*The Annals of Probability*
2004, Vol. 32, No. 1A, 243–264


# SHARP ERROR TERMS AND NECCESARY CONDITIONS FOR EXPONENTIAL HITTING TIMES IN MIXING PROCESSES[1]

BY MIGUEL ABADI

*University of São Paulo*

We prove an upper bound for the error in the exponential approximation of the hitting time law of a rare event in $\alpha$-mixing processes with exponential decay, $\phi$-mixing processes with a summable function $\phi$ and for general $\psi$-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 allows 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 $\alpha$-mixing processes with exponential decay, $\phi$-mixing processes with a summable function $\phi$ and for general $\psi$-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 $\inf_x\{x\mathbb{P}\{A\} + *(x)\} \times (t \vee 1)e^{-t}$, where $\mathbb{P}\{A\}$ is the measure of the cylinder, $*$ stands for $\phi$ or $\psi$, according to the assumed mixing properties of the process and $t$ is the time scale.

We recall that an irreducible and aperiodic finite state Markov chain is $\psi$-mixing with exponential decay. Moreover, Gibbs states with a Hölder continuous potential are exponentially $\psi$-mixing. See Bowen (1975) for definitions and properties. Also, chains with complete connections, as defined in Bressaud, Fernandez and Galves (1999) are exponentially $\psi$-mixing. We refer to Doukhan (1995) for examples and references of $\alpha$-mixing processes that are not $\phi$-mixing, $\phi$-mixing processes that are not $\psi$-mixing and $\psi$-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 Doeblin (1940), who

Received December 2001; revised January 2003.

[1]Supported in part by FAPESP Grant 98-3382-0 and Grant 99-11962-9 and by a fellowship of the Ministere de la Recherche-Ministere de l'Education Nationale, France.

*AMS 2000 subject classifications.* 60F05, 60G10, 37A50.

*Key words and phrases.* Mixing processes, hitting times, repetition times, rare event, exponential approximation.

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 Bellman and Harris (1951) and Harris (1953).

In the nineties several authors presented uniform upper bounds for this approximation for processes with mixing dynamics. Among them Galves and Schmitt (1997), Hirata, Saussol and Vaienti (1999), Collet, Galves and Schmitt (1999) and Abadi (2001). For a detailed review on these 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, say $\delta(A)$, for the exponential approximation to the hitting time law of a cylinder set $A$.

Our technique is a modification of that in Galves and Schmitt (1997), but essential new ideas are introduced. One of them allows 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 result gives a bound for the decay of the tail of the distribution. Our time-dependent bound is relevant for $t$ of order $t < 1/\varepsilon(A)$, with $\varepsilon(A)$ defined in Theorem 1, while the previous uniform bounds, as a way of estimating the distribution, are interesting only for values of $t$ of order $t \ll \log(1/\delta(A))$. Moreover, it tourns out that $\varepsilon(A) \ll \delta(A)$.

Another important application of this nonuniform 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 similar result for hitting time. We show that the scaling factor for hitting times can be written as $\xi_A \mathbb{P}\{A\}$, where $\xi_A$ is bounded below and above by two strictly positive constants $\Xi_1$ and $\Xi_2$, respectively, independent of $A$, $n$ and $t$. Moreover, the bound as a function of $t$ allow us to prove the convergence to $1/\xi_A$ of the ratio between the expected hitting time and the expected return time. We recall that in Abadi (2001) we presented an example in which $\xi_A \neq 1$. Also, in Asselah and Dai Pra (1997a, b) the same was shown to hold for the symmetric simple exclusion process and for the independent spin flip system on $\mathbb{Z}$. A real computable estimate for the parameter $\xi_A$ was proven in Abadi (2001). This estimate also says that $\xi_A$ is smaller than or equal to one. This proves that in general hitting times are longer than return times [see, e.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 which can recur very fast.

A uniform upper bound, strictly sharper than those presented in the previously 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 answer this question it should be presented as 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 an example which shows that the *exact* rate of convergence is the one given by the uniform upper bound.

A common point in several works such as Chen (1975), Galves and Schmitt (1997) and 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 provide a weaker condition for the convergence to the exponential law.

In the last section we consider the problem of the exponential limit law for $\alpha$-mixing processes. Loosely speaking, we prove that this limit holds if the tail of the coefficients $\alpha$ decays faster than the measure of the cylinders [see condition (22)]. So, these kind of processes present a substantial difference with respect to that of $\psi$ and $\phi$: the exponential does not hold for all the cylinders, but only for those which 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 $\psi$-mixing and $\phi$-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 $X_n : \Omega \to \mathcal{E}$ 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 $\sigma$-algebra over $\Omega$ generated by cylinders. Moreover, we denote by $\mathcal{F}_I$ the $\sigma$-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,

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

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

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

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

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

*$\phi$-mixing* if

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

and $\psi$*-mixing* if

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

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.

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

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

Clearly, $\psi$-mixing implies $\phi$-mixing that implies $\alpha$-mixing. We recall that mixing implies ergodicity, and ergodicity ensures that $\tau_A$ is $\mathbb{P}$-almost surely finite provided that $\mathbb{P}\{A\} > 0$, [see, e.g., Cornfeld, Fomin and Sinai (1982)].

We shall use the classical 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 $\{X_r^s = x_r^s\} = \{X_r = x_r, \dots, X_s = x_s\}$. Also we write for two measurable subsets $A$ and $B$ of $\Omega$, the conditional probability of $B$ given $A$ as $\mathbb{P}\{B \mid A\} = \mathbb{P}_A\{B\} = \mathbb{P}\{B\cap A\}/\mathbb{P}\{A\}$.

As usual, the mean of a random variable $X$ will be denoted by $\mathbb{E}(X)$. 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 $\phi$ and $\psi$ processes we shall replace $\phi$ or $\psi$ by a $*$.

## 3. Hitting times for $\psi$-mixing and $\phi$-mixing processes.

In this section we derive an upper bound for the difference between the rescaled hitting time law and the mean one exponential 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 $t$.

We prove an 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 $\psi$-mixing or $\phi$-mixing with $\phi$ 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*:

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

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

$\Delta$ is the time the process needs to lose memory. This time must be large enough to allow the process to lose memory. On the other hand, it must be small relative to the measure of $A$. The factor $\varepsilon(A)$ in the upper bound shows 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 losing 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 $\Xi_1$ and $\Xi_2$ are independent of $n$, $A$ and $t$.

