Inequalities for the occurrence times of rare events in mixing processes. The state of the art.*

Miguel Abadi† Antonio Galves‡

June 9, 2000

Dedicated to Joel Lebowitz
on his 70th anniversary

Abstract

The first occurrence time of a rare event in a mixing process typically has a distribution which can be well approximated by the exponential law. In this paper we review recent theorems giving upper bounds for the error term of this approximation. We shall focus on papers treating the problem in a general mixing framework.

Running title: Rare events in mixing processes.
Keywords: Mixing processes, occurrence time of a rare event, exponential approximation.
AMS subject classification: 60G10, 60F05, 37A50

1 Introduction.

The first occurrence time of a rare event in a mixing process typically has a distribution which can be well approximated by the exponential law. In this paper we review recent theorems giving upper bounds for the error term of this approximation. We shall focus on papers treating the problem in a general mixing framework.

*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 Pronex Project "Critical phenomena in probability and stochastic processes" (grant 41.96.0923.00)
†Supported by CAPES (Ph.D. grant)
‡Work partially supported by CNPq (research grant 301301/79)

The pioneer paper in this area is Döeblin (1940), who studied the Poisson approximation for the Gauss transformation. In the context of Markov chains, the convergence of the occurrence time of a rare event to the exponential law was first studied by Bellmann and Harris (1951) and Harris (1953). Then after a long period in which apparently nothing appeared in the area, several papers and books studying the problem for different types of Markov processes were published, starting with Keilson (1979), Aldous (1982), Korolyuk and Sil'vestrov (1984), Cassandro, Galves, Olivieri and Vares (1984), Kipnis and Newman (1985), Cogburn (1985), Lebowitz and Schonmann (1987), Olivieri and Scoppola (1995) and (1996) among others. In the context of dynamical systems the question was considered by Galves and Schmitt (1990), Pitskel (1991), Collet, Galves and Schmitt (1992), Collet and Galves (1993), Hirata (1993) and Collet and Galves (1995). Part of these papers were motivated by the modeling of physical phenomena like metastability or intermittency. With the exception of Aldous (1982), all the others are only interested in a qualitative result, and don't present upper bounds for the exponential approximation.

In the nineties, several papers presented bounds for the error of the exponential approximation. Aldous and Brown (1992, 1993) studied the case of reversible and finite Markov chains. Results for different types of interacting Markovian systems appeared in Ferrari, Galves and Landim (1994), Ferrari, Galves and Liggett (1995), Asselah and Dai Pra (1997), Simonis (1998).

The subject is considered in the context of different types of mixing processes in recent papers by different authors. Galves and Schmitt (1997) consider ϕ-mixing dynamical system with respect to a summable function ϕ. Collet, Galves and Schmitt (1999) consider one-dimensional Gibbs states, with Hölder continuous potentials. Haydn (1999) worked with rational maps on the Riemann sphere. Paccaut considers the case of weighted maps of the interval. Hirata, Saussol and Vaienti (1999) prove the Poisson approximation for the successive entrance times of a rare event. Their proof holds in a very general framework and provides sharp bounds for sets which does not recur very fast. Finally the recent paper by Abadi (2000) proves the exponential approximation for any ϕ-mixing processes and also for α-mixing processes with summable α. As far as we know this is the best result available in the field.

This paper is organized as follows. The basic definitions and notation is presented in section 2. Then in the next sections we review Aldous (1982), Aldous and Brown (1992), Galves and Schmitt (1997), Collet, Galves and

Schmitt (1999), Hirata, Saussol and Vaienti (1999) and Abadi (2000).

2 Basic definitions and notation.

Let $\mathcal{E}$ be a finite set. Put $\Omega = \mathcal{E}^{\mathbb{Z}}$. For each $n \in \mathbb{Z}$, define $X_n : \Omega \to \Re$ the n-th coordinate projection. We denote by $T : \Omega \to \Omega$ the one-step left shift operator.

We denote by $\mathcal{F}$ the σ-algebra over Ω generated by cylinders. Moreover we denote with $\mathcal{F}_I$ the σ-algebra generated by cylinders with coordinates in I, $I \subseteq \mathbb{Z}$.

Let $\mathcal{C}_n$ be the set of all cylinders A of the form N

$$A = \{X_0 = a_0, \ldots, X_{n-1} = a_{n-1}\} \ ,$$

with $a_i \in \mathcal{E},\ i = 1, \ldots, n$.

Let $\mathbb{P}$ be a probability measure over $\mathcal{F}$, such that

$$\mathbb{P}(A) = \mathbb{P}(T^{-1}(A)) \ ,$$

for all $A \in \mathcal{F}$. This is called a *stationary* measure in the literature. To avoid uninteresting pathologies we shall also assume that

$$0 < \mathbb{P}\{X_n = a\} < 1 \ ,$$

for all $a \in \mathcal{E}$.

