Sharp error terms for return time statistics under mixing conditions *

Miguel Abadi[†] Nicolas Vergne[‡]

Abstract

We describe the statistics of repetition times of a string of symbols in a stochastic process. We consider a string A of length n and prove: 1) The time elapsed until the process starting with A repeats A, denoted by τ_A, has a distribution which can be well approximated by a degenerated law at the origin and an exponential law. 2) The number of consecutive repetitions of A, denoted by S_A, has a distribution which is approximately a geometric law. We provide sharp error terms for each of these approximations. The errors we obtain are point-wise and allow to get also approximations for all the moments of τ_A and S_A. Our results hold for processes that verify the ϕ mixing condition.

keyword: Mixing, recurrence, rare event, return time.

1 Introduction

This paper describes the statistics of return times of a string of symbols in a mixing stochastic process with a finite alphabet. Generaly speaking, the study of the time elapsed until the first occurrence of an event with small probability in dependent processes has a long history which can be traced out in [11]. Recently and exhaustive analysis of this statistics was motivated for applications in different areas as entropy estimation, genomic analysis, computer science, linguistic, among others. The typical result is:

$$\lim_{n\to\infty} I\!\!P\left(\tau_{A_n} > t\, b_n \mid \mu_0\right) = e^{-t} \ , \tag{1}$$

where τ_{A_n} is the first time the process hits a given measurable set A_n, $n \in I\!\!N$ and such that the measure $I\!\!P\left(A_n\right)$ go to zero as $n \to \infty$, $\{b_n\}_{n\in I\!\!N}$ is a suitable re-scaling sequence of positive numbers and μ_0 is a given initial condition. From

*Work done as part of the activities of the project Projeto Temtico-PRONEX *"Stochastic behavior, critical phenomena and rhythmic patterns identification in natural languages"*. Pros. FAPESP 03/09930-9.

[†]IMECC, Universidade de Campinas, Pa Srgio Buarque de Holanda 651 Cid. Univ. CP 6065, 13083-859, Campinas SP, Brazil. miguel@ime.unicamp.br

[‡]Université d'Evry Val d'Essonne, Département Mathématiques, Laboratoire Statistique et Génome, 91 000 Evry, France. vergne@genopole.cnrs.fr

the point of view of applications, a fundamental task is to know the rate of convergence of the above limit. A detailed review of such results appearing in the litterature can be find in [3].

Is the purposse of this paper to present, for *any* n-string A:

- A new sharp upper bound for the above rate of convergence in general ϕ-mixing processes that holds when $\mu_0 = A$.
- A sharp upper bound between th e law of the number of *consecutive* visits to to A and a geometric law.

When μ_0 is taken as A, we refer to the distribution in (1) as the *return time*. In general it can not be well approximated by an exponential law. This was firstly noted in [13], where it is proven the convergence of the number of visits to a small cylinder around a point to the Poisson law for axiom A diffeomorphisms. The result holds for *almost* every point. Then, it is proven that for periodic points, the asymptotic limit law of the return time differs of the one-level Poisson law, namely e^{-t}.

When in equation (1) the initial condition is the ergodic measure of the process, τ_A is called the *hitting time* of A. In [14] it is proven a rate of convergence of the return time as function of the distance between the hitting time and return time laws but it does not converge for cylinders around periodic points.

Our first result concerns the rate of convergence of limit in (1) when $\mu_0 = A$ for any n-string A. We prove that $I\!P(\tau_A > t/I\!P(A) \mid A)$ converges to a mixture of Dirac law at the origin and an exponential law. Namely, for large n

$$I\!P\left(\tau_A > \frac{t}{I\!P(A)} \mid A\right) \approx (1-\zeta)\delta + \zeta e^{-\zeta t} \ . \tag{2}$$

δ is the Dirac measure at the origin. ζ is a parameter related to the self-repeating properties of the string A. It worth noting that the parameter of the exponential law is exactly the weight of the convex combination. So far, the self-repeating properties of a string appears as a major factor to describe the statistical properties of the return time. For instance, if a string admits to overlaps itself, then it will turn out in the sequel that $\zeta \neq 1$ and the return time distribution approximates the above mixture of laws. However, for a word which does not overlap itself, it will turn out that $\zeta = 1$ and the return time distribution approximates a purely exponential law. We explore the notion and properties of overlapping in section 3 of this paper. It worth noting that contrary to aforementioned works, our result applies to *all* strings.

Our error estimate decays exponentially fast in t for all $t > 0$. As a byproduct we obtain *explicit expressions* for *all the moments* of the return time. This also appears as a generalization of the famous Kac's lemma (see [15]) which states that the *first* moment of the return time to a n-string A of positive measure is equal to $I\!P(A)^{-1}$ and the result in [6] which only presents conditions for the *existence* of the moments of return times. Further, [14] proves that hitting and return times coincide if and only if the return time converges to the exponential law. We extend this result establishing that the laws of hitting and return time coincide if and only if the Dirac measure of the return time law is absent which is equivalent to consider a non-self-repeating string.

