SHARP ERROR TERMS AND NECCESARY CONDITIONS FOR EXPONENTIAL HITTING TIMES IN MIXING PROCESSES *

BY MIGUEL ABADI

Centre de Physique Téorique, CNRS, Luminy

Abstract

We prove an upper bound for the error in the exponential approximation of the hitting time law of a rare event in α-mixing processes with exponential decay, ϕ-mixing processes with a summable function ϕ and for general ψ-mixing processes with a finite alphabet. In the first case the bound is uniform as a function of the measure of the event. In the last two cases the bound depends also on the time scale t. This allow us to get further statistical properties as the ratio convergence of the expected hitting time and the expected return time. A uniform bound is a consequence. We present an example that shows that this bound is sharp. We also prove that second moments are not necessary for having the exponential law. Moreover, we prove a necessary condition for having the exponential limit law.

1 Introduction.

This paper proves an upper bound for the difference between the exponential law and the law of the first occurrence of a long string of symbols in a stochastic process with a finite alphabet. Our result stands for α-mixing processes with exponential decay, ϕ-mixing processes with a summable function ϕ and for general ψ-mixing processes.

We prove that for any n-cylinder set A, the law of its hitting time, suitably rescaled, can be uniformly approximated by a mean one exponential law. The error in the approximation of this law is bounded from above by $\inf_x\{x I\!P\{A\} + *(x)\}\ (t \vee 1)\ e^{-t}$, where $I\!P\{A\}$ is the measure of the cylinder, $*$ is the function of the decay of correlation of the process and t is the time scale.

We recall that an irreducible and aperiodic finite state Markov chain is ψ-mixing with exponential decay. Moreover, Gibbs states with a potential with exponential variations are

*This work was part of the author's Ph. D. *(Capes Ph. D. grant)* thesis under the advisoring of A. Galves who suggested the problem. 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 "Fenômenos Críticos em Probabilidade e Processos Estocásticos" (grant 41.96.0923.00).
Running title. Hitting times in mixing processes.
AMS subject classification. 60F05, 60G10, 37A50.
Key words and phrases. Mixing processes, occurrence time of a rare event, exponential approximation.

exponentially ψ-mixing. See Bowen (1975) for definitions and properties. Also, chain with complete connections, as defined in Bressaud, Fernandez and Galves (1999) are exponentially ψ-mixing. We refer to Doukhan (1995) for examples and references of α-mixing processes that are not ϕ-mixing, ϕ-mixing processes that are not ψ-mixing and ψ-mixing processes decaying at any rate.

The question of finding the limit law for the first occurrence of an event of small probability has a long history. The pioneer in this area was 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).

In the nineties several authors present uniform upper bounds for this approximation for processes with mixing dynamics. Among them Galves and Schmitt (1997), Hirata, Saussol and Vaienti (1998), Collet, Galves and Schmitt (1999) and Abadi (2001). For a detailed review on this works we refer the reader to Abadi and Galves (2001). Haydn (1999) proves the exponential approximation for rational maps. All those papers present *uniform* upper bounds for the exponential approximation to the hitting time law. However such a bound as a way of estimating the distribution is interesting only for small values of t.

Our technique is a modification of that in Galves and Schmitt (1997), but essential new ideas new ideas are introduced. One of them allow us to present a new point of view of the problem. We do not prove a uniform bound but rather a bound that depends also on the time scale t. This new approach provides more statistical properties. One of them is a bound for the decay of the tail of the distribution. Another important application of this non-uniform bound is the estimation of the expected *hitting* time. Kac's lemma (Kac, 1947) states that the expected *return* time is the inverse of the measure. So the right scaling factor for return times is $\mathbb{P}\{A\}$. There is no equivalent result for hitting time. We show that the scaling factor for hitting times can be written as $\xi_A \mathbb{P}\{A\}$, where ξ_A is bounded below and above by two strictly positive constants Ξ_1 and Ξ_2, respectively, independent of A, n and t. Moreover, the bound in function of t allow us to prove the convergence of the ratio between the expected hitting time and the expected return time to $1/\xi_A$. We recall that in Abadi (2001) it was presented an example in which $\xi_A \neq 1$. Also in In Asselah and Dai Pra (1997 a and b) the same was shown to hold for the symmetric simple exclusion process and for the spin flip system on $\mathbb{Z}$. A real computable estimation for the parameter ξ_A was proven in Abadi (2001). This estimate says also that ξ_A is smaller or equal that one. This proves that in general hitting times are longer than return times (see v.g. Shields (1996)). Therefore, this approximation is a powerful tool for computing the expected value of a hitting time. We emphasize that this holds for any kind of n-cylinder, even for those who can recur very fast.

A uniform upper bound, strictly sharper than the presented in previous mentioned works, can be immediately obtained from our result. This leads us to an important old question: what is the exact uniform rate of convergence to the exponential law. In order to solve this subject it should be presented a lower bound for the difference between hitting and exponential laws. A lower bound of order $\mathbb{P}\{A\}$ was proven in Abadi (2001). This order is reached asymptotically by the upper bound. However we are interested in the finite approximation. We present and example which shows that the *exact* rate of convergence is the one given by the uniform upper bound.

A common point in several works as Chen (1975), Galves and Schmitt (1997), Abadi (2001), is the control of the second moment of the function N, where N is the number of occurrences of the rare event. We prove that this is not necessary. Moreover we provided a *necessary and sufficient* condition for the convergence to the exponential law.

In the last section we consider the problem of the exponential limit law for α-mixing processes. Loosely, we prove that this limit holds if the tail of the coefficients α decay faster than the measure of the cylinders (see condition (22)). So, this kind of processes presents a substantial difference with respect to that of ψ and ϕ: the exponential does not holds for all the cylinder but for those who satisfy this condition.

This paper is organized as follows. In section 2 we establish our framework. In section 3 we state and prove the exponential approximation theorem for ψ-mixing and ϕ-mixing processes. In section 4 we present an example that shows the sharpness of the upper bound. In section 5 we prove the convergence of the ratio between the expected hitting time and the expected return time. In section 6 we prove that the second moment of N is not necessary for having an exponential law and we present a necessary and sufficient condition for it. In section 7 we prove the uniformly mixing case.

2 The framework

Let $\mathcal{E}$ be a finite set. Put $\Omega = \mathcal{E}^{\mathbb{Z}}$. For each $n \in \mathbb{Z}$, let $X_n : \Omega \to \Re$ be 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 by $\mathcal{F}_I$ the σ-algebra generated by cylinders with coordinates in I, $I \subseteq \mathbb{Z}$.

For a subset $A \subseteq \Omega$ we say that $A \in \mathcal{C}_n$ if and only if,

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

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

We consider a stationary probability measure $I\!\!P$ over $\mathcal{F}$. We shall assume that there are no singletons of probability 0 or 1.

Let $\alpha = (\alpha(l))_{l \geq 0},\ \phi = (\phi(l))_{l \geq 0}$ and $\psi = (\psi(l))_{l \geq 0}$ be three decreasing sequences converging to zero of positive real numbers. We shall say that the process $\{X_m\}_{m \in \mathbb{Z}}$ is *α-mixing* or uniform mixing if, for all integers $n \geq 1$ and $l \geq 0$, the following holds

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

ϕ*-mixing* if

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

and ψ*-mixing* if

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

where in the above expressions the supremum is taken over the sets B and C, such that $I\!\!P\{B\} > 0$ in the second case and such that $I\!\!P\{B\}I\!\!P\{C\} > 0$ in the third one.

Given $A \in \mathcal{C}_n$, we define the *entrance time* $\tau_A : \Omega \to I\!\!N \cup \{\infty\}$ as the following random variable defined on the probability space $(\Omega, \mathcal{F}, I\!\!P)$. For any $\omega \in \Omega$,

$$\tau_A(\omega) = \inf\{k \geq 1 : T^k(\omega) \in A\} \ .$$

Clearly, ψ-mixing implies ϕ-mixing that implies α-mixing. We recall that mixing (at any rate) implies ergodicity, and ergodicity ensures that τ_A is $I\!\!P$-almost surely finite (see e.g. Cornfeld, Fomin and Sinai, 1982). We remark that ϕ-mixing has positive entropy while *a priori* general α-mixing does not.