We shall say that $(\Omega, \mathcal{F}, \mathbb{P}, T)$ is *γ-mixing* or uniform mixing if, for all integers $n \geq 1$ and $l \geq 0$, the following inequality holds

$$\sup_{B \in \mathcal{F}_{\{0,.,n\}}, C \in \mathcal{F}_{\{n \geq 0\}}} \left| \mathbb{P}\left(B \cap T^{-(n+l+1)}C\right) - \mathbb{P}(B)\mathbb{P}(C) \right| = \gamma(l) \ ,$$

α-mixing if

$$\sup_{B \in \mathcal{F}_{\{0,.,n\}}, C \in \mathcal{F}_{\{n \geq 0\}}} \frac{\left| \mathbb{P}\left(B \cap T^{-(n+l+1)}C\right) - \mathbb{P}(B)\mathbb{P}(C) \right|}{\mathbb{P}(B)} = \alpha(l) \ ,$$

and *ϕ-mixing* if

$$\sup_{B \in \mathcal{F}_{\{0,.,n\}}, C \in \mathcal{F}_{\{n \geq 0\}}} \frac{\left| \mathbb{P}\left(B \cap T^{-(n+l+1)}C\right) - \mathbb{P}(B)\mathbb{P}(C) \right|}{\mathbb{P}(B)\mathbb{P}(C)} = \phi(l) \ ,$$

where in the above expressions the supremum is taken over the sets B and C, such that $\mathbb{P}(B) > 0$ in the second case and such that $\mathbb{P}(B)\mathbb{P}(C) > 0$ in the third one, and $\gamma = (\gamma(l))_{l\geq 0}$, $\alpha = (\alpha(l))_{l\geq 0}$ and $\phi = (\phi(l))_{l\geq 0}$ are decreasing sequences of positive real numbers converging to zero.

Clearly, ϕ-mixing implies α-mixing and this last one implies uniform mixing. We observe that an irreducible and aperiodic finite state Markov chain is ϕ-mixing with exponential decay. Moreover, Gibbs states on $\Omega = \mathcal{E}^{\mathbb{Z}}$, with a potential with summable variations are exponentially ϕ-mixing.

Remark. In the literature the mixing conditions have received different names. For example, Doukhan (1995) calls α-mixing (respectively ϕ, ψ), what we call γ-mixing (respectively α, ϕ). Also, Shields(1996) use ψ-mixing to denote what we call ϕ-mixing. Our terminology is the same adopted by Galves and Schmitt (1997) Collet, Galves and Schmitt (1999) and Hirata, Saussol and Vaienti (1999).

Given $A \in \mathcal{C}_n$, we define the *entrance time* $\tau_A : \Omega \to \mathbb{N}$ as follows

$$\tau_A(\omega) = \inf\{k \geq 0 : T^k(\omega) \in A\} \ .$$

We shall use the classic probabilistic shorthand notation for events defined through random variables. We shall write $\{\tau_A = m\}$ instead of $\{\omega \in \Omega : \tau_A(\omega) = m\}$, and $T^{-k}(A)$ instead of $\{\omega \in \Omega : T^k(\omega) \in A\}$. As usual, the mean of a random variable X will be called its *expectation* and denoted by $\mathbb{E}(X)$. We also denote the conditional measure with respect to the event A

$$\mathbb{P}_A\{W\} = \mathbb{P}\{A \cap W\} / \mathbb{P}\{A\} \ ,$$

for any $W \subseteq \mathcal{F}$.

We conclude this section with some references. The nice book by Shields (1996) is a wonderful introduction to mixing processes and to Ergodic Theory. We refer to Doukhan (1995) for examples and references of α-mixing processes that are not ϕ-mixing and for ϕ-mixing processes decaying at any rate. For definitions and basic properties of Gibbsian models we refer the reader to Bowen (1975).

3 Aldous 1982.

As far as we know, Aldous (1982) is the first paper which consider the problem of estimating the error term in the exponential approximation. The result holds when (X_n) is an irreducible Markov chain taking values in the finite set $\mathcal{E}$. Let us define η as the relaxation rate of the chain, namely,

$$\eta := \min_{n \in \mathbb{N}} \left\{ \| \mathbb{P}\left(X_n = \cdot | X_0 = a\right) - \pi\left(\cdot\right) \| \leq \frac{1}{e^2} \text{ for all } a \in \mathcal{E} \right\} , \qquad (1)$$

where π is the probability measure on $\mathcal{E}$ invariant with respect to the chain and

$$\|\mu - \nu\| = \frac{1}{2} \sum_{a \in \mathcal{E}} |\mu(a) - \nu(a)|$$

is the total variation norm.