There is an immediate corollary that states a uniform bound for the velocity of convergence of the law of $\tau_A$ 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 $t = 0$, and its value is $C\varepsilon(A)$. $\square$

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

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

(provided that the limit exists) where

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

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

The theorem and corollary stated above 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 $\phi$-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*.

COROLLARY 4. *If the process* $\{X_m\}_{m\in\mathbb{Z}}$ *is exponentially* $\phi$*-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

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

By Corollary 2, $d(n,\omega) \le \varepsilon(A_n(\omega)) \le x\mathbb{P}\{A_n(\omega)\} + \phi(x)$. Then $\beta(\omega) \ge 1$ for every $\omega$. By Theorem 3 in Abadi (2001), $\beta(\omega) \le 1$. The corollary follows. □

3.2. *Basic bounds on the measure of cylinders.* 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* $\Gamma$, *such that for any fixed positive integer* $n$ *and any* $A \in \mathcal{C}_n$, *the following inequality holds*:

$$\mathbb{P}\{A\} \le Ce^{-\Gamma n}.$$

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*:

$$\sum_{k=1}^{n} \mathbb{P}\{A \cap T^{-k}A\} \le C'\mathbb{P}\{A\}.$$

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

$$\mathbb{P}\{\tau_A \le f\} \le f\mathbb{P}\{A\}. \tag{2}$$

For any positive integer $f$ let us define,

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

where $\mathbb{1}_A$ 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

$$\{\tau_A \le f\} = \{N_f \ge 1\}.$$

By the Schwarz inequality,

$$(\mathbb{E}(N_f))^2 = \bigl(\mathbb{E}\bigl(N_f \mathbb{1}_{\{N_f\ge 1\}}\bigr)\bigr)^2 \le \mathbb{E}(N_f^2)\mathbb{E}\bigl(\mathbb{1}^2_{\{N_f\ge 1\}}\bigr) = \mathbb{E}(N_f^2)\mathbb{P}\{\tau_A \le f\}. \tag{3}$$

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*:

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

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

For any cylinder set $A$ and $f > 0$ define

$$\lambda_{A,f} = \frac{-\log \mathbb{P}\{\tau_A > f\}}{f\mathbb{P}\{A\}}.$$

LEMMA 9. *Under the conditions of Theorem* 1, *there exist two constants* $\Lambda_1, \Lambda_2$, *with* $0 < \Lambda_1 \leq 1 \leq \Lambda_2 < \infty$, *and a positive integer* $n_0$, *such that*, *for all* $n \geq n_0$, *all* $A \in \mathcal{C}_n$ *and* $f \leq 1/(2\mathbb{P}\{A\})$, *the following inequalities hold*:

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

*where* $\Lambda_1$ *and* $\Lambda_2$ *are independent of* $n$, $A$ *and* $f$.

3.3. *The independence property.*

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

$$\begin{aligned}&|\mathbb{P}\{\tau_A > t+s\} - \mathbb{P}\{\tau_A > t\}\mathbb{P}\{\tau_A > s\}|\\ &\qquad \leq 3\mathbb{P}\{\tau_A > t - 2\Delta\}(\mathbb{P}\{\tau_A \leq \Delta\} + \phi(\Delta)\}.\end{aligned}$$

PROOF. We use the triangle inequality

$$\begin{aligned}&|\mathbb{P}\{\tau_A > t+s\} - \mathbb{P}\{\tau_A > t\}\mathbb{P}\{\tau_A > s\}|\\ &\quad \leq |\mathbb{P}\{\tau_A > t+s\} - \mathbb{P}\{\tau_A > t - \Delta \cap \tau_A \circ T^t > s\}|\\ &\qquad + |\mathbb{P}\{\tau_A > t - \Delta \cap \tau_A \circ T^t > s\} - \mathbb{P}\{\tau_A > t - \Delta\}\mathbb{P}\{\tau_A > s\}|\\ &\qquad + |\mathbb{P}\{\tau_A > t - \Delta\}\mathbb{P}\{\tau_A > s\} - \mathbb{P}\{\tau_A > t\}\mathbb{P}\{\tau_A > s\}|.\end{aligned}$$

The first term can be computed and then bounded from above using the mixing property

$$\begin{aligned}&\mathbb{P}\{\tau_A > t - \Delta \cap \tau_A \circ T^{t-\Delta} \leq \Delta \cap \ \tau_A \circ T^t > s\}\\ &\quad \leq \mathbb{P}\{\tau_A > t - 2\Delta \cap \tau_A \circ T^{t-\Delta} \leq \Delta\}\\ &\quad \leq \mathbb{P}\{\tau_A > t - 2\Delta\}(\mathbb{P}\{\tau_A \leq \Delta\} + \phi(\Delta)\}.\end{aligned}$$

We use the mixing property to bound the second term by $\mathbb{P}\{\tau_A > t - \Delta\}\phi(\Delta)$. 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 $\phi$-mixing process. There exists a finite constant $C > 0$, such that for any $n$, $A \in \mathcal{C}_n$ and any $f \in (4n, 1/(2\mathbb{P}\{A\}))$, such that*

$$\phi(f/4) \le \mathbb{P}\{\tau_A \le f/4 \cap \tau_A \circ T^{f/4} > f/2\}. \tag{5}$$

*There exists a $\Delta > 0$, with $n < \Delta \le f/4$, such that for all positive integers $k$, the following inequalities hold*:

$$\left|\mathbb{P}\{\tau_A > kf\} - \mathbb{P}\{\tau_A > f - 2\Delta\}^k\right| \le C\varepsilon(A)k\mathbb{P}\{\tau_A > f - 2\Delta\}^k \tag{6}$$

*and*

$$\left|\mathbb{P}\{\tau_A > kf\} - \mathbb{P}\{\tau_A > f\}^k\right| \le C\varepsilon(A)k\mathbb{P}\{\tau_A > f - 2\Delta\}^k. \tag{7}$$

We recall the definition of $\varepsilon(A)$ in Theorem 1.

REMARK 1. Both inequalities provide an approximation between the hitting time law and a geometric law for $t = kf$. 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 $f$ rather than $f - 2\Delta$). We will use both in the proof of Theorem 1.

REMARK 2. Loosely speaking,

$$\begin{aligned}\mathbb{P}\{\tau_A \le f/4 \cap \tau_A \circ T^{f/4} > f/2\} &\approx \mathbb{P}\{\tau_A \le f/4\}\mathbb{P}\{\tau_A \circ T^{f/4} > f/2\}\\ &\approx Cf\mathbb{P}\{A\}.\end{aligned}$$

The first approximation follows by a similar argument to the proof of Lemma 10. The second one follows by Lemma 9. So, for $f = 1/(2\mathbb{P}\{A\})$, the left-hand side of (5) is small while the right-hand side is of order of a constant. Therefore, an $f$ satisfying (5) exists.