Our framework is the class of ϕ-mixing processes. For instance, irreducible and aperiodic finite state Markov chains are known to be ψ-mixing with exponential decay. Moreover, Gibbs states which have summable variations are ψ-mixing (see [17]). They have exponential decay if they have Hölder continuous potential (see [5]). We refer the reader to [10] for a source of examples of mixing processes. However, sometimes the ψ-mixing condition is very restricted hypothesis difficult to test. We establish our result under the more general ϕ-mixing condition. The error term is explicitly expressed as a function of the mixing rate ϕ.

The base of our proof is a sharp approximation on the rate of convergence of the hitting time to an exponential law proven in [2].

The self-repeating phenomena in the distribution of the return time leads us to consider the problem of the sojourn time. Our second result states that the law of the number of consecutive repetitions of the string A, denoted by S_A, converges to a geometric law. Namely

$$I\!P\left(S_A = k \mid A\right) \approx (1-\rho)\rho^k \ .$$

ρ is the probability that the string repeats itself. The parameter of this law does *not* depend on the length nor on the measure of the string but just upon its self-repeating properties. Furthermore we show that under suitable conditions one has $\rho \approx 1 - \zeta$. As far as we know, this is the first result on this subject for dependent processes.

As in our previous result, the error bound we obtain decreases geometrically fast with k=the number of consecutive visits to the string. This decay on the error bound allows us to obtain an approximation for *all the moments* of S_A for those of a geometrically distributed random variable.

Our results are applied in a forthcoming paper: In [4] the authors prove large deviations and fluctuations properties of the repetition time function introduced by Wyner and Ziv in [18] and further by Ornstein and Weiss in [16], and get entropy estimators.

This paper is organized as follows. In section 2 we establish our framework. In section 3 we describe the self-repeating properties needed to state the return time result. In section 4 we establish the approximation for the return time law. This is Theorem 2. Finally, in section 5 we state and prove the geometric approximation for the consecutive repetitions of a string. This is Theorem 24.

2 Framework and notations

Let $\mathcal{E}$ be a finite set. Put $\Omega = \mathcal{E}^{Z\!\!Z}$. For each $x = (x_m)_{m \in Z\!\!Z} \in \Omega$ and $m \in Z\!\!Z$, let $X_m : \Omega \to \mathcal{E}$ be the m-th coordinate projection, that is $X_m(x) = x_m$. We denote by $T : \Omega \to \Omega$ the one-step-left shift operator, namely $(T(x))_m = x_{m+1}$.

We denote by $\mathcal{F}$ the σ-algebra over Ω generated by strings. Moreover we denote by $\mathcal{F}_I$ the σ-algebra generated by strings 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 $a_i \in \mathcal{E},\ i = 0, \ldots , n-1$.

We consider an invariant probability measure $I\!\!P$ over $\mathcal{F}$. We shall assume without loss of generality that there is no singleton of probability 0.

We say that the process $\{X_m\}_{m\in Z}$ is ϕ-mixing if the sequence

$$\phi(l) = \sup |I\!\!P_B(C) - I\!\!P(C)| \ ,$$

converges to zero. The supremum is taken over B and C such that $B \in \mathcal{F}_{\{0,.,n\}}, n \in I\!\!N, I\!\!P(B) > 0, C \in \mathcal{F}_{\{m\geq n+l+1\}}$.

For two measurables V and W, we denote as usual $I\!\!P(V|W) = I\!\!P_W(V) = I\!\!P(V; W) / I\!\!P(W)$ the conditional measure of V given W. We write $I\!\!P(V; W) = I\!\!P(V \cap W)$. We also write $V^c = \Omega \backslash V$, for the complement of V.

We use the probabilistic notation: $\{X_n^m = x_n^m\} = \{X_n = x_n, \ldots , X_m = x_m\}$. For a n-string $A = \{X_0^{n-1} = x_0^{n-1}\}$ and $1 \leq w \leq n$, we write $A^{(w)} = \{X_{n-w}^{n-1} = x_{n-w}^{n-1}\}$ for the w-string belonging to the σ-algebra $\mathcal{F}_{\{n-w,\ldots,n-1\}}$ and consisting of the *last* w symbols of A.

The conditional mean of a r.v. X with respect to any measurable V will be denoted by $I\!\!E_V(X)$ and we put $I\!\!E(X)$ when $V = \Omega$. Wherever it is not ambiguous we will write C and c for different positive constants even in the same sequence of equalities/inequalities. For brevity we put $(a \vee b) = \max\{a, b\}$ and $(a \wedge b) = \min\{a, b\}$.

3 Periodicity

Definition 1 *Let $A \in \mathcal{C}_n$. We define the* periodicity *of A (with respect to T) as the number $\tau(A)$ defined as follows:*

$$\tau(A) = \min\left\{k \in \{1, \ldots , n\} \mid A \cap\ T^{-k}(A) \neq \emptyset\right\} \ .$$

Let us take $A \in \mathcal{C}_n$, and write $n = q\,\tau(A) + r$, with $q = [n/\tau(A)]$ and $0 \leq r < \tau(A)$. Thus

$$A = \left\{X_0^{\tau(A)-1} = X_{\tau(A)}^{2\tau(A)-1} = \ldots = X_{(q-1)\tau(A)}^{q\tau(A)-1} = a_0^{\tau(A)-1}\ ;\ X_{q\tau(A)}^{n-1} = a_0^{r-1}\right\} \ .$$

