(n) \ I\!P\{\overline{\tau_A} > t - 1\} \ .$$

As we did for the first term, the third term in the right hand side of (13) can be bounded by

$$\begin{aligned} & I\!P\{\overline{X}_0 = A\} I\!P\{\overline{X}_1 = A \ , \overline{X}_k \neq A; 2 \leq k \leq t\} \\ \leq \ & I\!P\{\overline{X}_0 = A\} I\!P\{\overline{X}_1 = A\} \left(1 + \psi(0)\right) I\!P\{\overline{X}_k \neq A; 2 \leq k \leq t\} \\ = \ & C \ I\!P\{A\}^2 \ I\!P\{\overline{\tau_A} > t - 1\} \ . \end{aligned} \tag{15}$$

up2

A similar upper bound holds for $t = 1$. To prove it we first observe that

$$\left| I\!P\{\overline{X}_0 = A \ , \ \overline{\tau_A} > 1\} - I\!P\{\overline{X}_0 = A\} I\!P\{\overline{\tau_A} > 1\} \right| = \left| I\!P\{\overline{X}_0 = A \ , \ \overline{\tau_A} = 1\} - I\!P\{\overline{X}_0 = A\} I\!P\{\overline{\tau_A} = 1\}. \right| \ .$$

Then we use the the mixing property to conclude that

$$\left| I\!P\{\overline{X}_0 = A \ , \ \overline{\tau_A} = 1\} - I\!P\{\overline{X}_0 = A\} I\!P\{\overline{\tau_A} = 1\} \right| \leq (2 + \psi(0)) I\!P\{A\}^2 \ . \tag{16}$$

up3

Since $I\!P\{\overline{\tau_A} > 0\} = 1$, this last inequality is the upper bound we wanted for $t = 1$.

Inequalities (14), (15)and (16) together imply that

$$\left| I\!P_A\{\overline{\tau_A} > t\} - e^{-I\!P\{A\}t} \right| \leq C(\psi(n) + I\!P\{A\}) I\!P\{\overline{\tau_A} > t - 1\} \ , \tag{17}$$

up4

where C is a postive constant. To conclude the proof of Theorem 4 it is enough to use Theorem 3 to the right hand side of inequality (17).

5 Proof of Theorem 1.

teo:ent

By definition

$$I\!E(\log \overline{R_n}) = \sum_{t\geq 1} \log t \; I\!P\left\{\overline{R_n} = t\right\} = \sum_{t\geq 1} (\log(t+1) - \log t)\, I\!P\left\{\overline{R_n} > t\right\} \; .$$

Also, by definition

$$I\!P\left\{\overline{R_n} > t\right\} = \sum_{A\in\mathcal{C}_n} I\!P\{A\} \; I\!P_A\left\{\overline{\tau_A} > t\right\} \; .$$

Thus

$$I\!E(\log \overline{R_n}) = \sum_{A} I\!P\{A\} \; \sum_{t\geq 1} (\log(t+1) - \log t)\, I\!P_A\left\{\overline{\tau_A} > t\right\} \; .$$

We define

$$nu(\epsilon) = \epsilon \sum_{k\geq 1} \log k \; e^{-\epsilon(k-1)} \; .$$

Clearly, we have that

$$\Xi = \lim_{\epsilon\to\infty} \left[\widetilde{\nu}(\epsilon) + \log\epsilon\right] \; .$$

Therefore

$$\sum_{A} I\!P\{A\} \sum_{t\geq 1} (\log(t+1) - \log t)\, e^{-I\!P\{A\}t} = \sum_{A} \left(1 - e^{-I\!P\{A\}}\right) \nu(I\!P\{A\}) \; . \qquad (18)$$

ident

Hence, by the triangle inequality

$$\left| I\!E(\log \overline{R_n}) - H_n - \Xi \right| \leq I + II + III \; ,$$

where the terms I, II and III are defined below.

We use Theorem 4 to get an upper bound for the first term

$$
\begin{aligned}
I &:= \left| \sum_A I\!P\{A\} \sum_{t\geq 1} (\log(t+1) - \log t) \left[I\!P_A\{\overline{\tau_A} > t\} - e^{-I\!P\{A\}t} \right] \right| \qquad (19) \quad \boxed{\texttt{I}} \\
&\leq \sum_A I\!P\{A\} \sum_{t\geq 1} \frac{1}{t} \left| I\!P_A\{\overline{\tau_A} > t\} - e^{-I\!P\{A\}t} \right| \\
&\leq \sum_A I\!P\{A\} \sum_{t\geq 1} \frac{1}{t} (\varepsilon_n + \psi(n)) \ f(A,t) \\
&\leq C \max_A (\varepsilon(A) + \psi(n)) \\
&\leq C \ (\varepsilon_n + \psi(n)) \ .
\end{aligned}
$$