PROOF OF PROPOSITION 11. For $k = 1$ (7) is obvious. Let us fix $\Delta = \Delta_f \in [n, f/4]$ such that,

$$\phi(\Delta) = \mathbb{P}\{\tau_A \le \Delta \cap \tau_A \circ T^{\Delta} > f - 2\Delta\}. \tag{8}$$

We notice that the left-hand side of the above equality is decreasing on $\Delta$, while the right-hand side is increasing on $\Delta$. Then, such a $\Delta$ exists by condition (5). Moreover,

$$\phi(\Delta) \le \mathbb{P}\{\tau_A \le \Delta\} \le \Delta\mathbb{P}\{A\}.$$

We denote by $\Delta_m$ the $x$ that realizes $\inf_x[x\mathbb{P}\{A\}+\phi(x)]$. In other words, $\phi(\Delta_m)=\Delta_m\mathbb{P}\{A\}$. Since, $\phi(x)$ is decreasing on $x$ and $x\mathbb{P}\{A\}$ is increasing on $x$, we have $\phi(\Delta)\le\phi(\Delta_m)\le\varepsilon(A)$. Therefore, by stationarity, for $k=1$, (6) is

$$
\begin{aligned}
&\mathbb{P}\{\tau_A > f-2\Delta\} - \mathbb{P}\{\tau_A > f\}\\
&\qquad = \mathbb{P}\{\tau_A \le 2\Delta \cap \tau_A\circ T^{2\Delta} > f-2\Delta\}\\
&\qquad = \mathbb{P}\{\tau_A \le \Delta \cap \tau_A\circ T^{\Delta} > f-2\Delta\} + \mathbb{P}\{\tau_A \le \Delta \cap \tau_A\circ T^{\Delta} > f-\Delta\}\\
&\qquad \le 2\mathbb{P}\{\tau_A \le \Delta \cap \tau_A\circ T^{\Delta} > f-2\Delta\}\\
&\qquad = 2\phi(\Delta) \le \varepsilon(A).
\end{aligned} \tag{9}
$$

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

$$\mathcal{N}_j^i = \{\tau_A\circ T^{if+j\Delta} > f-j\Delta\}.$$

We also write for the sake of simplicity

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

First, we note that, by stationarity, $\mathbb{P}\{\mathcal{N}\}=\mathbb{P}\{\mathcal{N}_2^i\}$ for all nonnegative integers $i$. We now look at (6). By the triangle inequality,

$$\left|\mathbb{P}\{\tau_A > kf\} - \mathbb{P}\{\mathcal{N}\}^k\right| \tag{10}$$

$$\le \sum_{j=0}^{k-2} \mathbb{P}\{\mathcal{N}\}^j \left|\mathbb{P}\{\tau_A > (k-j)f\} - \mathbb{P}\{\tau_A > (k-j-1)f \cap \mathcal{N}_2^{k-j-1}\}\right| \tag{11}$$

$$
\begin{aligned}
&+ \sum_{j=0}^{k-2} \mathbb{P}\{\mathcal{N}\}^j \big|\mathbb{P}\{\tau_A > (k-j-1)f \cap \mathcal{N}_2^{k-j-1}\}\\
&\qquad - \mathbb{P}\{\tau_A > (k-j-1)f\}\mathbb{P}\{\mathcal{N}_2^0\}\big|
\end{aligned} \tag{12}
$$

$$+ \mathbb{P}\{\mathcal{N}\}^{k-1}|\mathbb{P}\{\tau_A > f\} - \mathbb{P}\{\mathcal{N}\}|. \tag{13}$$

The idea of the proof is to iterate the proof of Lemma 10. In the rightmost term of (11) we introduced a gap (of length $2\Delta$). In the rightmost term of (13) we factored the measure on the left-hand side in the modulus. Term (13) is just a small technical correction.

The rightmost factor in term (13) is

$$\mathbb{P}\{\mathcal{N}\cap\tau_A\circ T^{f-2\Delta} \le 2\Delta\}.$$

We can bound the rightmost factor in the sum (11) by

$$\mathbb{P}\left\{\bigcap_{i=0}^{k-j-2} \mathcal{N}_1^i \cap \left(\tau_A\circ T^{(k-j-1)f} \le 2\Delta\right) \cap \mathcal{N}_2^{k-j-1}\right\}. \tag{14}$$

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

$$\mathbb{P}\left\{\bigcap_{i=0}^{k-j-2} \mathcal{N}_1^i\right\}\phi(2\Delta-n) \le \mathbb{P}\left\{\bigcap_{i=0}^{k-j-2} \mathcal{N}_1^i \cap \Omega\right\}\phi(\Delta). \tag{15}$$

Finally, we iterate the $\phi$-mixing property to get the bound

$$\begin{aligned}
\mathbb{P}\left\{\bigcap_{i=0}^{\ell} \mathcal{N}_1^i \cap B\right\} &\le \mathbb{P}\left\{\bigcap_{i=0}^{\ell-1} \mathcal{N}_1^i \cap B\right\} \\
&\le \prod_{i=1}^{\ell-1}[\mathbb{P}\{\mathcal{N}_1^i\}+\phi(\Delta)][\mathbb{P}\{B\}+\phi(\Delta)] \\
&= [\mathbb{P}\{\mathcal{N}_1^0\}+\phi(\Delta)]^{\ell-1}[\mathbb{P}\{B\}+\phi(\Delta)],
\end{aligned} \tag{16}$$

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

$$\begin{aligned}
\mathbb{P}\{\mathcal{N}\} &= \mathbb{P}\{\tau_A > f-2\Delta\} \\
&= \mathbb{P}\{\tau_A > f-\Delta\} + \mathbb{P}\{\tau_A \le \Delta \cap \tau_A \circ T^{\Delta} > f-2\Delta\} \\
&= \mathbb{P}\{\mathcal{N}_1^0\} + \phi(\Delta).
\end{aligned}$$