So, we say that A has *period* $\tau(A)$ and *rest* r. We remark that periods can be "read backward" (and for the purpose of section 5 it will be more useful to do it in this way), that is

$$\begin{aligned} A &= \left\{X_0^{r-1} = a_{n-r}^{n-1}\ ;\ X_{n-q\tau(A)}^{n-(q-1)\tau(A)-1} = \ldots = X_{n-2\tau(A)}^{n-\tau(A)-1} = X_{n-\tau(A)}^{n-1} = a_{n-\tau(A)}^{n-1}\right\} \\ &= \left\{T^{q\tau(A)}A^{(r)}; T^{(q-1)\tau(A)}A^{(\tau(A))}; \ldots ; T^{2\tau(A)}A^{(\tau(A))}; T^{\tau(A)}A^{(\tau(A))}; A^{(\tau(A))}\right\} \ . \end{aligned}$$

We recall the definition of $A^{(w)}, 1 \leq w \leq n$, at the end of section 2. For instance

$$A = (\overbrace{\texttt{aaaabb}}^{\text{period}}\overbrace{\texttt{aaaabb}}^{\text{period}}\overbrace{\texttt{aaa}}^{\text{rest}}) = (\underbrace{\overbrace{\texttt{aaa}}^{\text{rest}}}_{T^{12}A^{(3)}}\underbrace{\overbrace{\texttt{abbaaa}}^{\text{period}}}_{T^{6}A^{(6)}}\underbrace{\overbrace{\texttt{abbaaa}}^{\text{period}}}_{A^{(6)}}) \ . \tag{3}$$

In the middle of the above equality, periods are read forward while in the right hand side periods are read backward.

Consider the set of overlapping positions of A:

$$\left\{k \in \{1, \ldots, n-1\} \mid A \cap \ T^{-k}(A) \neq \emptyset\right\} = \{\tau(A), \ldots, [n/\tau(A)]\tau(A)\} \cup \mathcal{R}(A) \ ,$$

where

$$\mathcal{R}(A) = \left\{k \in \{[n/\tau(A)]\tau(A)+1, \ldots, n-1\} \mid A \cap \ T^{-k}(A) \neq \emptyset\right\} \ .$$

The set $\{\tau(A), \ldots, [n/\tau(A)]\tau(A)\}$ is called the set of principal periods of A while $\mathcal{R}(A)$ is called the set of secondary periods of A. Furthermore, put $r_A = \#\mathcal{R}(A)$. Observe that one has $0 \leq r_A < n/2$. Returns before $\tau(A)$ are not possible, thus, $I\!\!P_A\left(\tau_A < \tau(A)\right) = 0$. Still, if A does not return at time $\tau(A)$, then it can not return at times $k\tau(A)$, with $1 \leq k \leq [n/\tau(A)]$, so one has

$$I\!\!P_A\left(\tau(A) < \tau_A \leq [n/\tau(A)]\tau(A)\right) = 0.$$

The first possible return after $\tau(A)$ is

$$n_A = \begin{cases} \min \mathcal{R}(A) & \mathcal{R}(A) \neq \emptyset \\ n_A = n & \mathcal{R}(A) = \emptyset \end{cases} \ .$$

Furthermore, by definition of $\mathcal{R}(A)$ one has $A \bigcap \mathcal{R}(A)^c = \emptyset$. Thus

$$I\!\!P_A\left(\{[n/\tau(A)]\tau(A)+1 \leq \tau_A \leq n-1\} \cap \{\tau_A \notin \mathcal{R}(A)\}\right) = 0.$$

We finally remark that $\{T^{-i}A \cap T^{-j}A \mid i,j \in \mathcal{R}(A)\} = \emptyset$. Otherwise it would contradict the fact that the first return to A is $\tau(A)$ since for $i,j \in \mathcal{R}(A)$ one has $|i-j| < \tau(A)$. We conclude that

$$I\!\!P_A\left(T^{-i}A \cap T^{-j}A \mid i,j \in \mathcal{R}(A)\right) = 0. \tag{4}$$

4 Return times

Given $A \in \mathcal{C}_n$, we define the *hitting time* $\tau_A : \Omega \to I\!\!N \cup \{\infty\}$ as the following random variable: For any $x \in \Omega$

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

The *return time* is the hitting time restricted to the set A, namely $\tau_A|_A$. We remark the difference between τ_A and $\tau(A)$: while $\tau_A(x)$ is the first time A appears in x, $\tau(A)$ is the first overlapping position of A.

For $A \in \mathcal{C}_n$ define

$$\zeta_A \stackrel{def}{=} I\!\!P_A(\tau_A \neq \tau(A)) = I\!\!P_A(\tau_A > \tau(A)) \ .$$

The equality follows by the comment at the end of the previous section.