We shall use the classic probabilistic shorthand notation for events defined through random variables. We shall write $\{\tau_A = m\}$ instead of $\{\omega \in \Omega : \tau_A(\omega) = m\}$, $T^{-k}(A) = \{\omega \in \Omega : T^k(\omega) \in A\}$ and $\{X_r^s = x_r^s\} = \{X_r = x_r, \ldots, X_s = x_s\}$. Also we write for two measurables subsets A and B of Ω, the conditional probability of B given A as $I\!\!P\{B \mid A\} = I\!\!P_A\{B\} = I\!\!P\{B \cap A\} / I\!\!P\{A\}$.

As usual, the mean of a random variable X will be denoted by $I\!\!E(X)$. Wherever it is not ambiguous we will write C and c for different positive constants even in the same sequence of equalities/inequalities. Where a property holds for ϕ and ψ processes we shall replace ϕ or ψ by a $*$.

3 Hitting times for ψ-mixing and ϕ-mixing processes.

In this section we derive an upper bound for the difference between the re-scaled hitting time law and the exponential mean one law. Up to the present, the convergence to the exponential law was proved first without any rate of convergence and after with bounds depending on the event A but uniform on the time scale t.

Those kind of results, as a way of estimating the law of the hitting time, are interesting just for small values of t but are meaningless for large ones. We prove and exponential bound not uniform but depending also on t and derive further statistical properties from this fact.

3.1 Results

Theorem 1 *Let $\{X_m\}_{m\in\mathbb{Z}}$ be ψ-mixing or ϕ-mixing with ϕ summable. Then, there exist constants $C>0$, $0<\Xi_1<1<\Xi_2<\infty$, such that for all $n\in\mathbb{N}$, $A\in\mathcal{C}_n$ and $t>0$ there exists $\xi_A\in[\Xi_1,\Xi_2]$, for which the following inequality holds,*

$$\left|\mathbb{P}\left\{\tau_A>\frac{t}{\xi_A\mathbb{P}\{A\}}\right\}-e^{-t}\right|\le C\;\varepsilon(A)\;e^{-t}\,(t\vee 1)\;. \tag{1}$$

where $\varepsilon(A):=\inf_{n\le\Delta\le 1/(4\mathbb{P}\{A\})}[\Delta\mathbb{P}\{A\}+*(\Delta)]$.

Δ is the time the process needs to loose memory. This time must be large enough to allow the process loose memory. On the other hand, it must be small relative to the measure of A. What express the factor $\varepsilon(A)$ in the upper bound is that the rate of convergence to the exponential law is given by a trade off between the length of this time and the velocity of loosing memory of the process.

Remarks First, we emphasize the fact that the theorem holds *for all* cylinders $A\in\mathcal{C}_n$. Second, we remark the fact that the constants Ξ_1 and Ξ_2 are independent of n, A and t.

It is immediate a corollary that state a uniform bound for the velocity of convergence of the law of τ_A to the exponential law.

Corollary 2 *Under the above conditions, the following inequality holds,*

$$\sup_{t>0}\left|\mathbb{P}\left\{\tau_A>\frac{t}{\xi_A\mathbb{P}\{A\}}\right\}-e^{-t}\right|\le C\;\varepsilon(A)\;.$$

Proof. The maximum with respect to t of the upper bound in (1) is attained in $t=1$, and its value is $C\;\varepsilon(A)$.

The aboves theorem and corollary establish error bounds for the *finite* approximation of the hitting time to the exponential law. We could be interested in the limit order of convergence to the exponential law. In Saussol (2001) it was proved that this limit order is 1 *almost everywhere* for return times for Gibbs measures with Hölder continuous potential. As an easy corollary of

Theorem 1 we obtain that the same holds for the limit order of convergence of hitting times to the exponential law. The result holds for any exponentially ϕ-mixing process. In particular it holds for irreducible and aperiodic finite-state Markov chains and Gibbs measure with Hölder continuous potential. Moreover, in this case the result holds for *every point.*

Definition 3 *The* limit order of convergence *of the hitting time distribution to the exponential law is defined as*

$$\beta(\omega) = \lim_{n \to 0} \frac{\log d(n, \omega)}{\log I\!P\{A_n(\omega)\}} ,$$

(provided that the limit exists) where

$$d(n, \omega) = \sup_{t \geq 0} \left| I\!P \left\{ \tau_{A_n}(\omega) > \frac{t}{\xi_{A_n(\omega)} I\!P\{A_n(\omega)\}} \right\} - e^{-t} \right| ,$$

and where $A_n(\omega)$ is the n-cylinder $\{X_0^{n-1} = \omega_0^{n-1}\}$.

Corollary 4 *If the process $\{X_m\}_{m \in \mathbb{Z}}$ is exponentially ϕ-mixing, then the limit order of convergence of the hitting time distribution to the exponential law exists and $\beta \equiv 1$.*

Proof. By hypothesis $\phi(n) = Ce^{-cn}$, for some positive constants C, c. We take

$$x = -\frac{\log I\!P\{A_n(\omega)\}}{c} .$$

By Corollary 2, $d(n, \omega) \leq \varepsilon(A_n(\omega)) \leq x I\!P\{A_n(\omega)\} + \phi(x)$. Then $\beta(\omega) \geq 1$ for every ω. By Theorem 3 in Abadi (2001), $\beta(\omega) \leq 1$. The corollary follows.

3.2 The re-scaling parameter

For easy reference to the reader we quote in this section some lemmas that are proven in Abadi (2001) and that will be used in the proof of Theorem 1.

Lemma 5 *Under the conditions of Theorem 1, there exist strictly positive constants C and Γ such that for any fixed positive integer n and any $A \in \mathcal{C}_n$ the following inequality holds*

$$I\!P\{A\} \leq Ce^{-\Gamma n} .$$

Lemma 6 *Under the conditions of Theorem 1, there exists a strictly positive constant C', such that for any fixed positive integer n and any $A \in \mathcal{C}_n$ the following inequality holds,*

$$\sum_{k=1}^{n} I\!P\left\{A \cap T^{-k} A\right\} \leq C' I\!P\{A\} \ .$$

Lemma 7 *Let $\{X_m\}_{m\in\mathbb{Z}}$ be any stationary process. For any real number $f \geq 1$ the following inequality holds,*

$$I\!P\{\tau_A \leq f\} \leq f\, I\!P\{A\} \ . \tag{2}$$

For any positive integer f let us define,

$$N_f(\omega) = \sum_{l=1}^{f} 1\!\!1_{T^{-l}(A)}(\omega) \ ,$$

where $1\!\!1_A$ is the indicator function of the set A. For any $\omega \in \Omega$, $N_f(\omega)$ is the number of times the process visits A, during the first f steps. We remark that,

$$\{\tau_A \leq f\} = \{N_f \geq 1\} \ .$$

By Schwartz inequality,

$$(I\!E(N_f))^2 = \left(I\!E(N_f \ 1\!\!1_{\{N_f\geq 1\}})\right)^2 \leq I\!E(N_f^2) \ I\!E(1\!\!1^2_{\{N_f\geq 1\}}) = I\!E(N_f^2) \ I\!P\{\tau_A \leq f\} \ . \tag{3}$$

Lemma 8 *Under the conditions of Theorem 1, for any positive integer n, any cylinder $A \in \mathcal{C}_n$, and any $f > 0$, the following inequality holds,*

$$\frac{(I\!E(N_f))^2}{I\!E(N_f^2)} \geq \frac{[f]^2 I\!P\{A\}^2}{f\, I\!P\{A\} + C' f\, I\!P\{A\} + C(f\, I\!P\{A\})^2 + 2Kf\, I\!P\{A\}}, \tag{4}$$

where $C' > 0, C > 0$ e $K > 0$ are constants independent of n, A and f.

Let $f = f_A$ be a positive integer such that

$$f_A I\!P\{A\} \leq \frac{1}{2} \ . \tag{5}$$

Define

$$\lambda_{A,f} = \frac{-\log I\!P\left\{\tau_A > f_A\right\}}{f_A I\!P\{A\}} \ .$$

Lemma 9 *Under the conditions of Theorem 1, there exist two constants Λ_1, Λ_2, with $0 < \Lambda_1 \leq 1 \leq \Lambda_2 < \infty$, an a positive integer n_0, such that, for all $n \geq n_0$, all $A \in \mathcal{C}_n$, and all f as in (5) the following inequalities hold,*

$$\Lambda_1 \leq \lambda_{A,f} \leq \Lambda_2 \ ,$$

where Λ_1 and Λ_2 are independent of n, A and f_A.