The following theorem holds

Theorem. *For any subset* $A \subset \mathcal{E}$

$$\sup_{t>0} \left| \mathbb{P}\left\{\tau_A > t\right\} - e^{-t/\mathbb{E}(\tau_A)} \right| \leq \frac{C\eta}{\mathbb{E}(\tau_A)} \left(1 + \log^+ \left(\frac{\mathbb{E}(\tau_A)}{\eta}\right)\right) .$$

A simple example will help understanding the behavior of this upper bound. Consider a Markov chain $(X_n)_{n \in \mathbb{N}}$ with state space $\mathcal{E} = \{a, b\}$ and with transition probability matrix

$$Q = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

with $0 < p + q < 1$. Let $A = \{a\}$. We have that $\pi(a) = q/(p+q)$. So, the interesting case is when $p/q \to \infty$. In this case the expectation of the hitting time is explicitly computable and has the value

$$\mathbb{E}\left(\tau_A\right) = \frac{p}{(p+q)q}.$$

Moreover, an elementary computation shows that

$$\| \mathbb{P}\left(X_n = \cdot | X_0 = a\right) - \pi\left(\cdot\right) \| = (1 - (p+q))^n .$$

It follows that

$$\eta \approx \frac{-2}{\log\left(1 - (p+q)\right)} .$$

Remembering that $p << q$, we obtain the estimate

$$\frac{\mathbb{E}(\tau_A)}{\eta} \approx \frac{p}{q} ,$$

and therefore the upper bound in (1) is estimated by

$$\frac{C\,p}{q}\left(1+\log^{+}\left(\frac{q}{2p}\right)\right) \approx -\pi(a)\log\pi(a)\ .$$

Just for comparison a direct estimation of error term can be done, since the hitting time, conditioned on $\{\tau_A > 0\}$ has a geometric distribution with parameter $1-q$. The mean value theorem gives the following bound for the error of the approximation.

$$\left|-\log(1-q) - \frac{1}{\mathbb{E}(\tau_A)}\right| \max\left\{-\log(1-q)\ ;\ \frac{1}{\mathbb{E}(\tau_A)}\right\}^{-1} \approx \pi(a)\ .$$

4 Aldous and Brown 1992

In a further paper Aldous and Brown (1992) study the hitting time of a rare event $A \subset \mathcal{E}$ in the following setting. $\{X_t : t \geq 0\}$ is a continuous-time finite-state Markov chain with transition matrix $Q = (q(i,j); i,j \in \mathcal{E})$ and where $q(i,i) = -\sum_{i\neq j} q(i,j)$. Here $\eta = 1/\lambda_2$, where λ_2 is the second eigenvalue of $-Q$.

Theorem. *For any subset* $A \subset \mathcal{E}$

$$\sup_{t>0}\left|\mathbb{P}\left\{\tau_A > t\right\} - e^{-t/\mathbb{E}(\tau_A)}\right| \leq \frac{\eta/\mathbb{E}(\tau_A)}{1+\eta/\mathbb{E}(\tau_A)} \leq \eta/\mathbb{E}(\tau_A)\ . \qquad (2)$$

This upper bound is useful only if the relaxation time of the chain is much smaller than the expectation of the occurrence time of the rare event.

An example can help understanding the type of difficulties one can find to apply this upper bound. In Ferrari, Galves and Ligget (1995) there is an upper bound for the symmetric simple exclusion process. The process can be informally defined as follows. Infinitely many particles perform interacting random walks on $\mathbb{Z}$. At most one particle is allowed in each site $x \in \mathbb{Z}$ and, at rate one, the contents of sites x and $x+1$ are interchanged. Hence if both sites are occupied or both sites are empty, nothing happens but if one of the sites is occupied and the other is empty, the interchange is seen as a jump of the particle to the empty site.

The only extremal invariant measures for this system are the product measures μ_ρ, $\rho \in [0,1]$. Under the measure μ_ρ, the probability that a site is occupied by a particle is ρ and the occupation variables of different sites are independent random variables.

Consider the symmetric simple exclusion process over $\mathbb{Z}$ starting with the equilibrium product measure μ_ρ with density $\rho \in (0,1)$. Let τ_N the first time for which the interval $\{1,\dots,N\}$ is totally occupied. Since the symmetric simple exclusion process is reversible with respect to μ_ρ, it is tempting to apply Aldous and Brown approach to obtain the error for the law of τ_N. Unfortunately the process takes values on $\{0,1\}^{\mathbb{Z}}$ which is not finite, and therefore a direct application of the result of Aldous and Brown is excluded.