It would be useful for the reader to note now that according to the comments of the previous section, one has

$$\tau_A|_A \ \in \ \{\tau(A)\} \cup \mathcal{R}(A) \cup \{k \in I\!\!N \mid k \geq n\} \ . \tag{5}$$

Theorem 2 *Let $\{X_m\}_{m \in Z\!\!Z}$ be a ϕ-mixing process. Then, there exist a strictly positive constant C_1 such that for any $A \in \mathcal{C}_n$, the following inequality holds for all t:*

$$\left| I\!\!P_A\left(\tau_A > t\right) - 1\!\!1_{\{t<\tau(A)\}} - 1\!\!1_{\{t \geq \tau(A)\}} \zeta_A e^{-\zeta_A I\!\!P(A)(t-\tau(A))} \right| \leq C_1 \epsilon(A) f(A,t), \tag{6}$$

where $f(A,t) = I\!\!P(A) t e^{-(\zeta_A - \epsilon(A)) I\!\!P(A) t}$ and

$$\epsilon(A) \stackrel{def}{=} \inf_{0 \leq w \leq n_A} \left[(r_A + n) I\!\!P(A^{(w)}) + \phi\left(n_A - w\right) \right] \ . \tag{7}$$

We postpone an example showing the sharpness of $\epsilon(A)$ after Lemma 13.

Remark 3 *$A^{(n_A)}$ is the part of the string A which does not overlap itself in $A \cap T^{-n_A} A$. Note that n_A is the position of the first possible return if the process does not comes back at $\tau(A)$. Recall that $n_A = n$ if $\mathcal{R}(A) = \emptyset$. $I\!\!P(A^{(w)})$ with $1 \leq w \leq n_A$ is the part of the string $A^{(n_A)}$ after taking out its first w letters (this will be to create a gap of length w to use the mixing property).*

Remark 4 *When $\mathcal{R}(A) = \emptyset$, namely, A does not overlapps itself, the error of Theorem 2 reduces to $\epsilon(A) = \inf_{0 \leq w \leq n} \left[n I\!\!P(A^{(w)}) + \phi\left(n - w\right) \right]$.*

Remark 5 *In the error term of the theorem, $\epsilon(A)$ provides a bound which shows the convergence uniform in t of the return time law to that mixture of laws as the length of the string growths. The factor $I\!\!P(A)t$ provides an extra bound for values of t smaller than $1/I\!\!P(A)$. The factor $e^{-(\zeta_A - \epsilon(A)))I\!\!P(A)t}$ provides an extra bound for values of t larger than $1/I\!\!P(A)$.*

Remark 6 *On one hand $I\!\!P(A) \leq Ce^{-cn}$ (see [1]). On the other hand, by construction $n_A > n/2$. Further $\phi(n) \to 0$ as $n \to \infty$. Taking for instance $w = n/4$ in (7) we warranty the smallness of $\epsilon(A)$ for large enough n.*

Corollary 7 *Let the process $\{X_m\}_{m \in Z\!\!Z}$ be ϕ-mixing. Let $\beta > 0$. Then, for any $A \in \mathcal{C}_n$, the β-moment of the re-scaled time $I\!\!P(A)\tau_A$ converges, as $n \to \infty$, to $\Gamma(\beta+1)/\zeta_A^{\beta-1}$. Moreover*

$$\left| I\!\!P(A)^\beta I\!\!E_A(\tau_A^\beta) - \frac{\Gamma(\beta+1)}{\zeta_A^{\beta-1}} \right| \leq \epsilon^*(A) \frac{C\beta \ e^{2\epsilon(A)(\beta+1)/\zeta_A}}{\zeta_A^2} \ \frac{\Gamma(\beta+1)}{\zeta_A^{\beta-1}} \ , \tag{8}$$

where $\epsilon^*(A) = (\epsilon(A) \vee (nI\!\!P(A))^\beta)$, C is a constant and Γ is the analytic gamma function.

Remark 8 *In particular, the corollary establishes that all the moments of the return time are finite.*

Remark 9 *In the special case when $\beta = 1$, the above corollary establishes a weak version of Kac's Lemma (see [15]).*

Remark 10 *For each β fixed and n large enough one has $\beta\, e^{2\epsilon(A)(\beta+1)}/\zeta_A^2$ is close to β/ζ_A^2. Thus in virtue of inequality (8), the corollary reads not just as a difference result but also as a ratio result.*

The next corollary extends Theorem 2.1 in [14].

Corollary 11 *Let the process $\{X_m\}_{m\in\mathbb{Z}}$ be ϕ-mixing. There exists a constant $C > 0$ such that, for each $A \in \mathcal{C}_n$, the following conditions are equivalent:*

(a) $\left|\mathbb{P}_A(\tau_A > t) - e^{-\mathbb{P}(A)t}\right| \le C\ \epsilon(A)\ f(A,t)$,

(b) $\left|\mathbb{P}_A(\tau_A > t) - \mathbb{P}(\tau_A > t)\right| \le C\ \epsilon(A)\ f(A,t)$,

(c) $\left|\mathbb{P}(\tau_A > t) - e^{-\mathbb{P}(A)t}\right| \le C\ \epsilon(A)\ f(A,t)$,

(d) $|\zeta_A - 1| \le C\ \epsilon(A)$.

Moreover, if $\{A_n\}_{n\in\mathbb{N}}$ is a sequence of strings such that $\mathbb{P}(A_n) \to 0$ as $n \to \infty$, then the following conditions are equivalent:
($\tilde{a}$) the return time law of A_n converges to a parameter one exponential law,
($\tilde{b}$) the return time law and the hitting time law of A_n converge to the same law,
($\tilde{c}$) the hitting time law of A_n converges to a parameter one exponential law,
($\tilde{d}$) $\lim_{n\to\infty} \zeta_{A_n} = 1$.