3.3 The independence property

Lemma 10 *Let $\{X_m\}_{m\in\mathbb{Z}}$ be a ϕ-mixing process (at any rate). Let $A \in \mathcal{C}_n$ and $t, s \in I\!N$ such that $t > n$. For all $\Delta \in I\!N$, with $n \leq 2\Delta \leq t$, the following inequality holds,*

$$\left|I\!P\left\{\tau_A > t+s\right\} - I\!P\left\{\tau_A > t\right\} I\!P\left\{\tau_A > s\right\}\right| \leq 3 \ I\!P\left\{\tau_A > t - 2\Delta\right\} \ \left(I\!P\left\{\tau_A \leq \Delta\right\} + \phi(\Delta)\right\} \ .$$

Proof. We use the triangular inequality.

$$\begin{aligned}
& \left|I\!P\left\{\tau_A > t+s\right\} - I\!P\left\{\tau_A > t\right\} \ I\!P\left\{\tau_A > s\right\}\right| \\
\leq \ & \left|I\!P\left\{\tau_A > t+s\right\} - I\!P\left\{\tau_A > t - \Delta \bigcap \tau_A \circ T^t > s\right\}\right| \\
+ \ & \left|I\!P\left\{\tau_A > t - \Delta \bigcap \tau_A \circ T^t > s\right\} - I\!P\left\{\tau_A > t - \Delta\right\} I\!P\left\{\tau_A > s\right\}\right| \\
+ \ & \left|I\!P\left\{\tau_A > t - \Delta\right\} I\!P\left\{\tau_A > s\right\} - I\!P\left\{\tau_A > t\right\} I\!P\left\{\tau_A > s\right\}\right| \ .
\end{aligned}$$

The first term can be computed and then upper-bounded using the mixing property

$$
\begin{aligned}
& I\!P\left\{\tau_A > t-\Delta \bigcap \ \tau_A\circ T^{t-\Delta} \le \Delta \bigcap \ \tau_A \circ T^t > s\right\} \\
\le \ & I\!P\left\{\tau_A > t-2\Delta \bigcap \ \tau_A\circ T^{t-\Delta} \le \Delta\right\} \\
\le \ & I\!P\left\{\tau_A > t-2\Delta\right\} \ \left(I\!P\left\{\tau_A \le \Delta\right\} + \phi(\Delta)\right\} \ .
\end{aligned}
$$

We use the mixing property to bound the second term by $I\!P\{\tau_A > t-\Delta\} \ \phi(\Delta)$.

The third one is bounded as the first one. This ends the proof of the lemma.

The following proposition is an iterated version of the previous lemma.

Proposition 11 *Let* $\{X_m\}_{m\in\mathbb{Z}}$ *be a* ϕ*-mixing process. Let* $A \in \mathcal{C}_n$*. For any* f_A *such that*

$$
\phi(f_A/4) \le I\!P\left\{\tau_A \le f_A/4 \ \bigcap \ \tau_A\circ T^{f_A/4} > f_A/2\right\} \ , \tag{6}
$$

there exists a positive constants C *and a positive* $\Delta = \Delta_f$*, with* $n < \Delta \le f_A/4$*, such that for all positive integers* k*, the following inequalities hold,*

$$
\left| I\!P\left\{\tau_A > kf_A\right\} - I\!P\left\{\tau_A > f_A - 2\Delta\right\}^k \right| \le \varepsilon(A) \ k \ I\!P\left\{\tau_A > f_A - 2\Delta\right\}^k \ , \tag{7}
$$

and

$$
\left| I\!P\left\{\tau_A > kf_A\right\} - I\!P\left\{\tau_A > f_A\right\}^k \right| \le \varepsilon(A) \ k \ I\!P\left\{\tau_A > f_A - 2\Delta\right\}^k \ , \tag{8}
$$

where

$$
\varepsilon(A) = C \inf_{n \le x \le \frac{1}{I\!P\{A\}}} \left[x I\!P\{A\} + \phi(x) \right] \ .
$$

Remark 1. Both inequalities provide an approximation between the hitting time law and a geometric law for $t = kf_A$. The difference between them is that in the first one, the geometric inside the modulus is the *same* as in the upper bound, while in the second one, the geometric inside the modulus is *larger* than the one in the upper bound. That is, the second one gives a larger error. On the other hand it is technically clearer (since both quantities inside the modulus are in terms of f_A rather than $f_A - 2\Delta$). We will use both in the proof of Theorem 1.

Remark 2. Loosely,

$$\begin{aligned} I\!P\left\{\tau_A \le f_A/4 \bigcap \tau_A \circ T^{f_A/4} > f_A/2\right\} &\approx I\!P\left\{\tau_A \le f_A/4\right\} I\!P\left\{\tau_A \circ T^{f_A/4} > f_A/2\right\} \\ &\approx C\ f_A I\!P\{A\}\ , \end{aligned}$$

so, (6) holds if f_A is large enough to let the process evolve until $\phi(f_A/4)$ is small.

Proof. For $k = 1$ inequality (8) is obvious. Let us fix $\Delta = \Delta_f \in [n, f_A/4]$ such that,

$$\phi(\Delta) = I\!P\left\{\tau_A \le \Delta \bigcap \tau_A \circ T^{\Delta} > f_A - 2\Delta\right\}\ . \tag{9}$$

We notice that the left hand term of the above equation is decreasing on Δ while the right hand term is increasing on Δ. Then, such a Δ exists by condition (6). Moreover

$$\phi(\Delta) \le I\!P\left\{\tau_A \le \Delta\right\} \le \Delta I\!P\{A\}\ .$$

We denote with Δ_m the x that realizes $\inf_x\left[x I\!P\{A\} + \phi(x)\right]$. We have that $\Delta_f \ge \Delta_m$, then $\phi(\Delta_f) \le \phi(\Delta_m) \le \varepsilon(A)$. Therefore by stationarity, for $k = 1$ inequality (7) is

$$\begin{aligned} & I\!P\left\{\tau_A > f_A - 2\Delta\right\} - I\!P\left\{\tau_A > f_A\right\} \\ = \ & I\!P\left\{\tau_A \le 2\Delta \bigcap \tau_A \circ T^{2\Delta} > f_A - 2\Delta\right\} \\ = \ & I\!P\left\{\tau_A \le \Delta \bigcap \tau_A \circ T^{\Delta} > f_A - 2\Delta\right\} + I\!P\left\{\tau_A \le \Delta \bigcap \tau_A \circ T^{\Delta} > f_A - \Delta\right\} \\ \le \ & 2\ I\!P\left\{\tau_A \le \Delta \bigcap \tau_A \circ T^{\Delta} > f_A - 2\Delta\right\} \\ = \ & 2\ \phi(\Delta) \le\ \varepsilon(A)\ . \end{aligned} \tag{10}$$

Suppose $k \ge 2$. For each non negative integer i and for $j = 1, 2$, we write for a short-hand notation,

$$\mathcal{N}_j^i = \left\{\tau_A \circ T^{i\, f_A + j\, \Delta} > f_A - j\Delta\right\}\ ,$$

We also write for the sake of simplicity,

$$\mathcal{N} = \{\tau_A > f_A - 2\Delta\} \ .$$

First, we note that, by stationarity, $I\!P\{\mathcal{N}\} = I\!P\{\mathcal{N}_2^i\}$ for all non-negative integers i. We now look at inequality (7). By the triangular inequality,

$$\left| I\!P\{\tau_A > kf_A\} - I\!P\{\mathcal{N}\}^k \right| \tag{11}$$

$$\leq \sum_{j=0}^{k-2} I\!P\{\mathcal{N}\}^j \left| I\!P\{\tau_A > (k-j)f_A\} - I\!P\left\{\tau_A > (k-j-1)f_A \bigcap \mathcal{N}_2^{k-j-1}\right\}\right| \tag{12}$$

$$+ \sum_{j=0}^{k-2} I\!P\{\mathcal{N}\}^j \left| I\!P\left\{\tau_A > (k-j-1)f_A \bigcap \mathcal{N}_2^{k-j-1}\right\} - I\!P\{\tau_A > (k-j-1)f_A\} I\!P\left\{\mathcal{N}_2^0\right\}\right| \tag{13}$$

$$+ \ I\!P\{\mathcal{N}\}^{k-1} \left| I\!P\{\tau_A > f_A\} - I\!P\{\mathcal{N}\}\right| \ . \tag{14}$$

The idea of the proof is to iterate the proof of Lemma 10. In the most right term of (12) we introduced a gap (of length 2Δ). In the most right term of (13) we factored the measure on the left hand side in the modulus. Term (14) is just a small technical correction.