A possible way to overcome this difficulty is to approximate the infinite particle system by a sequence of symmetric simple exclusion processes taking values on the finite sets $\{0,1\}^A_N$, where $A_N = [-\ell_N, N+\ell_N]$ with any boundary condition that leaves invariant the measure μ_ρ restricted to the box. We need to estimate $\eta_N/\mathbb{E}(\tau_N)$ where $1/\eta_N$ is the second eigenvalue of the continuous time Markov chain in the finite box.

To apply this to obtain the result for the infinite system we need to take a box of length $\ell_N = (\mathbb{E}\tau_N)^{(1+\varepsilon)/2}$ for some $\varepsilon > 0$. This is necessary to guarantee that the time of occurrence of the rare event for the finite and the infinite systems behave in the same way. Indeed the influence of the boundary behaves *at least* as a simple symmetric random walk, and it takes a time of the order of the square of the length of the box to arrive to the center of the box. On the other hand, Lu and Yau (1993) proved that η_N is of the order at least of the square of the length of the box. Hence the upper bound in (2) diverges and the finite approximation cannot be used naively.

Actually, Ferrari, Galves and Ligget (1995) presents the following upper bound

Theorem. *There exists positive constants α', α'', A and A' independent of N and a sequence $\alpha_N \in [\alpha', \alpha'']$ such that*

$$\sup_{t\geq 0}\left|P\left\{\alpha_N\rho^N\tau_N > t\right\} - e^{-t}\right| \leq A\rho^{A'N}.$$

We remark that the scaling factor ρ^N is just the probability of the rare event by μ_ρ. This is suggested by Kac's Lemma(cf. [23]) which says that the expectation of the first return time to a given event is just the inverse of the probability of the event.

The approach followed by Ferrari, Galves and Ligget (1995) was successively used by Asselah and Dai Pra (1997) to obtain the error term for more general types of rare events for the symmetric simple exclusion process and by Simonis to study the same type of rare event for the contact process. This approach was also the starting point of the result by Galves and Schmitt (1997).

5 Galves and Schmitt 1997

This paper considers the general setting of ϕ-mixing processes, with ϕ summable. The result holds for any type of cylinder set A. The scaling factor is $\lambda_A I\!P(A)$ with λ_A belonging to an interval bounded away from zero (an containing the number one). The result is the following.

Theorem. *Let $(\Omega, \mathcal{F}, I\!P, T)$ be ϕ-mixing and let us suppose that the function $\phi : I\!N \to I\!R_+$ is summable. Then, there exist four strictly positive constants $\Lambda_1, \Lambda_2, \beta, C$, such that for any n and any $A \in \mathcal{C}_n$, there exists $\lambda(A) \in [\Lambda_1, \Lambda_2]$, for which the following inequality holds*

$$\sup_{t>0} \left| I\!P\left\{ \tau_A > \frac{t}{\lambda(A) I\!P(A)} \right\} - e^{-t} \right| \leq C I\!P(A)^{\beta} \ . \tag{3}$$

The proof of the theorem has two parts. First it is shown a factorization property holds for $I\!P\{\tau_A > f_A t\}$, namely

$$I\!P\{\tau_A > f_A(t+s)\} = I\!P\{\tau_A > f_A t\}\, I\!P\{\tau_A > f_A s\}$$

for all t and s in $\Re_+$, where f_A is a suitable scaling factor. This will follows as a consequence of the mixing property, if the event $\{\tau_A > f_A(t+s)\}$ is replaced by the event which says that the cylinder A is not visited between times 0 and $f_A t$ and between $f_A t + \Delta$ and $f_A(t+s)$, where Δ is a suitable gap. The game here is to compensate the length of the gap with the loss of memory of the process.

The second part of proof is to choose the right scaling factor f_A, in such a way that the limit is a non-degenerated distribution.

A natural question suggested by the theorem is the precise value of the parameter λ_A. Kac's lemma suggests that $\lambda_A = 1$. Collet, Galves and Schmitt (1999) shows that this is typically the case.

6 Collet, Galves and Schmitt 1999

Collet, Galves and Schmitt (1999) shows that the correction factor depends on the short time returns of the rare event to itself in the firsts steps. As a consequence, for a overwhelming set of cylinders the scaling factor λ_A in (3) is asymptotically one, as suggested by Kac's lemma.

Definition. Let $s \in I\!N$. Define $\mathcal{B}_n = \mathcal{B}_n(s)$ as the set of sequences of length n of elements in $\mathcal{E}$ which recur before time n/s, namely, $A \in \mathcal{B}_n$ if and only if there is an integer j, $1 \leq j \leq n/s$ such that $A \cap T^{-j}A \neq \emptyset$.