4.1 Preparatory results

Here we collect a number of results that will be useful for the proof of Theorem 2. The next lemma is a useful way to use the ϕ-mixing property.

Lemma 12 *Let $\{X_m\}_{m\in\mathbb{Z}}$ be a ϕ-mixing process. Suppose that $B \subseteq A \in \mathcal{F}_{\{0,\dots,b\}}, C \in \mathcal{F}_{\{b+g,\infty\}}$ with $b, g \in \mathbb{N}$. The following inequality holds:*

$$\mathbb{P}_A(B; C) \le \mathbb{P}_A(B)\left(\mathbb{P}(C) + \phi(g)\right) .$$

Proof Since $B \subseteq A$, obviously $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(B \cap C)$. By the ϕ-mixing property $\mathbb{P}(B; C) \le \mathbb{P}(B)\left(\mathbb{P}(C) + \phi(g)\right)$. Dividing the above inequality by $\mathbb{P}(A)$ the lemma follows. $\Box$

The following lemma says that returns over $\mathcal{R}(A)$ have small probability.

Lemma 13 *Let $\{X_m\}_{m\in\mathbb{Z}}$ be a ϕ-mixing process. For all $A \in \mathcal{C}_n$, the following inequality holds:*

$$\mathbb{P}_A(\tau_A \in \mathcal{R}(A)) \le \epsilon(A) . \tag{9}$$

Proof For any w such that $1 \leq w \leq n_A$

$$\begin{aligned} I\!\!P_A\left(\tau_A \in \mathcal{R}(A)\right) &= I\!\!P_A\left(\bigcup_{j\in\mathcal{R}(A)} T^{-j}A\right) \\ &\leq I\!\!P_A\left(\bigcup_{j\in\mathcal{R}(A)} T^{-j}A^{(w)}\right) \\ &\leq r_A I\!\!P\left(A^{(w)}\right) + \phi(n_A - w)\ . \end{aligned} \tag{10}$$

The equality follows by (4). Since $\{T^{-j}A\} \subset \{T^{-j}A^{(w)}\}$, first inequality follows. Second one follows by the above lemma with $B = A$ and $C = \cup_{j\in\mathcal{R}(A)}T^{-j}A^{(w)}$. This ends the proof since w is arbitrary. $\square$

Example 14 *Consider a process* $\{X_m\}_{m\in Z}$ *defined on the alphabet* $\mathcal{E} = \{a, b\}$. *Consider the string introduced in (3):*

$$A = ((X_0 ... X_{14}) = (\texttt{aaaabbaaaabbaaa}))\ .$$

Then, $n = 15$, $\tau(A) = 6$, $\mathcal{R}(A) = \{13, 14\}$, $r_A = 2$ *and* $n_A = 13$. *Thus*

$$A^{(13)} = ((X_2 ... X_{14}) = (\texttt{aabbaaaabbaaa}))\ .$$

The ϕ*-mixing property factorizes the probability*

$$I\!\!P_A\left(\bigcup_{j=13}^{14} T^{-j}A\right) = I\!\!P_A\left(\bigcup_{j=13}^{14} T^{-j}A^{(13)}\right) \leq I\!\!P_A\left(\bigcup_{j=13}^{14} T^{-j}A^{(w)}\right)\ .$$

In such case, a gap at $t = 15$ *of length* w *with* $0 \leq w \leq 13$ *is the best we can do to apply the* ϕ*-mixing property.*

The next lemma will be used to get the non-uniform factor $f(A, t)$ in the error term of Theorem 2.

Lemma 15 *Let* $\{X_m\}_{m\in Z}$ *be a* ϕ*-mixing process. Let* $B \in \mathcal{F}_{\{kf,\infty\}}$, *with* $k \in I\!\!N$ *and* $f > n$. *Then, for all* $n < g < f$, *the following inequality holds:*

$$I\!\!P_A\left(\tau_A > kf\ ;\ B\right) \leq \left[I\!\!P\left(\tau_A > f - g\right) + \phi(g)\right]^{k-1}\left[I\!\!P\left(B\right) + \phi(g)\right]\ . \tag{11}$$

Proof First introduce a gap of length g, then use Lemma 12 to get the inequalities

$$\begin{aligned} I\!\!P_A\left(\tau_A > kf\ ;\ B\right) &\leq I\!\!P_A\left(\tau_A > kf - g\ ;\ B\right) \\ &\leq I\!\!P_A\left(\tau_A > kf - g\right)\left[I\!\!P\left(B\right) + \phi(g)\right]\ . \end{aligned} \tag{12}$$

Apply the above procedure to $(\tau_A > (k-1)f)$ and $B = (\tau_A > f - g)$ to bound $I\!\!P_A\left(\tau_A > kf - g\right)$ by

$$I\!\!P_A\left(\tau_A > (k-1)f - g\right)\left[I\!\!P\left(\tau_A > f - g\right) + \phi(g)\right]\ .$$

Iterate this procedure to bound $I\!\!P_A(\tau_A > kf - g)$ by

$$I\!\!P_A(\tau_A > f - g)\left[I\!\!P(\tau_A > f - g) + \phi(g)\right]^{k-1} \leq \left[I\!\!P(\tau_A > f - g) + \phi(g)\right]^{k-1} .$$

This ends the proof of the Lemma. $\Box$

The next proposition establishes a relationship between *hitting* and *return* times with an error *uniform* in t. In particular, (b) says that they coincide if and only if $\zeta_A = 1$, namely, the string A is non-self-repeating.