The most right factor in term (14) is

$$I\!P\left\{\mathcal{N} \bigcap \ \tau_A \circ T^{f_A - 2\Delta} \leq 2\Delta\right\} \ .$$

We can bound the most right factor in the sum (12) by,

$$I\!P\left\{ \bigcap_{i=0}^{k-j-2} \mathcal{N}_1^i \ \bigcap \ \left(\tau_A \circ T^{(k-j-1)f_A} \leq 2\Delta\right) \ \bigcap \ \mathcal{N}_2^{k-j-1} \right\} \ . \tag{15}$$

Similarly we bound the most right factor in the sum (13) by,

$$I\!P\left\{ \bigcap_{i=0}^{k-j-2} \mathcal{N}_1^i \right\} \phi(2\Delta - n) \leq I\!P\left\{ \bigcap_{i=0}^{k-j-2} \mathcal{N}_1^i \ \bigcap \ \Omega \right\} \phi(\Delta) \ . \tag{16}$$

Finally, we iterate the ϕ-mixing property to get the bound,

$$
\begin{aligned}
I\!\!P\left\{\bigcap_{i=0}^{\ell} \mathcal{N}_1^i \bigcap B\right\} &\leq I\!\!P\left\{\bigcap_{i=0}^{\ell-1} \mathcal{N}_1^i \bigcap B\right\} \\
&\leq \prod_{i=1}^{\ell-1}\left[I\!\!P\left\{\mathcal{N}_1^i\right\}+\phi(\Delta)\right]\left[I\!\!P\{B\}+\phi(\Delta)\right] \\
&= \left[I\!\!P\left\{\mathcal{N}_1^0\right\}+\phi(\Delta)\right]^{\ell-1}\left[I\!\!P\{B\}+\phi(\Delta)\right] \ , \qquad (17)
\end{aligned}
$$

for any measurable $B \in \mathcal{F}_{\{(\ell+1)f_A,(\ell+2)f_A+n-1\}}$. Furthermore, note that,

$$
\begin{aligned}
I\!\!P\{\mathcal{N}\} &= I\!\!P\{\tau_A > f_A - 2\Delta\} \\
&= I\!\!P\{\tau_A > f_A - \Delta\} + I\!\!P\left\{\tau_A \leq \Delta \bigcap \tau_A \circ T^{\Delta} > f_A - 2\Delta\right\} \\
&= I\!\!P\left\{\mathcal{N}_1^0\right\} + \phi(\Delta) \ .
\end{aligned}
$$

Thus, we apply the above inequality to (17) and we obtain that

$$
\begin{aligned}
I\!\!P\left\{\bigcap_{i=0}^{\ell} \mathcal{N}_1^i \bigcap B\right\} &\leq I\!\!P\{\mathcal{N}\}^{\ell-1}\left[I\!\!P\{B\}+\phi(\Delta)\right] \qquad (18) \\
&= I\!\!P\{\mathcal{N}\}^{\ell+1}\left[I\!\!P\{B\}+\phi(\Delta)\right] \ \frac{1}{I\!\!P\{\mathcal{N}\}^2} \\
&\leq C \ I\!\!P\{\mathcal{N}\}^{\ell+1}\left[I\!\!P\{B\}+\phi(\Delta)\right] \ .
\end{aligned}
$$

where the last inequality follows since

$$
I\!\!P\{\mathcal{N}\} \geq I\!\!P\{\tau_A > f_A\} \geq I\!\!P\left\{\tau_A > \frac{1}{2I\!\!P\{A\}}\right\} \geq 1/2 \ ,
$$

so, $1/I\!\!P\{\mathcal{N}\}^2 \leq C$. Now, we apply inequality (18) to inequality(16) with $B = \Omega$, and to inequality (15) with

$$
B = \left\{\tau_A \circ T^{(k-j-1)f_A} \leq 2\Delta \ \bigcap \ \mathcal{N}_2^{k-j-1}\right\} \ .
$$

To obtain the most right factor on the right hand side of inequality (7) we just need to sum over k. On the other hand,

$$\begin{aligned} I\!P\left\{B\right\} + \phi(\Delta) &= I\!P\left\{\tau_A \le 2\Delta \bigcap \mathcal{N}_2^0\right\} + \phi(\Delta) \\ &= I\!P\left\{\tau_A \le \Delta \bigcap \mathcal{N}_1^0\right\} + I\!P\left\{\tau_A \circ T^{\Delta} \le \Delta \bigcap \mathcal{N}_2^0\right\} + \phi(\Delta) \\ &\le 3\ \phi(\Delta) \le \varepsilon(A)\ . \end{aligned}$$

This concludes the proof of the upper bound (7).

Inequality (8) follows by the Mean Value Theorem and inequality (10).

$$\left| I\!P\left\{\tau_A > f_A - 2\Delta\right\}^k - I\!P\left\{\tau_A > f_A\right\}^k \right| \le k\ \varepsilon(A)\ I\!P\left\{\tau_A > f_A - 2\Delta\right\}^{k-1}\ .$$

This ends the proof of the proposition.

3.4 Proof of Theorem 1

Let us fix $f_A = 1/(2I\!P\{A\})$ and Δ_f the one given by Proposition 11. We define,

$$\xi_A = \frac{-\log I\!P\left\{\tau_A > f_A - 2\Delta_f\right\}}{f_A I\!P\{A\}}\ .$$

We note that,

$$\xi_A = \lambda_{A,f-2\Delta_f}\ \frac{f_A - 2\Delta_f}{f_A}\ . \tag{19}$$

So, $\xi_A \in [\Lambda_1/2, \Lambda_2]$, where Λ_1, Λ_2, are given in Lemma 9. Thus, we define $\Xi_1 = \Lambda_1/2$, and $\Xi_2 = \Lambda_2$.

The structure of the proof of inequality (1) has three steps. First we prove for t of the form $t = k\ f_A$ with k a positive integer. This is basically Proposition 11. So we go straightforward to the second step. There we prove for t of the form $t = (k+p/q)\ f_A$ with k, p positive integers and $1 \le p \le q$ with $q := 1/(2\varepsilon(A))$. The basic tools are the Mean Value Theorem and Proposition 11. Finally we prove for the remaining t's. Basically we approximate such a t by one of the form $(k + p/q)\ f_A$.

Proof: t's of the form $t = \left(k + \frac{p}{q}\right)\ f_A$.

Let $t = (k + (p/q))\ f_A$, with k, p, q as was just told. Put $r = (p/q)\ f_A$.

$$
\begin{aligned}
& \left| I\!P\{\tau_A > t\} - e^{-\xi_A I\!P\{A\} t} \right| \\
= & \left| I\!P\{\tau_A > kf_A + r\} - I\!P\{\tau_A > f_A - 2\Delta\}^{k + \, r/f_A} \right| \\
\le & \left| I\!P\{\tau_A > kf_A + r\} - I\!P\{\tau_A > kf_A\}\, I\!P\{\tau_A > r\} \right| \\
+ & \left| I\!P\{\tau_A > kf_A\} - I\!P\{\tau_A > f_A - 2\Delta\}^k \right|\ I\!P\{\tau_A > r\} \\
+ & \left| I\!P\{\tau_A > r\} - I\!P\{\tau_A > f_A - 2\Delta\}^{k + \, r/f_A} \right|\ I\!P\{\tau_A > f_A - 2\Delta\}^k \ .
\end{aligned}
$$

Let k be a non-negative integer. The first term on the right hand side of the above inequality is bounded using Lemma 10 and inequality (7). The last one can be used since obviously $f_A = 1/(2I\!P\{A\})$ satisfies condition (6).

$$
\begin{aligned}
& \left| I\!P\{\tau_A > kf_A + r\} - I\!P\{\tau_A > kf_A\}\, I\!P\{\tau_A > r\} \right| \\
\le & \ \varepsilon(A)\ I\!P\{\tau_A > kf_A - 2\Delta\} \\
\le & \ \varepsilon(A)\ I\!P\{\tau_A > (k-1) f_A\} \\
\le & \ \varepsilon(A)\ e^{-\xi_A I\!P\{A\} t} \ .
\end{aligned}
$$

The last inequality follows by inequality (18) with $B = \Omega$.