Theorem. *Let $(\Omega, \mathcal{F}, I\!P, T)$ be exponentially ϕ-mixing and let us suppose that the function $\phi : I\!N \to I\!R_+$ is summable. Then, there exist two strictly positive constants c, C, such that for any n and any $A \in \mathcal{C}_n/\mathcal{B}_n$, there exists $\lambda(A) \in [\Lambda_1, \Lambda_2]$, for which the following inequality holds*

$$\sup_{t>0} \left| I\!P\left\{\tau_A > \frac{t}{I\!P(A)}\right\} - e^{-t} \right| \leq C e^{-cn} \ .$$

Moreover,

$$I\!P\{\mathcal{B}_n(3)\} \leq C e^{-cn} \ .$$

7 Hirata, Saussol and Vaienti 1999

This paper shows that the distribution of the first hitting time into A and the first return time to A, suitably rescaled, are close to an exponential one law if and only if both distributions are mutually close in the following sense,

Theorem. *Let us define $c(k, A) = |I\!P_A(\tau_A > k) - I\!P(\tau_A > k)|$ and set $c(A) = \sup_k c(k, A)$. The distribution of the rescaled first hitting time into the set A differs from the exponential-one law up to $d(A) := 4I\!P(A) + c(A)(1 + \log c(A)^{-1})$, namely :*

$$\sup_{t\geq 0} \left| I\!P\left(\tau_A > \frac{t}{I\!P(A)}\right) - e^{-t} \right| \leq d(U)$$

and the same holds starting from A :

$$\sup_{t\geq 0} \left| I\!P_A\left(\tau_A > \frac{t}{I\!P(A)}\right) - e^{-t} \right| \leq d(U).$$

Conversely, the difference between the rescaled hitting time and the rescaled return time can be bounded in terms of the distance $\tilde{c}(A) := \sup_{t\geq 0} |I\!P_A(\tau_A > t/I\!P(A)) - e^{-t}|$, namely,

$$c(A) \leq 2I\!P(A) + \tilde{c}(A)(2 + \log \tilde{c}(A)^{-1}) \ .$$

This theorem gives a *necessary and sufficient condition* to obtain the exponential law for the first hitting time and the first return time. That is $c(A) \to 0$, if and only if $\tilde{c}(A) \to 0$. The proof holds for any stationary measure $I\!P$, without any assumption with respect to its ergodicity. It is clear that with this is generality we can not expect to assure the convergence of $c(A)$. The next lemma gives an upper bound for $c(A)$, which indicates when we can expect convergence.

Lemma. *Let $A \subset \Omega$ a measurable set. The following estimate holds*

$$c(A) \leq \inf_{k \in I\!N} \{a_k(A) + b_k(A) + kI\!P(A)\}$$

where the quantities are defined by

$$\begin{aligned} a_k(A) &= I\!P_A(\bigcup_{j=1}^{k} T^{-j}A) = I\!P_A(\tau_A \leq k), \\ b_k(A) &= \sup_{W \in \mathcal{F}_{\{0,.,n\}}} |I\!P_A(T^{-k}W) - I\!P(W)| \end{aligned}$$

Note that $b_k(A) = \alpha(k-n)$, if the process is α-mixing, and $b_k(A) = \gamma(k-n)/I\!P(A)$, if the process is γ-mixing.

The key element in this upper bound is the quantity $a_k(A) = I\!P_A(\tau_A \leq k)$. For instance, if $A \in \mathcal{C}_n/\mathcal{B}_n(s)$, then $a_k(A) = 0$ for $k = 1, ..., s$. If the process looses memory fast enough, then $b_k(A)$ decreases to 0 and the theorem assures the asymptotic convergence to the exponential law.

It is important to remark that in many situations $c(A)$ does not converge to zero. The classical Ehrenfest model provides an example of this fact. The Ehrenfest model is a Markov chain, taking values on the set $\mathcal{E} = \{0,1\}^d$, where N is a positive integer. For any $x \in \mathcal{E}$ and any index $i \in \{1, \ldots, N\}$, the state x^i is defined as the string which coincides with x in all the coordinates with the exception of coordinate i. We define the transition probabilities of the chain as follows

$$Q(x,y) = \begin{cases} \frac{1}{2N} & \text{if } x^i = y \ , \quad \text{for some } \ i = 1, \ldots, N \ , \\ \frac{1}{2} & \text{if } \ x = y \ , \\ 0 & \text{otherwise} \ . \end{cases}$$

The invariant measure for the chain is the uniform distribution on $\mathcal{E}$. Now consider the event $A = \{(1, \ldots, 1)\}$. If N is large this is a rare event. Let us suppose that the chain starts with its invariant measure and define τ_A as the first time the chain visits A.