Proposition 16 *Let $\{X_m\}_{m\in Z}$ be a ϕ-mixing processes. Let $A \in \mathcal{C}_n$. Then the following holds:*

(a) For all $M, M' \geq g \geq n$,

$$\begin{aligned} & \left|I\!\!P_A(\tau_A > M + M') - I\!\!P_A(\tau_A > M)\, I\!\!P(\tau_A > M')\right| \\ \leq\ & I\!\!P_A(\tau_A > M - g)\, 2\left[gI\!\!P(A) + \phi(g)\right], \end{aligned}$$

and similarly

$$\begin{aligned} & \left|I\!\!P_A(\tau_A > M + M') - I\!\!P_A(\tau_A > M)\, I\!\!P(\tau_A > M' - g)\right| \\ \leq\ & I\!\!P_A(\tau_A > M - g)\left[gI\!\!P(A) + 2\phi(g)\right]. \end{aligned}$$

(b) For all $t \geq \tau(A) \in I\!\!N$,

$$\left|I\!\!P_A(\tau_A > t) - \zeta_A I\!\!P(\tau_A > t)\right| \leq 2\epsilon(A) \ . \tag{13}$$

Proof To simplify notation, for $t \in Z\!\!\!Z$ we write $\tau_A^{[t]}$ to mean $\tau_A \circ T^t$. We introduce a gap of length g after coordinate M to construct the following triangular inequality

$$\begin{aligned} & \left|I\!\!P_A(\tau_A > M + M') - I\!\!P_A(\tau_A > M)\, I\!\!P(\tau_A > M')\right| \\ \leq\ & \left|I\!\!P_A(\tau_A > M + M') - I\!\!P_A\left(\tau_A > M; \tau_A^{[M+g]} > M' - g\right)\right| && (14) \\ +\ & \left|I\!\!P_A\left(\tau_A > M; \tau_A^{[M+g]} > M' - g\right) - I\!\!P_A(\tau_A > M)\, I\!\!P(\tau_A > M' - g)\right| && (15) \\ +\ & I\!\!P_A(\tau_A > M)\left|I\!\!P(\tau_A > M' - g) - I\!\!P(\tau_A > M')\right| \ . && (16) \end{aligned}$$

Term (14) is bounded as in (12) by

$$I\!\!P_A\left(\tau_A > M; \tau_A^{[M]} \leq g\right) \leq I\!\!P_A(\tau_A > M - g)\left[gI\!\!P(A) + \phi(g)\right] \ .$$

Term (15) is bounded using the ϕ-mixing property by $I\!\!P_A(\tau_A > M)\,\phi(g)$. The modulus in (16) is bounded using stationarity by $I\!\!P(\tau_A \leq g) \leq gI\!\!P(A)$. This ends the proof of both inequalities of item (a).

Item (b) for $t \geq 2n$ is proven applying item (a) with $M = n$ and $M' = t - n$. Then, by stationarity $I\!\!P(\tau_A > t - n) - I\!\!P(\tau_A > t) \leq nI\!\!P(A)$. Further, $I\!\!P_A(\tau_A > n) - I\!\!P_A(\tau_A > \tau(A)) \leq \epsilon(A)$ by Lemma 13.

Consider now $\tau(A) \leq t < 2n$. Take any $1 \leq w \leq n_A$.

$$\begin{aligned}\zeta_A - I\!\!P_A(\tau_A > t) &= I\!\!P_A(\tau(A) < \tau_A \leq t) \\ &= I\!\!P_A(\tau_A \in \mathcal{R}(A) \cup (n \leq \tau_A < 2n)) \\ &\leq (r_A + n) I\!\!P\left(A^{(w)}\right) + \phi(n_A - w) \ . \end{aligned} \tag{17}$$

First and second equalities follow by the considerations of section 3. The inequality follows similarly to (10). □

The following two propositions are the key of the proof of Theorem 2.

Proposition 17 *Let $\{X_m\}_{m\in Z}$ be a ϕ-mixing process. Let $n < g < f$. Then the following inequality holds:*

$$\begin{aligned}&\left| I\!\!P_A(\tau_A > kf) - I\!\!P_A(\tau_A > f)\, I\!\!P(\tau_A > f-g)^{k-1} \right| \\ \leq\ & 2\left[g I\!\!P(A) + \phi(g)\right](k-1)\left(I\!\!P(\tau_A > f-g) + \phi(g)\right)^{k-2} \ .\end{aligned}$$

Proof The left hand side of the above inequality is bounded by

$$\sum_{j=2}^{k} \left| I\!\!P_A(\tau_A > jf) - I\!\!P_A(\tau_A > (j-1)f)\, I\!\!P(\tau_A > f-g) \right| I\!\!P(\tau_A > f-g)^{k-j} \ .$$

The modulus in the above sum is bounded by

$$2\left[g I\!\!P(A) + \phi(g)\right] I\!\!P_A(\tau_A > (j-1)f - g) \ ,$$

due to Proposition 16 (a). The right-most factor is bounded using Lemma 15 by $\left[I\!\!P(\tau_A > f-g) + \phi(g)\right]^{j-2}$. The conclusion follows. □

Let us define

$$\delta(A) = \inf_{n \leq y \leq 1/I\!\!P(A)} \left(y I\!\!P(A) + \phi(y)\right) \ .$$

In the proof of Theorem 2 we will make use of the following version of Proposition 17 proved in [2] for *hitting* times instead of *return* times as was done in Proposition 17. We quote it here for easy reference.