The second term is bounded by inequality (7) with $f_A = 1/(2I\!P\{A\})$,

$$
\begin{aligned}
& \left| I\!P\{\tau_A > kf_A\} - I\!P\{\tau_A > f_A - 2\Delta\}^k \right|\ I\!P\{\tau_A > r\} \\
\le & \ \varepsilon(A)\ k\ I\!P\{\tau_A > f_A - 2\Delta\}^k\ \ I\!P\{\tau_A > r\} \\
\le & \ \varepsilon(A)\ I\!P\{A\}\ t\ e^{-\xi_A I\!P\{A\} t} \ .
\end{aligned}
$$

The most left factor of the third term is bounded using the Mean Value Theorem applied to the function $h(x) = x^{1/q}$.

$$
\left| I\!P\{\tau_A > r\} - I\!P\{\tau_A > f_A - 2\Delta\}^{r/f_A} \right|
$$

$$
\begin{aligned}
&= \left| I\!P\left\{\tau_A > f_A \frac{p}{q}\right\}^q - I\!P\left\{\tau_A > f_A - 2\Delta\right\}^p \right| \times \frac{1}{q} \ \omega^{\frac{1}{q}-1} \\
&\leq \left| I\!P\left\{\tau_A > f_A \frac{p}{q}\right\}^q - I\!P\left\{\tau_A > f_A - 2\Delta\right\}^p \right| \times \qquad (20) \\
&\qquad\qquad \times \frac{1}{q} \min\left\{ I\!P\left\{\tau_A > f_A \frac{p}{q}\right\}^q ; I\!P\left\{\tau_A > f_A - 2\Delta\right\}^p \right\}^{\frac{1}{q}-1} \\
&= \frac{|b-a|}{a} \frac{1}{q} \ a^{\frac{1}{q}} ,
\end{aligned}
$$

where we call b and a the maximum and minimum respectively of the difference (20) and where ω is such that $a < \omega < b$. We have that $a^{1/q} \leq 1$, and

$$
\frac{|b-a|}{a} \frac{1}{q} \leq \frac{|b-c|}{a} \frac{1}{q} + \frac{|c-a|}{a} \frac{1}{q} ,
$$

for every c, real positive. We choose $c = I\!P\{\tau_A > p \ f_A\}$.

Consider now two cases. $i)$ When a is the most right term of the difference (20), and $ii)$ When a is the most left term of the difference (20).

Proof of case $i)$.

By equation (7) with $k = p$, and $f_A = 1/(2I\!P\{A\})$,

$$
\frac{|c-a|}{a} \frac{1}{q} \leq \frac{\varepsilon(A) \ p \ a}{a} \frac{1}{q} \leq \varepsilon(A) \ .
$$

Now we want to use inequality (8) with $k = q$, and $f_A = p/(2I\!P\{A\}q)$. Notice that to do this, we just need to verify that there exists a $\Delta = \Delta_{p,q} \in [n, p/(2I\!P\{A\}q)]$ that verifies condition (9). We observe that the term in the right hand side of (9) is actually a family function parametrized by f_A. Moreover, they are decreasing as a function of f_A. Then, such $\Delta_{p,q}$ exists and $\Delta_m < \Delta_{1,q} < \Delta_{2,q} < \ldots < \Delta_{q,q} = \Delta_{1/(2I\!P\{A\})}$. Furthermore, $\phi(\Delta_{p,q}) \leq \varepsilon(A)$. This inequality follows in the same way as was proved that $\phi(\Delta_f) \leq \varepsilon(A)$ at the beginning of the proof of Proposition 11, where also was defined Δ_m.

So, we use twice inequality (8). Finally we use inequality (7) with $k = p$, and with $f_A = 1/(2I\!P\{A\})$.

$$
\frac{|b-c|}{a} \frac{1}{q} \leq \frac{\varepsilon(A) \ q \ b}{a} \frac{1}{q} \leq \frac{\varepsilon(A)}{a} \frac{c}{1 - \ \varepsilon(A) \ q}
$$

$$\leq \frac{\varepsilon(A)}{a} \frac{1+\ \varepsilon(A)\ p}{1-\ \varepsilon(A)\ q}\ a \leq\ \varepsilon(A)\ .$$

The case *ii)* is treated in the same way. This proves that the third term is upper-bounded by,

$$C\ \varepsilon(A)\ I\!P\{\tau_A > f_A - 2\Delta\}^k \leq C\ \varepsilon(A)\ e^{-\xi_A I\!P\{A\}t}\ .$$

Proof: A general t.

Now, let t be any positive real. We write $t = k\ f_A + r$, with k a positive integer and r such that $0 \leq r < f_A$. We can choose a $\bar{t}$ such that $\bar{t} < t$ and $\bar{t} = (k + (p/q))\ f_A$ with p, q as before and such that $|t - \bar{t}| < \varepsilon(A)$.

$$\begin{aligned}
\left| I\!P\{\tau_A > t\} - e^{-\xi_A I\!P\{A\}t} \right| &\leq \left| I\!P\{\tau_A > t\} - I\!P\{\tau_A > \bar{t}\ \} \right| \\
&+ \left| I\!P\{\tau_A > \bar{t}\ \} - e^{-\xi_A I\!P\{A\}\bar{t}} \right| \\
&+ \left| e^{-\xi_A I\!P\{A\}\bar{t}} - e^{-\xi_A I\!P\{A\}t} \right|\ .
\end{aligned}$$

Since $I\!P\{\tau_A > t\}$ is constant in each interval $[k, k+1), k$ integer, we have,

$$\left| I\!P\{\tau_A > t\} - I\!P\{\tau_A > \bar{t}\ \} \right| = 0\ .$$

The upper bound for the third term is obtained using the Mean Value Theorem,

$$\begin{aligned}
\left| e^{-\xi_A I\!P\{A\}\bar{t}} - e^{-\xi_A I\!P\{A\}t} \right| &\leq \xi_A I\!P\{A\} \left(r - \frac{p}{q}\ f_A \right)\ e^{-\xi_A I\!P\{A\}\bar{t}} \\
&\leq \Lambda_2\ \varepsilon(A)\ e^{-\xi_A I\!P\{A\}t}\ .
\end{aligned}$$

Finally, the second term is bounded as in the first part of the proof. This, together with equation (19), ends the proof of the theorem.

4 Lower bound: an example of sharpness

Now we present an example that shows that the uniform bound established by Corollary 2 is sharp. For such an example, we must present a process, a rare event and a time t, for which the inequality actually is an equality.

Suppose that $\{X_m\}_{m\in\mathbb{Z}}$ is a Markov chain with state space $\mathcal{E} = \{0,1\}$, and transition matrix $Q = Q(i,j)$, with $i = 0,1;\ j = 0,1$; given by,

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

Consider the cylinder $A = \{X_0 = 0\}$. Denote with $I\!P$ the equilibrium measure of the chain. Then $I\!P\{A\} = q/(p+q)$. When $q << p$, we have $I\!P\{A\} << 1$. That is, A is rare. Then it is fair to ask for the difference of its hitting time law with respect to the exponential law. We fix $t = 1/I\!P\{A\}$. So, the proof of Theorem 1 for this time t, reduces to the proof of Proposition 11 with $k = 2$, and $f_A = 1/(2I\!P\{A\})$. We notice that the difference inside the modulus in (12) and (14) have always negative signal. Furthermore, a straightforward computation using the Markovian property both terms are equal to $C\ \Delta I\!P\{A\}$ for some positive constant C.