Thus, we apply the above equality to (16) and we obtain that

$$\begin{aligned}
\mathbb{P}\left\{\bigcap_{i=0}^{\ell} \mathcal{N}_1^i \cap B\right\} &\le \mathbb{P}\{\mathcal{N}\}^{\ell-1}[\mathbb{P}\{B\}+\phi(\Delta)] \\
&= \mathbb{P}\{\mathcal{N}\}^{\ell+1}[\mathbb{P}\{B\}+\phi(\Delta)]\frac{1}{\mathbb{P}\{\mathcal{N}\}^2} \\
&\le C\ \mathbb{P}\{\mathcal{N}\}^{\ell+1}[\mathbb{P}\{B\}+\phi(\Delta)],
\end{aligned} \tag{17}$$

where the last inequality follows since

$$\mathbb{P}\{\mathcal{N}\} \ge \mathbb{P}\{\tau_A > f\} \ge \mathbb{P}\left\{\tau_A > \frac{1}{2\mathbb{P}\{A\}}\right\} \ge 1/2.$$

On the other hand,

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

Therefore, we have the inequality

$$\mathbb{P}\left\{\bigcap_{i=0}^{\ell} \mathcal{N}_1^i \cap B\right\} \le \mathbb{P}\{\mathcal{N}\}^{\ell-1}\varepsilon(A). \tag{18}$$

Now, we apply (18) to (15) with $B = \Omega$, and to (14) with

$$B = \{\tau_A \circ T^{(k-j-1)f} \le 2\Delta \cap \mathcal{N}_2^{k-j-1}\}.$$

To obtain the rightmost factor on the right-hand side of (6) we just need to sum over-$k$. This concludes the proof of the upper bound (6).

Inequality (7) follows by the mean value theorem and (9).

$$|\mathbb{P}\{\tau_A > f - 2\Delta\}^k - \mathbb{P}\{\tau_A > f\}^k| \le k\varepsilon(A)\mathbb{P}\{\tau_A > f - 2\Delta\}^{k-1}.$$

This ends the proof of the proposition. □

3.4. *Proof of Theorem* 1. Let us fix $f = f_A = 1/(2\mathbb{P}\{A\})$ and $\Delta = \Delta_{A,f_A} = \Delta_A$ given by Proposition 11. We define

$$\xi_A = \frac{-\log \mathbb{P}\{\tau_A > f - 2\Delta\}}{f\mathbb{P}\{A\}}.$$

We note that

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

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

The structure of the proof of (1) has three steps. First, we prove for $t$ of the form $t = kf$ 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 $t = (k + p/q)f$ 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 $(k + p/q)f$.

PROOF [$t$'s of the form $t = (k + p/q)f$]. Let $t = (k + (p/q))f$, with $k, p, q$ as was just told. Put $r = (p/q)f$:

$$\begin{aligned}
&|\mathbb{P}\{\tau_A > t\} - e^{-\xi_A\mathbb{P}\{A\}t}| \\
&\quad = |\mathbb{P}\{\tau_A > kf + r\} - \mathbb{P}\{\tau_A > f - 2\Delta\}^{k + r/f}| \\
&\quad \le |\mathbb{P}\{\tau_A > kf + r\} - \mathbb{P}\{\tau_A > kf\}\mathbb{P}\{\tau_A > r\}| \\
&\qquad + |\mathbb{P}\{\tau_A > kf\} - \mathbb{P}\{\tau_A > f - 2\Delta\}^k|\mathbb{P}\{\tau_A > r\} \\
&\qquad + |\mathbb{P}\{\tau_A > r\} - \mathbb{P}\{\tau_A > f - 2\Delta\}^{r/f}|\mathbb{P}\{\tau_A > f - 2\Delta\}^k.
\end{aligned}$$

The first term on the right-hand side of the above inequality is bounded using Lemma 10 and (6). The last one can be used since, by Remark 2 on page 9, $f = 1/(2\mathbb{P}\{A\})$ satisfies condition (5).

$$\begin{aligned}
&|\mathbb{P}\{\tau_A > kf + r\} - \mathbb{P}\{\tau_A > kf\}\mathbb{P}\{\tau_A > r\}| \\
&\quad \le \varepsilon(A)\mathbb{P}\{\tau_A > kf - 2\Delta\} \\
&\quad \le \varepsilon(A)\mathbb{P}\{\tau_A > (k-1)f\} \\
&\quad \le \varepsilon(A)e^{-\xi_A \mathbb{P}\{A\}t}.
\end{aligned}$$

The last inequality follows by (17) with $B = \Omega$.

The second term is bounded by (6) with $f = 1/(2\mathbb{P}\{A\})$,

$$\begin{aligned}
&\left|\mathbb{P}\{\tau_A > kf\} - \mathbb{P}\{\tau_A > f - 2\Delta\}^k\right|\mathbb{P}\{\tau_A > r\} \\
&\quad \le \varepsilon(A)k\mathbb{P}\{\tau_A > f - 2\Delta\}^k\mathbb{P}\{\tau_A > r\} \\
&\quad \le \varepsilon(A)\mathbb{P}\{A\}te^{-\xi_A \mathbb{P}\{A\}t}.
\end{aligned}$$

The leftmost factor of the third term is bounded using the mean value theorem applied to the function $h(x) = x^{1/q}$,

$$\begin{aligned}
&\left|\mathbb{P}\{\tau_A > r\} - \mathbb{P}\{\tau_A > f - 2\Delta\}^{r/f}\right| \\
&\quad = \left|\mathbb{P}\left\{\tau_A > f\frac{p}{q}\right\}^q - \mathbb{P}\{\tau_A > f - 2\Delta\}^p\right|\frac{1}{q}\omega^{1/q-1} \\
&\quad \le \left|\mathbb{P}\left\{\tau_A > f\frac{p}{q}\right\}^q - \mathbb{P}\{\tau_A > f - 2\Delta\}^p\right| \qquad (20)\\
&\qquad \times \frac{1}{q}\min\left\{\mathbb{P}\left\{\tau_A > f\frac{p}{q}\right\}^q; \mathbb{P}\{\tau_A > f - 2\Delta\}^p\right\}^{1/q-1} \\
&\quad = \frac{|b-a|}{a}\,\frac{1}{q}a^{1/q},
\end{aligned}$$

where $b$ and $a$ are the maximum and minimum, respectively, of the difference (20) and $\omega \in (\min\{a,b\}, \max\{a,b\})$. We have that $a^{1/q} \le 1$, and

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

for every $c$, real positive. We choose $c = \mathbb{P}\{\tau_A > pf\}$.

Consider now two cases. (i) When $a$ is the rightmost term of the difference (20). (ii) When $a$ is the leftmost term of the difference (20).