Proposition 18 (Abadi, 2004) *Let $\{X_m\}_{m\in Z}$ be a ϕ-mixing process. There exists a finite constant $C > 0$, such that for any $f \in (4n, 1/(2I\!\!P(A))]$ such that*

$$\phi(f/4) \leq I\!\!P\left(\tau_A \leq f/4 \ ; \ \tau_A^{[f/4]} > f/2\right) \ , \tag{18}$$

there exists a $\Delta = \Delta(f) > 0$, with $n < \Delta \leq f/4$, such that for all positive integers k, the following inequalities hold:

$$\left| I\!\!P(\tau_A > kf) - I\!\!P(\tau_A > f - 2\Delta)^k \right| \leq C\delta(A) k I\!\!P(\tau_A > f - 2\Delta)^k \ , \tag{19}$$

and

$$\left| I\!\!P(\tau_A > kf) - I\!\!P(\tau_A > f)^k \right| \leq C\delta(A) k I\!\!P(\tau_A > f - 2\Delta)^k \ . \tag{20}$$

4.2 Proofs of Theorem 2 and corollaries

Proof of Theorem 2 We divide the proof according the different values of t: (i) $t < n$, (ii) $n \le t \le 1/(2I\!\!P(A))$ and (ii) $t > 1/(2I\!\!P(A))$. (Factor 2 is rather technical.)

Consider first $t \le n$. If $t \le \tau(A)$, (5) says that the left hand side of (6) is zero. If $\tau(A) < t \le [n/\tau(A)]\tau(A)$, also (5) implies that the left hand side of equation (6) is $\zeta_A - \zeta_A e^{-\zeta_A I\!\!P(A)(t-\tau(A))} \le I\!\!P(A)n$. If $[n/\tau(A)]\tau(A) < t < n$ it follows by (5) and Lemma 13 that the left hand side of (6) is bounded by $\epsilon(A)$.

Consider now $n \le t \le 1/(2I\!\!P(A))$. First write

$$I\!\!P_A\left(\tau_A > t\right) = \frac{I\!\!P_A\left(\tau_A > t\right)}{I\!\!P\left(\tau_A > t\right)} I\!\!P\left(\tau_A > t\right) = \rho_{t+1} I\!\!P\left(\tau_A > t\right) \ ,$$

and

$$\begin{aligned} I\!\!P\left(\tau_A > t\right) &= \prod_{i=\tau(A)+1}^{t} I\!\!P\left(\tau_A > i | \tau_A > i-1\right) \\ &= \prod_{i=\tau(A)+1}^{t} \left(1 - I\!\!P\left(T^{-i}A | \tau_A > i-1\right)\right) \\ &= \prod_{i=\tau(A)+1}^{t} \left(1 - \rho_i I\!\!P(A)\right) \ , \end{aligned}$$

where

$$\rho_i \stackrel{def}{=} \frac{I\!\!P_A\left(\tau_A > i-1\right)}{I\!\!P\left(\tau_A > i-1\right)} \ .$$

Further

$$\left|1 - \rho_i I\!\!P(A) - e^{-\zeta_A I\!\!P(A)}\right| \le \left|\rho_i - \zeta_A\right| I\!\!P(A) + \left|1 - \zeta_A I\!\!P(A) - e^{-\zeta_A I\!\!P(A)}\right| \ . \quad (21)$$

Firstly, by Proposition 16 (b) and the fact that $I\!\!P\left(\tau_A > i\right) \ge 1/2$ since $i \le 1/(2I\!\!P(A))$ we have

$$\left|\rho_i - \zeta_A\right| \le \frac{2\epsilon(A)}{I\!\!P\left(\tau_A > i\right)} \le 4\epsilon(A) \ .$$

for all $i = \tau(A)+1, \ldots, 1/(2I\!\!P(A))$. Secondly, note that $|1 - x - e^{-x}| \le x^2/2$ for $x > 0$ small enough. Apply it with $x = \zeta_A I\!\!P(A)$ to bound the most right term of (21) by $(\zeta_A I\!\!P(A))^2/2$. Further, since $|\prod a_i - \prod b_i| \le (\#i) \max |a_i - b_i|$ for $0 \le a_i, b_i \le 1$, we conclude that $|I\!\!P\left(\tau_A > t\right) - e^{-\zeta_A I\!\!P(A)t}|$ and therefore the left hand side of (6) are both bounded by

$$(t - \tau(A)) \left(4\epsilon(A) I\!\!P(A) + \frac{I\!\!P(A)^2}{2}\right) \le C\epsilon(A) I\!\!P(A) t \ e^{-\zeta_A I\!\!P(A)t} \ , \qquad (22)$$

for all $\tau_A \le t \le 1/(2I\!\!P(A))$. The inequality follows since $e^{-\zeta_A I\!\!P(A)t} \ge e^{-1/2}$.