We must see that the difference inside the modulus in (13) is negative too. so in this case, the triangular inequality that bounds (11) is actually an equality. This is equivalent to see that

$$I\!P\left\{\mathcal{N}_2^{k-j-1} \mid \tau_A > (k-j-1)f_A\right\} \le I\!P\left\{\mathcal{N}_2^{0}\right\} .$$

Using stationarity and the definition of A, this is equivalents to see that

$$I\!P\left\{X_{2\Delta+1}^{f_A} = 1 \mid X^0_{-(k-j-1)f_A} = 1\right\} \le I\!P\left\{X_{2\Delta+1}^{f_A} = 1\right\} .$$

But using the Markovian property this is equivalent to

$$\frac{p}{p+q} + (1-(p+q))^{2\Delta+1}\left(\frac{q}{p+q}\right) = Q^{2\Delta+1}(1,1) \le I\!P\{X_0 = 1\} = \frac{p}{p+q} .$$

That occurs if and only if $p + q \ge 1$. We remark that the computing of Q^k for any positive integer k can be done recursively (see Ferrari and Galves (1997)). This holds for every Δ. It follows that in this case the uniform upper bound $\varepsilon(A)$ defined on Theorem 1 is actually an equality.

5 Application: estimation of $\mathbb{E}(\tau_A)$

An important result in ergodic theory is the famous Kac's Lemma that state 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. There is no equivalent result for the case of the expected hitting time.

We present in the next corollary an estimation for the expected hitting time as a consequence of the non-uniform bound given in Theorem 1. Roughly speaking, if the hitting time law converges to an exponential law with parameter ξ_A, we must have that $\xi_A = 1/\mathbb{E}(\tau_A)$, since the parameter is the inverse of the expected value for the exponential law.

Corollary 12 *Under the conditions of Theorem 1 the following inequality holds,*

$$|\mathbb{E}\left(\xi_A \mathbb{P}\{A\}\ \tau_A\right) - 1| \leq C\ \varepsilon(A)\ ,$$

or equivalently,

$$\left|\frac{\mathbb{E}\left(\tau_A\right)}{\mathbb{E}_A\left(\tau_A\right)} - \frac{1}{\xi_A}\right| \leq C\ \varepsilon(A)\ ,$$

where $\varepsilon(A)$ is the same as in Theorem 1.

Proof. Using the Mean Value Theorem,

$$e^{-(1-\varepsilon(A))t} - e^{-t} = \varepsilon(A)\ t\ e^{-(1-\chi_A)t} \geq \varepsilon(A)\ t\ e^{-t}\ .$$

with $\chi_A \in (1-\varepsilon(A), 1)$. Applying the above inequality to (1) and integrating we have,

$$\begin{aligned} |\mathbb{E}\left(\xi_A \mathbb{P}\{A\}\ \tau_A\right) - 1| &\leq C\ \left[\mathbb{E}\left(\exp\{1-\varepsilon(A)\}\right) - \mathbb{E}\left(\exp\{1\}\right) + \ \varepsilon(A)\mathbb{E}\left(\exp\{1\}\right)\ \right] \\ &= C\ \left[\frac{1}{1-\varepsilon(A)} - 1 + \varepsilon(A)\right] \\ &\leq C\ \varepsilon(A)\ , \end{aligned}$$

where we denote with $\exp\{\nu\}$ an exponential random variable with parameter ν. This ends the proof of the first inequality. The second one is an immediate consequence of the previous one together with Kac's Lemma (see Kac, 1947).

The following theorem proves a computable approximation for ξ_A. Also shows that this parameter actually is no larger than 1.

Let us fix s, a positive integer. We denote with $I\!P_A$ the conditional probability measure on the event A. We define,

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

Theorem 13 (Abadi, 2001) *Let $\{X_m\}_{m\in\mathbb{Z}}$ be exponentially ϕ-mixing. Let s be a positive integer. Then, there exist strictly positive constants Ψ_1, Ψ_2, C_1 and c such that for any n and any $A \in \mathcal{C}_n$, $\zeta_{A,s} \in [\Psi_1, \Psi_2]$, and the following inequality holds*

$$\left| I\!P\left\{\tau_A > \frac{t}{\zeta_{A,s} I\!P\{A\}}\right\} - e^{-t}\right| \le C_1 e^{-cn} . \tag{21}$$

Remark 1 In Galves and Schmitt (1990) it was shown that the limit $I\!E(\tau_{A_n})/b_n$ exists, where τ_{A_n} is the first time an exponentially ψ-mixing process hits an n-cylinder A_n, for some increasing constants b_n. Corollary 12 generalizes this result to a larger class of processes, specifies values for the sequence b_n and provide also a velocity of convergence for that limit.

Remark 2 Theorem 1 says that when ξ_A is used as re-scaling factor and the process is exponentially ϕ-mixing the error in the approximation of the hitting time law is exponencial. However this quantity depends upon a large part of the process, say $X_1, \ldots, X_{f_A - 2\Delta}$ (and typically $f_A - 2\Delta \approx e^{cn}$ for some positive constant c). Theorem 13 says that, with a bit larger error, that is an exponential error but with a smaller constant of exponentiation, we can use $\zeta_{A,s}$ instead of ξ_A as re-scaling factor. We remark that, a priori, $\zeta_{A,s}$ is an easy value to compute since it depends on a small number of coordinates of the process, say $1, \ldots, n + n/s$ and on the combinatorial properties of A. For instance, $\zeta_{A,s} = 1$ for every non recurrent cylinder.

Remark 3 In Theorem 1 we prove that $\xi_A \in [\Xi_1, \Xi_2] \ni 1$. Theorem 13 proves that $\zeta_{A,s} \in [\Psi_1, 1]$. We can deduce that, actually $\xi_A \in [\Xi_1, 1 + C_2 e^{-cn}]$, where C_2 and c are the same constants appearing in (21). This, together with Corollary 12, proves a fact that was already known as a folklore and commented in Shields (1996). That is, in general, hitting times are longer than return times.

Remark 4 In Abadi (2001) it was presented an example of a process and a cylinder A such that $\xi_A < 1$. Moreover, a process an a cylinder A can be constructed such that $\xi_A < \delta < 1$, for all $0 < \delta < 1$. In Asselah and Dai Pra (1997 a and b) the same was shown to hold for the symmetric simple exclusion process and for the spin flip system on $\mathbb{Z}$ with no interactions, when τ_A is the first time the empirical density in a large box exceeds its equilibrium value. In both cases $\xi_A \neq 1$.

6 Second moments not necessary

The proof of Lemma 9 shows that the uniform lower bound Λ_1 for the parameter $\lambda_{A,f}$, depends upon the control of the second moment of the function N_f. We recall that N_f counts the number of occurrences of the event A up to time f. The finiteness of this quantity is *suffice* to prove the uniform lower bound. We prove in the next proposition a *necessary and suffice* condition for the convergence to the exponential law. Roughly speaking it says that the moment that must be controlled is that of order "log plus one". So, second moments are not necessary.

Proposition 14 *Let $\{X_m\}_{m\in\mathbb{Z}}$ be a stochastic stationary process. Let t be a positive integer. Denote $N = N_t$. Then,*

$$\mathbb{P}\{\tau_A \leq t\} = \exp\left\{ -\mathbb{E}\left(\frac{N}{t\mathbb{P}\{A\}} \log \frac{N}{t\mathbb{P}\{A\}} \right) \right\} \ ,$$

where we adopt the convention $(N\log N)\ (0) = 0$.

Proof. By Hölder's inequality with p and q conjugated we have,

$$\mathbb{E}N = \mathbb{E}\left(N 1\!\!1_{\{Sopp\mathbb{N}\}} \right) \leq \mathbb{E}(N^p)^{1/p}\ \mathbb{P}\{N \geq 1\}^{1/q} \ ,$$

and the equality holds in the limit when p goes 1.

Now, since by definition $\mathbb{E}N = t\mathbb{P}\{A\}$, by a straightforward computation we have,

$$\mathbb{P}\{\tau_A \leq t\} = \mathbb{P}\{\mathbb{N} \geq 1\} = \lim_{p\to 1} \frac{1}{\left(\mathbb{E}\left(\frac{\mathbb{N}}{t\mathbb{P}\{A\}} \right)^p \right)^{1/(p-1)}} = \frac{1}{\exp\left\{ \mathbb{E}\left(\frac{\mathbb{N}}{t\mathbb{P}\{A\}} \log \frac{\mathbb{N}}{t\mathbb{P}\{A\}} \right) \right\}} \ .$$

This ends the proof of the proposition.