To define II we use identity (18). The upper bounds for the second and third terms are obtained by straightforward computations

$$
\begin{aligned}
II &:= \left| \sum_A I\!P\{A\} \left[\frac{1 - e^{-I\!P\{A\}t}}{I\!P\{A\}} \ \nu(I\!P\{A\}) \ - \ \nu(I\!P\{A\}) \right] \right| \qquad (20) \quad \boxed{\texttt{II}} \\
&= \sum_A I\!P\{A\} \left| \frac{1 - e^{-I\!P\{A\}t} - I\!P\{A\}}{I\!P\{A\}} \ \nu(I\!P\{A\}) \right| \\
&\leq \sum_A I\!P\{A\} \ 2 I\!P\{A\} \nu(I\!P\{A\}) \\
&\leq \max_A 2 I\!P\{A\} \nu(I\!P\{A\}) \\
&\leq C \ e^{-cn} \ .
\end{aligned}
$$

Finally

$$
\begin{aligned}
III &:= \left| \sum_A I\!P\{A\} \ \nu(I\!P\{A\}) - H_n - \Xi \right| \\
&= \left| \sum_A I\!P\{A\} \ [\nu(I\!P\{A\}) + \log I\!P\{A\} - \Xi] \right| \\
&\leq \max_A |\nu(I\!P\{A\}) + \log I\!P\{A\} - \Xi| \\
&\leq C \ e^{-cn} \ .
\end{aligned}
$$

On the other hand

$$\begin{aligned}
& \left|\mathrm{Var}(\log \overline{R_n}) - \mathrm{Var}(\log I\!P\{X_1^n\}) - \Delta + \Xi^2\right| \\
\leq & \left| I\!E\left((\log \overline{R_n})^2\right) - \sum_A I\!P\{A\}\widetilde{\omega}(I\!P\{A\})\right| \\
+ & \left|\sum_A I\!P\{A\}\widetilde{\omega}(I\!P\{A\}) - \sum_A I\!P\{A\}(\log I\!P\{A\})^2 - 2\Xi H_n - \Delta\right| \\
+ & \left|H_n^2 + 2\Xi H_n + \Xi^2 - \left(I\!E(\log \overline{R_n})\right)^2\right| \ .
\end{aligned}$$

In the right hand side of the above inequality, the first term is bounded repeating the argument used in (19) and (20). The bound for the second is obtained using the definition of Δ. The upper bound of the third one is a consequence of inequality 4.

This ends the proof of the theorem.

Proof of Corollary 2. Let

$$A = (a_0, \ldots, a_{n-1}) \ ,$$

with $a_i \in \mathcal{E},\ i = 1, \ldots, n$. We first recall that for Markov chains one has $\psi(n) \leq Ce^{-cn}$. Second, we observe that

$$I\!P\{A\} = I\!P\{(a_0, \ldots, a_{n-1})\} = I\!P\{(a_0)\} \prod_{i=1}^{n-1} I\!P\{a_i | a_{i-1}, \ldots, a_0\} \ .$$

Define

$$F_n := H_n - H_{n-1} = \sum_{a_{n-1}, \ldots, a_0} I\!P\{(a_{n-1}, \ldots, a_0)\} \log I\!P\{a_{n-1} | a_{n-2}, \ldots, a_0\} \ ,$$

for all positive integers n, with $H_0 = 0$.

Then it si enough to observe that inequality (4) can be rewritten as

$$\left| \ I\!E\left(\log \overline{R_n}\right) - \sum_{i=1}^n F_i - \Xi \ \right| \leq \varepsilon_n \ ,$$

to conclude the proof.

6 Final remarks

Maurer's conjecture was originally formulated using logarihtms of base 2, which is a natural choice when we consider binary sources. We recall that using base 2 logarithms the values of the constants considered in this paper are $\Xi = -0.832746... = -\gamma/\ln 2$ where γ is Euler's constant and $\Delta = 4.117181...$

Remark 2. The return time $\overline{R_n}$ is defined considering the non-overlapping *adjacent* observations $(X_{kn}, \ldots, X_{kn+n-1})$;
$k \geq 1$. However we only need to consider that the n-blocks are non-overlapping, not necessarily adjancent. Let us fix a positive integer M and define the process

$$\left(X_k^M\right)_{k\in I\!N} \qquad \text{with} \qquad X_k^M = (X_{kMn}, \ldots, X_{(kMn)+n-1}) \quad .$$

with the corresponding retunr time

$$R_n^M = \inf\left\{k \geq 1 : (X_{kMn}, \ldots, X_{(kMn+n-1)} = (X_0, \ldots, X_{n-1})\right\}.$$

Similarly to $\overline{\tau_A}$ we can define τ_A^M. Clearly, Theorem 3 and Theorem 4 hold τ_A^M and so does Theorem 1, too. The larger is M, the more mixing is the process $\left(X_k^M\right)_{k\in I\!N}$) and the limits in Theorem 1 (and Theorems 3 and 4) take place faster. Statistically speaking, this means that considering R_n^M has the disadvantage that we need a larger sample to observe the repetition time but when this is not a problem, it has the advantage that the convergence given by Theorem 1 for R_n^M is faster than the convergence for $\overline{R_n}$.