Then a straightforward computation gives that,

$$\mathbb{P}_A(\tau_A > 1) \quad = \quad \frac{1}{2}$$

and

$$\mathbb{P}(\tau_A > 1) \quad \geq \quad 1 - \frac{N+1}{2N} \, .$$

Therefore, $c(1, A)$ converges to $1/2$, when N diverges.

Hirata, Saussol and Vaienti(1999) and Collet, Galves and Schmitt(1999) prove that $\mathbb{P}(A)$ is the right scaling factor for cylinders which do not recur too fast. Galves and Schmitt(1998) prove that the exponential approximation is true *for all* cylinders using the scaling factor $\lambda_A \mathbb{P}(A)$, where the possibility that $\lambda_A = 1$ is not excluded. Therefore the question remains. Is it true that the parameter λ_A can be different from 1? In Abadi (2000) a computable estimation of this value is presented and it is shown that λ_A can be really different from 1.

8 Abadi 2000

Abadi (2000) presents a modification of Galves and Schmitt's approach and obtains a sharper upper bound for the exponential approximation. The improvement is obtained in part by changing the scaling factor. Moreover, it is shown that the scaling factor can be written as $\xi_{A,\mu(t)} \mathbb{P}(A)$, where $\xi_{A,\mu(t)}$ is bounded below and above by two strictly positive constants Ξ_1 and Ξ_2, respectively, independent of A, n and t. This technique works for all type of cylinders, for any ϕ-mixing processes and also for α-mixing processes with summable α.

Kac's lemma says that the right scaling factor for return times is $\mathbb{P}(A)$. This suggests to use $\mathbb{P}(A)$ as the scaling factor for hitting times. We show that this is not always a good choice for first occurrence times.

In the main theorem of this paper it is proven that a better rate of convergence can be obtained by rescaling the first occurrence time by a certain correction factor and letting this factor have some (bounded) fluctuations.

Theorem 1. *Let $(\Omega, \mathcal{F}, I\!P, T)$ be ϕ-mixing or α-mixing with α summable. Then, there exists strictly positive constants Ξ_1, Ξ_2, C_1 and C_2 such that for any $n, A \in \mathcal{C}_n$ and $t > 0$ there exists $\xi_{A,\mu(t)} \in [\Xi_1, \Xi_2]$, for which the following inequality holds*

$$\sup_{t>0} \left| I\!P \left\{ \tau_A > \frac{t}{\xi_{A,\mu(t)} I\!P(A)} \right\} - e^{-t} \right| \leq C_2 \inf_{\Delta \geq n} \left[2\Delta I\!P(A) + *(\Delta - n) \right] \ .$$

*where * stands for α or ϕ. Moreover $e^{-\xi_{A,\mu(t)} I\!P(A) t}$ is left-continuous non-increasing function of t over each $[k/I\!P(A), (k+1)/I\!P(A)]$ and for all $k \in I\!N$*

$$\left| \lim_{t \to \left(\frac{k+1}{I\!P(A)}\right)^-} e^{-\xi_{A,\mu(t)}} - e^{-\xi_{A,\mu((k+1)/I\!P(A))}} \right| \leq 2\Delta I\!P(A) + *(\Delta - n) \ .$$

The use of $\xi_{A,\mu(t)} I\!P(A)$ as scaling factor provides a good approximation to the exponential law. However an explicit computation of this scaling factor is usually a most difficult task. To overcome this real difficulty it is presented a computable approximation $\zeta_{A,s}$ for this parameter in terms of the short recurrence times of each cylinder. A corollary presents conditions on the process and on the cylinders, assuring that $\zeta_{A,s} = 1$. In particular, it is proven that $\zeta_{A,s} = 1$ for an overwhelming set of cylinders. Finally the paper presents a class of examples in which the parameter is arbitrarily small.

The proof uses the following extensions of the result in Galves and Schmitt(1997).

Theorem 2. *Let $(\Omega, \mathcal{F}, I\!P, T)$ be ϕ-mixing. Then, there exists strictly positive constants $\Lambda_1, \Lambda_2, C_1$ and C_2 such that for any n and any $A \in \mathcal{C}_n$, there exists $\lambda_A \in [\Lambda_1, \Lambda_2]$, for which the following inequality holds*

$$C_1 I\!P(A) \leq \sup_{t>0} \left| I\!P \left\{ \tau_A > \frac{t}{\lambda_A I\!P(A)} \right\} - e^{-t} \right| \leq C_2 \, b(A)$$

where $b(A) = I\!P(A)^\eta$, $0 < \eta < 1$ if $\phi(n) \leq C/n^\epsilon$ for all $n > n_o$ with $\epsilon > 0$ and $b(A) = \phi^{1/2}(n)$ otherwise.