Corollary 15 *Let $\{X_m\}_{m\in\mathbb{Z}}$ be a stationary process. Let ξ_A defined as above. Let us denote $N = N_{\frac{1}{2\mathbb{P}\{A\}}}$. There exists. $\Xi_1 > 0$ such that $\Xi_1 \leq \xi_A$ for all $A \in \mathcal{C}_n$, if and only if, there exists a positive constant C such that $\mathbb{E}(N \log N) \leq C$. Moreover, in this case,*

$$\Xi_1 = \inf_A \exp\left\{ -\mathbb{E}\left(\mathbb{N} \log \mathbb{N} \right) \right\} \ .$$

Proof. We just need to notice that there is a positive constant C such that

$$C \; I\!P\left\{\tau_A \leq \frac{1}{2I\!P\{A\}}\right\} \leq \frac{1-e^{-\xi_A}}{2} \leq \xi_A \leq 1-e^{-\xi_A} \leq I\!P\left\{\tau_A \leq \frac{1}{2I\!P\{A\}}\right\} \; .$$

Then ξ_A remains bounded if and only if $I\!P\{\tau_A \leq 1/(2I\!P\{A\})\}$ does. Lemma 7 and the above proposition end the proof of the corollary.

7 Uniformly mixing processes

Uniformly mixing processes are a wide family of processes. Even not positive entropy is guaranteed for them. For instance, uniform mixing processes doesn't have, in general, the property of exponentially uniform decay of the cylinders as given in Lemma 5 for ϕ-mixing processes. Notice that, if this would be the case, they would have positive entropy and this is not the case. We recall that positive entropy is equivalent to the K-mixing property. On the other hand, there exist processes K-mixing, but not uniformly mixing (see v.g. definition of K-mixing, properties in Theorem 1, §8, Cap. 10; Cornfeld, Fomin and Sinai (1982)). Even so, still a sub-exponential uniform decay for the cylinders, in terms of the function α can be proved.

In such generality we can not expect to have the asymptotic exponential law. However, when the function α decays fast enough and the cylinder does not recur very fast we still can prove, with exponential rate, the convergence to the exponential law. We need the following definition of recurrence,

Definition 16 *Let $s \in I\!N$. For $1 \leq j \leq n/s$, define $\mathcal{B}_{n,j} = \mathcal{B}_{n,j}(s)$ as the set of $A \in \mathcal{C}_n$ which recur exactly at time j. Namely, $A \in \mathcal{B}_{n,j}$, if $A \cap T^{-j}A \neq \emptyset$, and $A \cap T^{-i}A = \emptyset$, for all $1 \leq i < j$. Define also $\mathcal{B}_n = \cup_{1\leq j \leq n/s}\mathcal{B}_{n,j}$.*

Theorem 17 *Let $\{X_m\}_{m\in\mathbb{Z}}$ be an α-mixing process with function α. Let s be a positive integer. Then, there exist strictly positive constants Λ_1, Λ_2 and C, such that for any $n \in I\!N$, and any $A \in \mathcal{E}^n \backslash \mathcal{B}_n(s)$ such that,*

$$\sum_{j=n/s}^{\sqrt{n\; I\!P\{A\}}} \alpha(j) \leq C \; I\!P\{A\} \; , \tag{22}$$

there exists a constant $\widetilde{\lambda_A} \in [\Lambda_1, \Lambda_2]$, and the following inequality holds,

$$\sup_{t\geq 0} \left| I\!P\left\{\tau_A > \frac{t}{\widetilde{\lambda_A} I\!P\{A\}}\right\} - e^{-t} \right| \leq C\sqrt{n \; I\!P\{A\}} \; . \tag{23}$$

One important difference between the exponential law for α-mixing processes in one hand and ϕ-mixing and ψ-mixing processes in the other is that in the second ones the convergence of the hitting time law to the exponential law holds for all cylinders while in the first case it holds for those cylinders which satisfy condition (22).

Another important difference between them is that in this case, we do not get upper bounds in function of t, as in the ϕ-mixing and ψ-mixing cases but uniform upper bound. So, for example, the estimation of the expectation of τ_A does not longer holds.

Loosely, we can see what does it mean the condition on the decay of the function α. If the process has positive entropy, then typically $\mathbb{P}\{A\} \approx e^{-hn}$ where h is the entropy. Moreover, if α decay exponentially fast with constant c, then $\sum_{j \geq n/s} \alpha(j) \approx Ce^{-cn/s}$. So, the hypothesis means $c/s > h$. That is, the system mix faster than the entropy.

On the other hand,

$$\#\mathcal{B}_n(s) = \sum_{j=1}^{n/s} \#\mathcal{B}_{n,j}(s) \leq \sum_{j=1}^{n/s} |\mathcal{E}|^j = \frac{|\mathcal{E}|^{n/s\ +1} - 1}{|\mathcal{E}| - 1} \ .$$

Thus we have,

$$\mathbb{P}\left\{\mathcal{B}_n(s)\right\} \leq \sum_{B \in \mathcal{B}_n(s)} \mathbb{P}\{B\} \approx \frac{\exp\left\{-n\left[h - \left(\log|\mathcal{E}|\right)/s\right]\right\}}{|\mathcal{E}| - 1} \ .$$

So, $\mathbb{P}\left\{\mathcal{B}_n(s)\right\} \to 0$, if $h_M/s < h$, where $h_M = \log|\mathcal{E}|$. Note that h_M is the maximum entropy the process could have. That is, we have an almost sure result if there is a positive real s such that $h_M < h\ s < c$.

Even when the uniform exponential decay does not hold for uniformly mixing processes we can prove the sub-exponential decay of them.

Lemma 18 *Let* $\{X_m\}_{m \in \mathbb{Z}}$ *be a* α*-mixing process. There exist strictly positive constants* C *and* Γ *such that for all positive integer* n *and for all* $A \in \mathcal{C}_n$ *the following inequality holds,*

$$\mathbb{P}\{A\} \leq C\left[e^{-\Gamma\sqrt[3]{n}} + \gamma(\sqrt[3]{n})\right] \ .$$

Proof. Let us write $A = \{X_0^{n-1} = a_0^{n-1}\}$. We note that $n = [\sqrt{n}\ ]^2 + r$, with $0 \leq r \leq 2n$. We have

$$\mathbb{P}\left\{X_0^{n-1}=a_0^{n-1}\right\} \leq \mathbb{P}\left\{X_0=a_0, X_{[\sqrt{n}\,]}=a_{[\sqrt{n}\,]}, X_{2[\sqrt{n}\,]}=a_{2[\sqrt{n}\,]}, \ldots\right\} .$$

Using the mixing property,

$$\mathbb{P}\{A\} \leq \rho^{[\sqrt{n}\,]}+\gamma([\sqrt{n}\,]-1) \frac{1-\rho^{[\sqrt{n}\,]}}{1-\rho} ,$$

where $\rho=\sup\{\mathbb{P}\{a_i\}: a_i \in \mathcal{E}\}$, and by hypothesis, $\rho<1$. Now, it is enough to note that, $\sqrt[3]{n} \leq[\sqrt{n}\,]-1$ for all $n>n_0$. This ends the proof of the lemma.

Proof of Theorem 17. Let us define $\widetilde{\lambda_A}=\widetilde{\lambda}_{A,f,\Delta}$ with Δ, and f_A, satisfying $n<\Delta \leq f_A/4$. The proof of the upper bound for the parameter $\widetilde{\lambda_A}$ follows as in Lemma 9.