An important result in ergodic theory is the famous Kac's Lemma (Kac, 1947) which states that for an ergodic system the expected *return* time to a measurable set with positive measure is the inverse of the measure of the event. We present in the next corollary an estimation for all the moments of the non-overlapping hitting and return times. This a consequence of the bounds given in Theorem 3 and Theorem 4. Roughly speaking, if the hitting (or return) time law is close to an exponential law with parameter $I\!P\{A\}$, we must have that $\beta!/I\!P\{A\}^\beta \approx I\!E({\tau_A}^\beta)$.

cor:mean **Corollary 6** *Let $\beta > 0$. Let $(X_m)_{m\in Z\!\!Z}$ be a ψ-mixing process. Then the following inequalities hold*

$$\left|I\!P\{A\}^\beta I\!E\left({\overline{\tau_A}}^\beta\right) - k!\right| \leq C_\beta\ \varepsilon(A)\ ,$$

and

$$\left|I\!P\{A\}^\beta I\!E_A\left({\overline{\tau_A}}^\beta\right) - k!\right| \leq C_\beta\ \varepsilon(A)\ ,$$

where $\varepsilon(A)$ is the same as in Theorem 3, and C_β is a positive constant that only depends on β.

Proof. Follows by direct integration using inequalities (6) and (8).

Acknowledgments

We are indebted to Pablo Ferrari and Davide Gabrielli for many interesting discussions on the subject. Special thanks to Davide Gabrielli for calling our attention to Wegenkittl's paper. We thank Pierre Collet ans Bernard Schmitt who pointed out a mistake in a previous version of Theorem 1. The authors thank FAPESP, FINEP, CNPq and an agreement USP-COFECUB for support. MA thanks the Centre de Physique Théorique, CNRS, Luminy and Université de Toulon et du Var for the hospitality during the preparation of this paper.

References

Ahit [1] M. Abadi (2001). Exponential approximation for hitting times in mixing processes. *Math. Phys. Elec. J.* **7** No. 2.

Asharp [2] M. Abadi (2003a). Sharp error terms and necessary conditions for exponential hitting times in mixing processes. *To appear in Ann. Probability*. Can be downloaded from *http://www.ime.usp.br/˜ abadi*.

Astatis [3] M. Abadi (2003b). Statistics and error terms of occurrence times in mixing processes. Submitted. Can be downloaded from *http://www.ime.usp.br/˜ abadi*.

AG [4] M. Abadi and A. Galves (2001). Inequalities for the occurrence times of rare events in mixing processes. The state of the art. *Markov Proc. Relat. Fields.* **7** . 1 97-112.

CN [5] G. Choe and D. Kim (2000). Average convergence rate of the first return time. *Coll. Math.* **84**, 159-171

CGS [6] P. Collet, A. Galves and B. Schmitt (1999). Repetition times for gibbsian sources. *Nonlinearity* **12**, 1225-1237.

Sinai [7] I. Cornfeld, S. Fomin and Y. Sinai (1982). Ergodic theory, Grundlähren der Mathematischen Wissenschaften, vol 245. *Springer-Verlag, New York.*

ChK [8] J. Coron and D. Naccache (1999). An accurate evaluation of Maurer's universal test. *Lec. Notes Compt Sci.* **1556**

GS [9] A. Galves and B. Schmitt (1997). Inequalities for hitting times in mixing dynamical systems. *Random Comput. Dyn.* **5**, 337-348.

Kac [10] M. Kac (1947). On the notion of recurrence in discrete stochastic processes. *Bull. Amer. Math. Soc.* **53**, 1002-1010.

Mau [11] U. Maurer (1992). A universal test for random bit generators. *J. Cryptology* **5**, 89-105.

Weg [12] S. Wegenkittl (2001). Entropy estimators and serial tests for ergodic chains. *IEEE Trans. Infor. Theory* **47**, 6 2480-2489.

Miguel Abadi and Antonio Galves
Instituto de Matemática e Estatística
Universidade de São Paulo
Rua do Matão 1010 São Paulo, SP, 05508-090 BRASIL
E-MAIL: abadi@ime.usp.br, galves@ime.usp.br