*Proof of case* (i). By (6) with $k = p$ and $f = 1/(2\mathbb{P}\{A\})$,

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

Now we want to use (7) with $k = q$ and $f = p/(2\mathbb{P}\{A\}q)$. Notice that to do this, we just need to verify that there exists a $\Delta = \Delta_{p,q} \in [n, p/(2\mathbb{P}\{A\}q)]$ that verifies condition (8). We observe that the term in the right-hand side of (8) is actually a family function parametrized by $f_A$. Moreover, they are decreasing as a function of $f$. Then, such $\Delta_{p,q}$ exists and $\Delta_m < \Delta_{1,q} < \Delta_{2,q} < \cdots < \Delta_{q,q} = \Delta_{1/(2\mathbb{P}\{A\})}$. Furthermore, $\phi(\Delta_{p,q}) \le \varepsilon(A)$. This inequality follows in the same way as was proved that $\phi(\Delta_f) \le \varepsilon(A)$ at the beginning of the proof of Proposition 11, where $\Delta_m$ was defined.

So, we use (7) twice. Finally, we use (6) with $k = p$ and with $f = 1/(2\mathbb{P}\{A\})$:

$$\begin{aligned}\frac{|b-c|}{a}\frac{1}{q} &\le \frac{\varepsilon(A)qb}{a}\frac{1}{q} \le \frac{\varepsilon(A)}{a}\frac{c}{1-\varepsilon(A)q}\\ &\le \frac{\varepsilon(A)}{a}\frac{1+\varepsilon(A)p}{1-\varepsilon(A)q}a \le \varepsilon(A).\end{aligned}$$

Case (ii) is treated in the same way. This proves that the third term is bounded from above by

$$C\varepsilon(A)\mathbb{P}\{\tau_A > f - 2\Delta\}^k \le C\varepsilon(A)e^{-\xi_A\mathbb{P}\{A\}t}.$$

*Proof* (A general $t$). Now, let $t$ be any positive real. We write $t = kf + r$, with $k$ a positive integer and $r$ such that $0 \le r < f$. We can choose a $\bar{t}$ such that $\bar{t} < t$ and $\bar{t} = (k + (p/q))f_A$ with $p, q$ as before.

$$\begin{aligned}\left|\mathbb{P}\{\tau_A > t\} - e^{-\xi_A\mathbb{P}\{A\}t}\right| &\le |\mathbb{P}\{\tau_A > t\} - \mathbb{P}\{\tau_A > \bar{t}\}|\\ &\quad + \left|\mathbb{P}\{\tau_A > \bar{t}\} - e^{-\xi_A\mathbb{P}\{A\}\bar{t}}\right|\\ &\quad + \left|e^{-\xi_A\mathbb{P}\{A\}\bar{t}} - e^{-\xi_A\mathbb{P}\{A\}t}\right|.\end{aligned}$$

The first term on the right-hand side of the above inequality is bounded applying (18),

$$\begin{aligned}|\mathbb{P}\{\tau_A > t\} - \mathbb{P}\{\tau_A > \bar{t}\}| &= \mathbb{P}\{\tau_A > \bar{t} \cap \tau_A \circ T^{\bar{t}} \le t - \bar{t}\}\\ &\le \mathbb{P}\{\tau_A > kf \cap \tau_A \circ T^{\bar{t}} \le \Delta\}\\ &\le C\varepsilon(A)e^{-\xi_A\mathbb{P}\{A\}t}.\end{aligned}$$

The upper bound for the third term is obtained using the mean value theorem,

$$\begin{aligned}\left|e^{-\xi_A\mathbb{P}\{A\}\bar{t}} - e^{-\xi_A\mathbb{P}\{A\}t}\right| &\le \xi_A\mathbb{P}\{A\}\left(r - \frac{p}{q}f\right)e^{-\xi_A\mathbb{P}\{A\}\bar{t}}\\ &\le \Lambda_2\,\varepsilon(A)\,e^{-\xi_A\mathbb{P}\{A\}t}.\end{aligned}$$

Finally, the second term is bounded as in the first part of the proof. This, together with (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 $\mathbb{P}$ the equilibrium measure of the chain. Then $\mathbb{P}\{A\} = q/(p+q)$. When $q \ll p$, we have $\mathbb{P}\{A\} \ll 1$. That is, $A$ is rare. Then it is fair to ask about the difference between its hitting time law and the exponential law. We fix $t = 1/\mathbb{P}\{A\}$. So, the proof of Theorem 1 for this time $t$ reduces to the proof of Proposition 11 with $k = 2$ and $f_A = 1/(2\mathbb{P}\{A\})$. We notice that the difference inside the modulus in (11) and (13) always has negative sign. Furthermore, a straightforward computation using the Markovian property shows that both terms are *equal*, respectively, to $C_1\Delta\mathbb{P}\{A\}$ and $C_2\Delta\mathbb{P}\{A\}$ for positive constant $C_1, C_2$.

We must see that the difference inside the modulus in (12) is negative too. So, in this case, the triangle inequality that bounds (10) is actually an equality. This is equivalent to

$$\mathbb{P}\{\mathcal{N}_2^{k-j-1} | \tau_A > (k-j-1) f_A\} \le \mathbb{P}\{\mathcal{N}_2^0\}.$$

Using stationarity and the definition of $A$, this is equivalent to

$$\mathbb{P}\{X_{2\Delta+1}^{f_A} = 1 | X_{-(k-j-1)f_A}^0 = 1\} \le \mathbb{P}\{X_{2\Delta+1}^{f_A} = 1\}.$$

But using the Markovian property this is equivalent to

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

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

**5. Application: estimate of $\mathbb{E}(\tau_A)$.** An important result in ergodic theory is the famous Kac's lemma that 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. There is no equivalent result for the case of the expected hitting time.

We present in the next corollary an estimate for the expected hitting time as a consequence of the nonuniform bound given in Theorem 1. Roughly speaking, if the hitting time law converges to an exponential law with parameter $\xi_A$, 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*:

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

*or equivalently*,

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

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

PROOF. Using the mean value theorem,

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

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

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

where we denote with $\exp\{\nu\}$ an exponential random variable with parameter $\nu$. This ends the proof of the first inequality. The second one is an immediate consequence of the previous one together with Kac's lemma and the fact that $\xi_A$ is bounded from below. $\square$

The following theorem proves a computable approximation for $\xi_A$. It also shows that this parameter actually is no larger than 1.