To obtain the lower bound for $\widetilde{\lambda_A}$ it is enough by (3) to find a lower bound for the left-hand term in inequality (4).

$$\begin{aligned}
\mathbb{E}(N_f^2) = & \sum_{\ell=1}^{[f]} \mathbb{E}\left(\mathbb{1}_{T^{-\ell}(A)}^2\right) \\
& +2 \sum_{\ell=1}^{n/s}([f]-\ell)\, \mathbb{E}\left(\mathbb{1}_A\, \mathbb{1}_{T^{-\ell}(A)}\right) \\
& +2 \sum_{\ell=n/s+1}^{2n}([f]-\ell)\, \mathbb{E}\left(\mathbb{1}_A\, \mathbb{1}_{T^{-\ell}(A)}\right) \\
& +2 \sum_{\ell=2n+1}^{[f]}([f]-\ell)\, \mathbb{E}\left(\mathbb{1}_A\, \mathbb{1}_{T^{-\ell}(A)}\right) .
\end{aligned}$$

By definition the first term in the above decomposition is $[f]\mathbb{P}\{A\}$ and the second, by hypothesis, is zero. The remaining terms are bounded using the mixing property. The third term is bounded by,

$$\begin{aligned}
& \sum_{\ell=n/s+1}^{2n}([f]-\ell)\left[\mathbb{P}\{A\} \mathbb{P}\left\{A^{(n/2s)}\right\}+\gamma\left(\ell-\frac{n}{2s}\right)\right] \qquad (24)\\
\leq\ & f\, \mathbb{P}\{A\}\, n\, \mathbb{P}\left\{A^{(n/2s)}\right\}+f \sum_{j=n/2s}^{(2n-n)/2s} \gamma(j) .
\end{aligned}$$

Finally, the last term is, bounded by,

$$\sum_{\ell=2n+1}^{[f]} ([f]-\ell)\left[I\!P\{A\}^2+\gamma(\ell-n)\right] \le \frac{f^2\, I\!P\{A\}^2}{2} + f \sum_{j=n+1}^{[f]-n} \gamma(j)\ . \tag{25}$$

By Lemma 8, we must prove that the terms on the right-hand side of (24) and (25) are smaller than or equal to $f I\!P\{A\}$. By Lemma 18, $n\, I\!P(A^{(n/2s)}) \to 0$. Immediately, we conclude that there is $0 < \Lambda_1 < \widetilde{\lambda_A}$, for all $A \in \mathcal{C}_n$ and for all n if there is a constant C, such that

$$\frac{\sum_{j=n/2s}^{t} \gamma(j)}{I\!P\{A\}} \le C < \infty\ .$$

This ends the proof of the bounds for $\widetilde{\lambda_A}$. Now we note that,

$$\begin{aligned} I\!P\left\{\tau_A \le (f_A - 2\Delta)\right\} &\ge \frac{\left(I\!E(N_{f_A-2\Delta})\right)^2}{I\!E(N^2_{f_A-2\Delta})} \\ &\ge C(f_A-2\Delta) I\!P\{A\} \\ &\ge C f_A I\!P\{A\}\ . \end{aligned} \tag{26}$$

The first inequality follows by (3). The second one by the above computations. The third one since $\Delta \le f_A/4$.

For proving inequality (23), we follow the triangle inequality (11), emphasizing that it holds for any $n \le \Delta \le f_A/4$ without assuming (6).

Proof: t's of the form $t = k\ f_A$.

For $t = kf_A$, with k positive integer, we bound the sum (12) by,

$$\sum_{j=0}^{k-2} I\!P\left\{\mathcal{N}\right\}^j \Delta I\!P\{A\} \le \frac{\Delta I\!P\{A\}}{1 - I\!P\left\{\mathcal{N}\right\}} \le C\frac{\Delta I\!P\{A\}}{f_A I\!P\{A\}}\ ,$$

where the last inequality follows by (26). Similarly, we bound the sum (13) by,

$$\sum_{j=0}^{k-2} I\!P\left\{\mathcal{N}\right\}^j \alpha(\Delta) \le \frac{\alpha(\Delta)}{1 - I\!P\left\{\mathcal{N}\right\}} \le C\frac{\alpha(\Delta)}{f_A I\!P\{A\}}\ .$$

Finally, we bound term (14) by

$$\mathbb{P}\{\mathcal{N}\}^{k-1}\,\Delta\mathbb{P}\{A\} \leq C\frac{\Delta\mathbb{P}\{A\}}{f_A\mathbb{P}\{A\}}\;.$$

Thus, summing the three bounds above we get that,

$$\left|\mathbb{P}\{\tau_A > kf_A\} - \mathbb{P}\{\mathcal{N}\}^k\right| \leq C\frac{\Delta\mathbb{P}\{A\} + \alpha(\Delta)}{f_A\mathbb{P}\{A\}}\;.$$

Proof: A general t.

Now, for a time t of the form $t = kf_A + r$, with k and r integers, $0 \leq r < f_A$, we write

$$\begin{aligned}\left|\mathbb{P}\{\tau_A > t\} - e^{-\widetilde{\lambda_A}\mathbb{P}\{A\}t}\right| &\leq \left|\mathbb{P}\{\tau_A > kf_A + r\} - \mathbb{P}\{\tau_A > kf_A\}\right| \\ &+ \left|\mathbb{P}\{\tau_A > kf_A\} - \mathbb{P}\{\mathcal{N}\}^k\right| \\ &+ \left|e^{-\widetilde{\lambda_A}\mathbb{P}\{A\}kf_A} - e^{-\widetilde{\lambda_A}\mathbb{P}\{A\}t}\right|\;.\end{aligned}$$

The first term is bounded using equation (2) and the mixing property,

$$\begin{aligned}\left|\mathbb{P}\{\tau_A > kf_A + r\} - \mathbb{P}\{\tau_A > kf_A\}\right| &= \mathbb{P}\left\{\tau_A > kf_A \bigcap \tau_A \circ T^{kf_A} \leq r\right\} \\ &\leq \mathbb{P}\{\tau_A \leq r\} \\ &\leq r\,\mathbb{P}\{A\} \\ &\leq f_A\mathbb{P}\{A\}\;. \end{aligned} \tag{27}$$

A similar bound is obtained for the third term using the Mean Value Theorem,

$$\begin{aligned}\left|e^{-\widetilde{\lambda_A}\mathbb{P}\{A\}kf_A} - e^{-\widetilde{\lambda_A}\mathbb{P}\{A\}t}\right| &\leq \widetilde{\lambda_A}\mathbb{P}\{A\}\;r\;e^{-\widetilde{\lambda_A}\mathbb{P}\{A\}kf_A} \\ &\leq C\Lambda_2\mathbb{P}\{A\}\;f_A\;e^{-\widetilde{\lambda_A}\mathbb{P}\{A\}t} \\ &\leq C\Lambda_2\mathbb{P}\{A\}\;f_A\;.\end{aligned}$$

The sum of the three bounds gives the expression for the bound for all t,

$$C\left[\frac{\Delta I\!P\{A\}+\alpha(\Delta)}{f_A I\!P\{A\}}+f_A I\!P\{A\}\right]\ .$$

To end the proof we just need to optimize the above expression. Note that the Δ that realize $\inf_{n\le\Delta\le 1/(4I\!P\{A\})}\{\Delta I\!P\{A\}+\alpha(\Delta)\}$, is n since by hypothesis,

$$\frac{\alpha(n)}{n}<\alpha(n)<\sum_{j=n/s}^{\sqrt{n\ I\!P\{A\}}}\alpha(j)\le C\ I\!P\{A\}\ .$$

For that chosen Δ, we choose f_A such that,

$$\frac{n}{f_A}=f_A I\!P\{A\}\ .$$

This ends the proof of the theorem.

References

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

[2] 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.

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

[4] A. Asselah and P. Dai Pra (1997b). Occurrence of rare events in ergodic interacting spin systems. *Ann. Inst. H. Poincaré* **33** 6, 727-751.

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

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

[7] X. Bressaud, R. Fernandez and A. Galves (1999). Speed of $\overline{d}$-convergence for Markov approximations of chains with complete connections. A coupling approach. *Stochastic Process. Appl.* **83**. 1, 127–138.

[8] L. Chen (1975). Poisson approximation for dependent trials. *Ann. Probab.* **3**, 534-345.

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

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

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

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

[13] P. A. Ferrari and A. Galves (1997). Acoplamento e Processos Estocásticos. *21 Colóquio Brasileiro de Matemática* IMPA.

[14] A. Galves and B. Schmitt (1990). Occurrence time of rare events for mixing dynamical systems. *Ann. Inst. H. Poincaré* **52**, 3, 267-281.

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

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

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

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

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

[20] B. Saussol (2001). On fluctuations and the exponential statistics of return times. *Nonlinearity.* **14** 179-191.

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

CENTRE DE PHYSIQUE TÉORIQUE
LUMINY, CASE 907
13288 MARSEILLE CEDEX 9, FRANCE
E-MAIL: abadi@cpt.univ-mrs.fr