Then upper bound in Galves and Schmitt(1997) is $I\!P(A)^\beta$ with $0 < \beta < 1$ for ϕ summable systems. Theorem 2 extends that result for *any* type of decreasing function ϕ, and improves the upper bound.

Theorem 3. *Let $(\Omega, \mathcal{F}, I\!P, T)$ be α-mixing and let us suppose that the function $\alpha : I\!N \to I\!R_+$ is summable. Then, there exists strictly positive constants $\Lambda_1, \Lambda_2, C_1$, C_2 and κ such that for any n and any $A \in \mathcal{C}_n$, there exists $\lambda_A \in [\Lambda_1, \Lambda_2]$, for which the following inequality holds*

$$C_1 I\!P(A) \leq \sup_{t>0} \left| I\!P\left\{ \tau_A > \frac{t}{\lambda_A I\!P(A)} \right\} - e^{-t} \right| \leq C_2 I\!P(A)^{\kappa}$$

We emphasize that in both theorems the constants $\Lambda_1, \Lambda_2, C_1, C_2, \eta$ and κ are independent of n and A. As usual, $\Lambda_1, \Lambda_2, C_1$, and C_2 stand for (possibly) different constants in each theorem.

The difference between theorem 1 and theorems 2 and 3 is that the first one provides a more accurate upper bound, given that we use $\xi_{A,\mu(t)}$ instead of λ_A as a scaling factor.

The next theorem gives an estimation of $\xi_{A,\mu(t)}$. Let s be a positive integer. We define

$$\zeta_{A,s} = I\!P_A \left\{ \tau_A > \frac{n}{s} \right\} = \frac{I\!P\left\{\tau_A > \frac{n}{s} \cap A\right\}}{I\!P(A)} .$$

Theorem 4. *Let $(\Omega, \mathcal{F}, I\!P, T)$ be exponentially α-mixing. Let s be a positive integer. Then, there exist strictly positive constants Ψ_1, Ψ_2, C_1, C_2 and c such that for any n and any $A \in \mathcal{C}_n$, there exists $\zeta_{A,s} \in [\Psi_1, \Psi_2]$, for which the following inequality holds*

$$C_1 I\!P(A) \leq \sup_{t>0} \left| I\!P\left\{\tau_A > t\right\} - e^{-t\zeta_{A,s} I\!P(A)} \right| \leq C_2 e^{-cn} .$$

Now we prove that the cylinders for which Theorem 4 holds, are typical in the following sense

Lemma. *Let the system be α-mixing (at any rate). There exist $s \in I\!N$ and two positive constants C and c such that*

$$I\!P\{\mathcal{B}_n(s)\} \leq C e^{-cn} .$$

We now provide an example of a process and a sequence of cylinders A_n such that $\lambda_{A_n} < 1$.

Let $\mathcal{E} = \{0,1\}$ and $A = A_n = \{X_0 = 1, \ldots, X_{n-1} = 1\}$ $\forall n \in I\!N$. Suppose that X_n are i.i.d. with $I\!P(X_n = 1) = p$. Then $\zeta_{A,s} = 1 - p$. Note that in this case, the result by Hirata, Saussol and Vaienti gives $a_k(A_n) = 1 - p$ for all $k = 1, \ldots, n-1$. Then, the hitting rescaled by $I\!P(A)$ does *not* converge in law to the exponential distribution. To obtain we need to use $(1-p)I\!P(A)$ as scaling factor.

Acknowledgments. It is a pleasure to thank Amine Asselah, Xavier Bressaud, Pierre Collet, Pablo Ferrari, Nicolai Haydn, Frédéric Paccaut, Benoit Saussol, Bernard Schmitt, Elisabetta Scoppola and Sandro Vaienti for many illuminating discussions about the subject.

The authors thanks the hospitality of the Centre de Physique Théorique of the CNRS, at Marseille-Luminy, and the organizers of the meeting *Temps de retour, entropie et compléxité* where parts of this paper have been presented.

References

[1] M. Abadi (2000). Exponential approximation for hitting times in mixing processes. *Preprint IME-USP*. Can be retrieved from *http://rene.ma.utexas.edu/mp_arc/c /00/00-211.ps.gz*

[2] D. J. Aldous (1982). Markov chains with almost exponential hitting times. *Stochastic Process. Appl.* **13** 305-310.

[3] A. Asselah and P. Dai Pra (1997). Sharp estimates for the occurrence time of rare events for simple symmetric exclusion. *Stochastic Processes Appl.* **71** 259-273.

[4] D. J. Aldous and M. Brown (1992). Inequalities for rare events in time-reversible Markov chains I. *Stochastic Inequalities* IMS Lecture Notes **22** 1-16.