Let us fix $s$, a positive integer. We denote with $\mathbb{P}_A$ the conditional probability measure on the event $A$. We define

$$\zeta_{A,s} := \mathbb{P}_A\left\{\tau_A > \frac{n}{s}\right\} = \mathbb{P}\left\{\left\{\tau_A > \frac{n}{s}\right\}\middle| A\right\}.$$

THEOREM 13 [Abadi (2001)]. *Let* $\{X_m\}_{m\in\mathbb{Z}}$ *be exponentially* $\phi$*-mixing. Let* $s$ *be a positive integer. Then, there exist strictly positive constants* $\Psi_1, \Psi_2, C_1$ *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*:

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

REMARK 1. In Galves and Schmitt (1997) it was shown that the limit $\mathbb{E}(\tau_{A_n})/b_n$ exists, where $\tau_{A_n}$ is the first time an exponentially $\psi$-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 also provides a velocity of convergence for that limit.

REMARK 2. Theorem 1 says that when $\xi_A$ is used as a re-scaling factor and the process is exponentially $\phi$-mixing, the error in the approximation of the hitting time law is exponential. However, this quantity depends upon a large part of the process, say $X_1, \dots, X_{f_A - 2\Delta}$ (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 $\xi_A$ as a 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 $1, \dots, n + n/s$ and on the combinatorial properties of $A$. For instance, $\zeta_{A,s} = 1$ for every nonrecurrent 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 as an example of a process and a cylinder $A$ such that $\xi_A < 1$. Moreover, a process as a cylinder $A$ can be constructed such that $\xi_A < \delta < 1$, for all $0 < \delta < 1$. In Asselah and Dai Pra (1997a, 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 $\tau_A$ 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 are not necessary.** The proof of Lemma 9 shows that the uniform lower bound $\Lambda_1$ for the parameter $\lambda_{A,f}$, depends on 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 *sufficient* to prove the uniform lower bound. We prove in the next proposition a weaker 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 stationary stochastic process. Let $t$ be a positive integer. Denote $N = N_t$. Then,*

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

*where we adopt the convention* $0 \log 0 = 0$.

PROOF. By Hölder's inequality with $p$ and $q$ conjugate, we have

$$\mathbb{E}N = \mathbb{E}(N \mathbb{1}_{\{N \ge 1\}}) \le \mathbb{E}(N^p)^{1/p} \mathbb{P}\{N \ge 1\}^{1/q}.$$

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

$$\begin{aligned}\mathbb{P}\{\tau_A \le t\} &= \mathbb{P}\{N \ge 1\} \\ &\ge \lim_{p\to 1} \frac{1}{(\mathbb{E}(N/(t\mathbb{P}\{A\}))^p)^{1/(p-1)}} = \frac{1}{\exp\{\mathbb{E}(N/(t\mathbb{P}\{A\})\log N/(t\mathbb{P}\{A\}))\}}.\end{aligned}$$

This ends the proof of the proposition. □

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

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

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

$$C\mathbb{P}\left\{\tau_A \le \frac{1}{2\mathbb{P}\{A\}}\right\} \le \frac{1-e^{-\xi_A}}{2} \le \xi_A \le 1 - e^{-\xi_A} \le \mathbb{P}\left\{\tau_A \le \frac{1}{2\mathbb{P}\{A\}}\right\}.$$

Then $\xi_A$ remains bounded if and only if $\mathbb{P}\{\tau_A \le 1/(2\mathbb{P}\{A\})\}$ does. Lemma 7 and the above proposition end the proof of the corollary. □

**7. Uniformly mixing processes.** When the function $\alpha$ 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 $s \in \mathbb{N}$. For $1 \le j \le n/s$, define $\mathcal{B}_{n,j}$ as the set of $A \in \mathcal{C}_n$ which recur exactly at time $j$. Namely, $A \in \mathcal{B}_{n,j}$, if $A \cap T^{-j}A \ne \varnothing$ and $A \cap T^{-i}A = \varnothing$, for all $1 \le i < j$. Define also $\mathcal{B}_n(s) = \bigcup_{1\le j\le n/s} \mathcal{B}_{n,j}$.

For each positive integer $s$, define

$$\mathcal{F}_s = \left\{A \in \mathcal{C}_n \,\middle|\, \sum_{j=n/(2s)}^{\infty} \alpha(j) \le \mathbb{P}\{A\}\right\}. \tag{22}$$

THEOREM 17. *Let $\{X_m\}_{m\in\mathbb{Z}}$ be an $\alpha$-mixing process with function $\alpha$. Assume that $\alpha(j) \le 1/j^4$, for all $j \in \mathbb{N}$. There exist strictly positive constants $\Lambda_1$, $\Lambda_2$ and $C$, such that for any $n \in \mathbb{N}$ and any $A \in (\mathcal{E}^n \setminus \mathcal{B}_n(s)) \cap \mathcal{F}_s$, for which $n^2\sqrt{\mathbb{P}\{A\}} \le 1$, there exists a $\lambda_A \in [\Lambda_1, \Lambda_2]$, such that the following inequality holds:*

$$\sup_{t\ge 0}\left|\mathbb{P}\left\{\tau_A > \frac{t}{\lambda_A\mathbb{P}\{A\}}\right\} - e^{-t}\right| \le Cn\sqrt{\mathbb{P}\{A\}}. \tag{23}$$

One important difference between the exponential law for $\alpha$-mixing processes at one hand and $\phi$-mixing and $\psi$-mixing processes at the other is that in this case, we do not get upper bounds in function of $t$, as in the $\phi$-mixing and $\psi$-mixing cases, but instead uniform upper bounds.

Another important difference between them is that in the latter cases 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 $A \in (\mathcal{E}^n \setminus \mathcal{B}_n(s)) \cap \mathcal{F}_s$.

Loosely speaking, if $h > 0$ is the entropy of process, one has typically $\mathbb{P}\{A\} \approx e^{-hn}$. Then $\sqrt{\mathbb{P}\{A\}}n^2 < 1$. Moreover, if $A \in \mathcal{F}_s$, one has that $\alpha$ actually decays exponentially fast. Hence, $\alpha(j) \leq 1/j^4$. Therefore, the two constraints of the theorem are not restrictive. Moreover, if $\alpha$ decays exponentially fast with constant $c$, then $\sum_{j \geq n/s} \alpha(j) \approx Ce^{-cn/s}$. So, (22) means $c/s > h$.

On the other hand,

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

Thus, we have

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

So, if $h_M/s < h$, then $\mathbb{P}\{\bigcup_{B \in \mathcal{B}_n(s)} B\} \to 0$ as $n \to \infty$, where $h_M = \log|\mathcal{E}|$ 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 $h_M < h\ s < c$.