[5] D. J. Aldous and M. Brown (1993). Inequalities for rare events in time-reversible Markov chains II. *Stochastic Processes Appl.* **44** 15-25.

[6] Bellman and T. E. Harris (1951). Recurrence times for the Ehrenfest model. *Pacific J. Math.* **1** 179–193.

[7] R. Bowen (1975). Equilibrium states and the ergodic theory of Anosov diffeomorphisms. *Lecture Notes in Math.* **470**. Springer-Verlag, New York.

[8] M. Cassandro, A. Galves, E. Olivieri and M. E. Vares (1984). Metastable behavior of stochastic dynamics: a pathwise approach. *J. Stat. Phys.* **35**:603-633.

[9] R. Cogburn (1985). On the distribution of first passage and return times for small sets. *Ann. Probab.* **13** 1219-1223.

[10] P. Collet and A. Galves (1993). Statistics of close visits to the indifferent fixed point of an interval map. *J. Stat.Phys.* **72**, 459-478.

[11] P. Collet and A. Galves (1995) *Asymptotic distribution of entrance times for expanding maps of the interval. Dynamical Systems and Applications*, WSSIAA **4**, 139-152.

[12] P. Collet, A. Galves and B. Schmitt (1992). Unpredictability of the occurrence time of a long laminar period in a model of temporal intermittency. *Ann. Inst. H. Poincaré, Section A*, **57** 3:319-331.

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

[14] W. Döeblin (1940). Remarques sur la théorie métrique des fractions continues. *Compositio Math.* **7** 353-371.

[15] P. Doukhan (1995) Mixing. Properties and examples. *Lecture Notes in Statistics* **85**. Springer-Verlag.

[16] P. A. Ferrari, A. Galves and C. Landim (1994). Exponential waiting time for a big gap in a one dimensional zero range process. *Ann. Probab.* **22** 1:284-288.

[17] P. A. Ferrari, A. Galves and T. Liggett (1995). Exponential waiting time for filling a large interval in the symmetric simple exclusion process. *Ann. Inst. H. Poincaré* **31** 1:155-175.

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

[19] N. Haydn (1999). The distribution of the first return time for rational maps. *J. Stat. Phys.* **94** 5-6:1027-1036.

[20] M. Hirata (1993). Poisson law for Axiom A diffeomorphism. *Ergod. Th. Dyn. Sys.* **13**. 533-556.

[21] M. Hirata, B. Saussol and S. Vaienti (1999). Statistics of return times: a general framework and new applications. *Comm. Math. Phys.* **206** 33-55.

[22] T. E. Harris (1953). First passage and recurrence distributions. *Trans. Amer. Math. Soc.* **73** 471-486.

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

[24] J. Keilson (1979) *Markov Chain Models. Rarity and Exponentiality.* Springer-Verlag.

[25] C. Kipnis and C. Newman (1985). The metastable behavior of infrequently observed, weakly random, one-dimensional diffusion processes. *SIAM J. Appl. Math.* **45**:972.

[26] D. V. Korolyuk and D. S. Sil'vestrov (1984). Entry times into asymptotically receding domains for ergodic Markov chains. *Theory Prob. Appl.* **19** 432-442.

[27] J. L. Lebowitz and R. H. Schonmann (1987). On the asymptotics of occurrence times of rare events for stochastic spin systems. *J. Stat. Phys.* **48** 727-751.

[28] E. Olivieri and E. Scoppola (1995). Markov chains with exponentially small transition probabilities: first exit problem from a general domain. I. The reversible case. *J. Stat. Phys.* **79** 613-647.

[29] E. Olivieri and E. Scoppola (1996) Markov chains with exponentially small transition probabilities: first exit problem from a general domain. II. The general case. *J. Statist. Phys.* **84** , no. 5-6, 987–1041.

[30] F. Paccaut (1999). Statistics of return times for weighted maps of the interval *Preprint Laboratorie de Topologie, Université de Bourgogne.*

[31] B. Pitskel (1991). poisson law for Markov chains. *Ergod. Th. Dyn. Sys.* **11**, 501-513.

[32] P. Shields (1996) The ergodic theory of discrete sample paths. *Graduate Studies in Mathematics* **Volume 13** A.M.S.

[33] A. Simonis (1998). Filling the hypercube in the supercritical contact process in equilibrium. *Markov Process. Rel. Fields* **4**1:113-130.

Miguel Abadi and Antonio Galves
Instituto de Matemática e Estatística
Universidade de São Paulo
BP 66281
05315-970 São Paulo, Brasil
e-mail: `abadi@ime.usp.br` and `galves@ime.usp.br`