We prove the uniform subexponential decay of the cylinders.

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

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

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

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

Using the mixing property,

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

where $\rho := \sup\{\mathbb{P}\{a_i\} : a_i \in \mathcal{E}\} < 1$ since the process is ergodic. 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. $\square$

For any cylinder set $A$ define

$$\lambda_A = \frac{-\log \mathbb{P}\{\tau_A > 1/\sqrt{\mathbb{P}\{A\}}\}}{\sqrt{\mathbb{P}\{A\}}}.$$

LEMMA 19. *Under the conditions of Theorem* 17, *there exist two constants* $\Lambda_1, \Lambda_2$, *with* $0 < \Lambda_1 \le 1 \le \Lambda_2 < \infty$, *and a positive integer* $n_0$, *such that, for any* $n \ge n_0$ *and any* $A \in (\mathcal{E}^n \setminus \mathcal{B}_n(s)) \cap \mathcal{F}_s$, *for which* $n^2\sqrt{\mathbb{P}\{A\}} \le 1$, *the following inequalities hold*:

$$\Lambda_1 \le \lambda_A \le \Lambda_2.$$

We remark that $\Lambda_1$ and $\Lambda_2$ are independent of $n$ and $A$.

PROOF OF LEMMA 19. By Taylor's expansion,

$$1 - e^{-x} \le x \le 1 - e^{-x} + 2(1 - e^{-x})^2, \tag{24}$$

for $0 \le x \le \log 3$. Applying it to $x = -\log \mathbb{P}\{\tau_A > 1/\sqrt{\mathbb{P}\{A\}}\}$, together with Lemma 7, we have that

$$\lambda_A \le 1 + 2\sqrt{\mathbb{P}\{A\}},$$

and we can take $\Lambda_2 = 3$.

In what follows let us write $N$ for $N_{1/\sqrt{\mathbb{P}\{A\}}}$. To obtain the lower bound for $\lambda_A$ we first compute

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

By definition the first term in the above decomposition is bounded from above by $\sqrt{\mathbb{P}\{A\}}$. The second term, by hypothesis, is zero. The remaining terms are bounded using the mixing property. Denote by $A^{(k)}$ the cylinder defined by the last $k$ ($k \le n$) letters of $A$. The third term is bounded by

$$2 \sum_{\ell=n/s+1}^{2n} \left( \left[ \frac{1}{\sqrt{\mathbb{P}\{A\}}} \right] - \ell \right) \left[ \mathbb{P}\{A\}\mathbb{P}\{A^{(n/2s)}\} + \alpha\left(\ell - \frac{n}{2s}\right) \right]$$

$$\le 4n\sqrt{\mathbb{P}\{A\}}\mathbb{P}\{A^{(n/2s)}\} + \frac{2}{\sqrt{\mathbb{P}\{A\}}} \sum_{j=n/2s}^{2n-n/(2s)} \alpha(j).$$

By Lemma 18 and the polinomial decay of $\alpha$, one has that for $n$ large enough, $n\mathbb{P}(A^{(n/2s)}) < 1$.

Finally, the last term is bounded by

$$2 \sum_{\ell=2n+1}^{[1/\sqrt{\mathbb{P}\{A\}}]} \left( \left[ \frac{1}{\sqrt{\mathbb{P}\{A\}}} \right] - \ell \right) [\mathbb{P}\{A\}^2 + \alpha(\ell - n)]$$

$$\le 2\mathbb{P}\{A\} + \frac{2}{\sqrt{\mathbb{P}\{A\}}} \sum_{j=n+1}^{[1/\sqrt{\mathbb{P}\{A\}}]} \alpha(j).$$

With the above inequalities we have

$$\mathbb{E}(N^2) \le 7\sqrt{\mathbb{P}\{A\}} + \frac{4}{\sqrt{\mathbb{P}\{A\}}} \sum_{j=n/(2s)}^{[1/\sqrt{\mathbb{P}\{A\}}]} \alpha(j) \le 11\sqrt{\mathbb{P}\{A\}}.$$

The last inequality follows by condition (22). Therefore, we have that

$$\lambda_A \ge \frac{\mathbb{P}\{\tau_A \le 1/\sqrt{\mathbb{P}\{A\}}\}}{\sqrt{\mathbb{P}\{A\}}} \ge \frac{(\mathbb{E}(N))^2}{\sqrt{\mathbb{P}\{A\}}\,\mathbb{E}(N^2)} \ge \frac{1}{11}. \tag{25}$$

The first inequality follows by (24). The second one follows by (3). This ends the proof of the lemma. □

PROOF OF THEOREM 17. Without loss of generality we can assume that $n$ is large. For proving inequality (23), we follow the triangle inequality (10), emphasizing that it holds for $\Delta = n$ and $f = 1/\sqrt{\mathbb{P}\{A\}}$ without assuming (5). In such case $\mathbb{P}\{\mathcal{N}\} = \mathbb{P}\{\tau_A > 1/\sqrt{\mathbb{P}\{A\}} - 2n\}$. By stationarity and Lemma 7,

$$\mathbb{P}\{\mathcal{N}\} - \mathbb{P}\{\tau_A > 1/\sqrt{\mathbb{P}\{A\}}\} \le \mathbb{P}\{\tau_A \le 2n\} \le 2n\mathbb{P}\{A\}.$$

By (25) we have

$$1 - \mathbb{P}\{\mathcal{N}\} \ge \tfrac{1}{11}\sqrt{\mathbb{P}\{A\}} - 2n\mathbb{P}\{A\}.$$

For $n$ large enough, $0 < \sqrt{\mathbb{P}\{A\}}/11 - 2n\mathbb{P}\{A\} \le 1/2$. So, we bound the sum (11) by

$$\sum_{j=0}^{\infty} \mathbb{P}\{\mathcal{N}\}^j\, n\mathbb{P}\{A\} = \frac{n\mathbb{P}\{A\}}{1 - \mathbb{P}\{\mathcal{N}\}} \le 2\, n\sqrt{\mathbb{P}\{A\}}. \tag{26}$$

Similarly, we bound the sum (12) by

$$\sum_{j=0}^{\infty} \mathbb{P}\{\mathcal{N}\}^j \alpha(n) \le \frac{\alpha(n)}{1-\mathbb{P}\{\mathcal{N}\}} \le 2\sqrt{\mathbb{P}\{A\}},$$

where the last inequality follows by condition (22). Finally, we bound term (13) by $n\mathbb{P}\{A\}$.

Write $t = k\frac{1}{\sqrt{\mathbb{P}\{A\}}} + r$, with $k$ and $r$ integers, $0 \le r < \frac{1}{\sqrt{\mathbb{P}\{A\}}}$. We have

$$\begin{aligned}
&\left|\mathbb{P}\{\tau_A > t\} - e^{-\lambda_A \mathbb{P}\{A\} t}\right| \\
&\quad \le \left|\mathbb{P}\left\{\tau_A > k\frac{1}{\sqrt{\mathbb{P}\{A\}}} + r\right\} - \mathbb{P}\left\{\tau_A > k\frac{1}{\sqrt{\mathbb{P}\{A\}}}\right\}\right| \\
&\qquad + \left|\mathbb{P}\left\{\tau_A > k\frac{1}{\sqrt{\mathbb{P}\{A\}}}\right\} - \mathbb{P}\{\mathcal{N}\}^k\right| \\
&\qquad + \left|\mathbb{P}\{\mathcal{N}\}^k - \mathbb{P}\left\{\tau_A > \frac{1}{\sqrt{\mathbb{P}\{A\}}}\right\}^k\right| \\
&\qquad + \left|e^{-\lambda_A \mathbb{P}\{A\} k/\sqrt{\mathbb{P}\{A\}}} - e^{-\lambda_A \mathbb{P}\{A\} t}\right|.
\end{aligned}$$

The first term is bounded using stationarity and (2),

$$\left|\mathbb{P}\left\{\tau_A > k\frac{1}{\sqrt{\mathbb{P}\{A\}}} + r\right\} - \mathbb{P}\left\{\tau_A > k\frac{1}{\sqrt{\mathbb{P}\{A\}}}\right\}\right| \le \sqrt{\mathbb{P}\{A\}}. \tag{27}$$

Bounds for the third and fourth terms are obtained using the mean value theorem,

$$\left|\mathbb{P}\{\tau_A > \mathcal{N}\}^k - \mathbb{P}\left\{\tau_A > \frac{1}{\sqrt{\mathbb{P}\{A\}}}\right\}^k\right| \le k2n\mathbb{P}\{A\}\mathbb{P}\{\tau_A > \mathcal{N}\}^{k-1} \le 2n\mathbb{P}\{A\},$$

$$\left|e^{-\lambda_A \mathbb{P}\{A\} k/\sqrt{\mathbb{P}\{A\}}} - e^{-\lambda_A \mathbb{P}\{A\} t}\right| \le \lambda_A \mathbb{P}\{A\} r e^{-\lambda_A \mathbb{P}\{A\} k/\sqrt{\mathbb{P}\{A\}}} \le \Lambda_2 \sqrt{\mathbb{P}\{A\}}.$$

The sum of the above bounds proves (23). Lemma 19 ends the proof of the theorem. □

**Acknowledgments.** Part of 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.

The author kindly thanks two anonymous referees for corrections and useful suggestions to a previous version of this paper.

## REFERENCES

ABADI, M. (2001). Exponential approximation for hitting times in mixing processes. *Math. Phys. Electron. J.* **7** No. 2.

ABADI, M. and GALVES, A. (2001). Inequalities for the occurrence times of rare events in mixing processes. The state of the art. *Markov Process. Related Fields* **7** 97–112.

ASSELAH, A. and DAI PRA, P. (1997a). Sharp estimates for the occurrence time of rare events for simple symmetric exclusion. *Stochastic Process. Appl.* **71** 259–273.

ASSELAH, A. and DAI PRA, P. (1997b). Occurrence of rare events in ergodic interacting spin systems. *Ann. Inst. H. Poincaré Probab. Statist.* **33** 727–751.

BELLMAN, R. and HARRIS, T. E. (1951). Recurrence times for the Ehrenfest model. *Pacific J. Math.* **1** 179–193.

BOWEN, R. (1975). *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*. *Lecture Notes in Math.* **470**. Springer, New York.

BRESSAUD, X., FERNANDEZ, R. and GALVES, A. (1999). Speed of $\overline{d}$-convergence for Markov approximations of chains with complete connections. A coupling approach. *Stochastic Process. Appl.* **83** 127–138.

CHEN, L. (1975). Poisson approximation for dependent trials. *Ann. Probab.* **3** 534–345.

COLLET, P., GALVES, A. and SCHMITT, B. (1999). Repetition times for Gibbsian sources. *Nonlinearity* **12** 1225–1237.

CORNFELD, I., FOMIN, S. and SINAI, Y. (1982). *Ergodic Theory*. Springer, New York.

DOEBLIN, W. (1940). Remarques sur la théorie métrique des fractions continues. *Compositio Math.* **7** 353–371.

DOUKHAN, P. (1995). *Mixing*. *Properties and Examples*. *Lecture Notes in Statist.* **85**. Springer, New York.

FERRARI, P. A. and GALVES, J. A. (1997). *Acoplamento e Processos Estocásticos.* IMPA, Rio de Jeneiro.

GALVES, A. and SCHMITT, B. (1990). Occurrence time of rare events for mixing dynamical systems. *Ann. Inst. H. Poincaré* **52** 267–281.

GALVES, A. and SCHMITT, B. (1997). Inequalities for hitting times in mixing dynamical systems. *Random Comput. Dyn.* **5** 337–348.

HARRIS, T. E. (1953). First passage and recurrence distributions. *Trans. Amer. Math. Soc.* **73** 471–486.

HAYDN, N. (1999). The distribution of the first return time for rational maps. *J. Statist. Phys.* **94** 1027–1036.

HIRATA, M., SAUSSOL, B. and VAIENTI, S. (1999). Statistics of return times: A general framework and new applications. *Comm. Math. Phys.* **206** 33–55.

KAC, M. (1947). On the notion of recurrence in discrete stochastic processes. *Bull. Amer. Math. Soc.* **53** 1002–1010.

SAUSSOL, B. (2001). On fluctuations and the exponential statistics of return times. *Nonlinearity* **14** 179–191.

SHIELDS, P. (1996). *The Ergodic Theory of Discrete Sample Paths*. **13**. Amer. Math. Soc., Providence, RI.

DEPARTAMENTO DE ESTATÍSTICA
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
UNIVERSIDADE DE SÃO PAULO
RUA DO MATÃO 1010
CIDADE UNIVERSITARIA
CEP 05508-090
SÃO PAULO SP
BRASIL
E-MAIL: abadi@ime.usp.br