Finally, consider $t > 1/(2I\!\!P(A))$. The proof has two steps. First we prove for t of the form $t = (k + p/q)\ f$ with $f = 1/(2I\!\!P(A))$, $k \in I\!\!N$, p a positive integer and $1 \le p \le q$ with $q := 1/(2\delta(A))$. The basic tools are the Mean Value Theorem (MVT) and the ϕ-mixing property. Then 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 $t = \left(k + \frac{p}{q}\right)\ f$.

Let $t = (k + (p/q))\ f$, with k, p, q and f as was just told. For brevity put $r = (p/q)\ f$. Let Δ be the one given by Proposition 18. Then

$$
\begin{aligned}
& \left| I\!\!P_A\left(\tau_A > t\right) - \zeta_A e^{-\zeta_A I\!\!P(A)t} \right| \\
= \ & \left| I\!\!P_A\left(\tau_A > kf + r\right) - \zeta_A e^{-(\zeta_A/2)t/f} \right| \\
\le \ & \left| I\!\!P_A\left(\tau_A > kf + r\right) - I\!\!P_A\left(\tau_A > kf\right) I\!\!P\left(\tau_A > r\right) \right| \\
+ \ & \left| I\!\!P_A\left(\tau_A > kf\right) - I\!\!P_A\left(\tau_A > f\right) I\!\!P\left(\tau_A > f - \Delta\right)^{k-1} \right| \ I\!\!P\left(\tau_A > r\right) \\
+ \ & \left| I\!\!P\left(\tau_A > f - \Delta\right)^{k-1} - I\!\!P\left(\tau_A > f - 2\Delta\right)^{k-1} \right| I\!\!P_A\left(\tau_A > f\right) I\!\!P\left(\tau_A > r\right) \\
+ \ & \left| I\!\!P_A\left(\tau_A > f\right) - \zeta_A I\!\!P\left(\tau_A > f - 2\Delta\right) \right| I\!\!P\left(\tau_A > f - 2\Delta\right)^{k-1} I\!\!P\left(\tau_A > r\right) \\
+ \ & \left| I\!\!P\left(\tau_A > r\right) - I\!\!P\left(\tau_A > f - 2\Delta\right)^{r/f} \right| \ \zeta_A I\!\!P\left(\tau_A > f - 2\Delta\right)^{k} \\
+ \ & \left| I\!\!P\left(\tau_A > f - 2\Delta\right)^{t/f} - e^{-(\zeta_A/2)t/f} \right| \ \zeta_A \ .
\end{aligned}
$$

The first term on the right hand side of the above inequality is bounded using first Proposition 16 (a) with $M = kf$, $M' = r$ and $g = \Delta$ and then Lemma 15 with $B = \{\tau_A > f - g\}$ by

$$2\left(\Delta I\!\!P(A) + \phi(\Delta)\right)\left(I\!\!P\left(\tau_A > f - \Delta\right) + \phi(\Delta)\right)^{k-1} \ .$$

The modulus in the second one is bounded using Proposition 17 by

$$2\left(\Delta I\!\!P(A) + \phi(\Delta)\right)(k-1)\left(I\!\!P\left(\tau_A > f - \Delta\right) + \phi(\Delta)\right)^{k-2} \ .$$

The modulus in the third one is bounded using the MVT by

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

The modulus in the fourth one is bounded using Proposition 16 (b) by $2\epsilon(A)$. The modulus in the fifth one is bounded by

$$C\delta(A) I\!\!P\left(\tau_A > f - 2\Delta\right)^{[k+(p/q)]/2} \ ,$$

as shown in the proof of Theorem 1 of [2] (see p. 254). The modulus in the sixth one is bounded using the MVT and (22) with $t = f - 2\Delta$ by

$$\epsilon(A)\frac{t}{f}\left(I\!\!P\left(\tau_A > f - 2\Delta\right) \vee e^{-(\zeta_A/2)}\right)^{(t/f)-1} \le C\epsilon(A) I\!\!P(A) t e^{-(\zeta_A - \epsilon(A)) I\!\!P(A)t} \ .$$

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 $0 \leq r < f$. We can choose a $\bar{t}$ such that $\bar{t} > t$ and $\bar{t} = (k + (p/q))\, f$ with p, q as before. Then

$$\begin{aligned}
& \left| I\!P_A (\tau_A > t) - \zeta_A e^{-\zeta_A I\!P(A) t} \right| \\
\leq\ & \left| I\!P_A (\tau_A > t) - I\!P_A (\tau_A > \bar{t}\ ) \right| \\
+\ & \left| I\!P_A (\tau_A > \bar{t}) - \zeta_A I\!P (\tau_A > f - 2\Delta)^{[k+(p/q)]/2} \right| \\
+\ & \zeta_A \left| I\!P (\tau_A > f - 2\Delta)^{[k+(p/q)]/2} - e^{-\zeta_A I\!P(A)\bar{t}} \right| \\
+\ & \zeta_A \left| e^{-\zeta_A I\!P(A)\bar{t}} - e^{-\zeta_A I\!P(A) t} \right| \ .
\end{aligned}$$

The first term on the right hand side of the above inequality is bounded applying Lemma 15

$$\begin{aligned}
& \left| I\!P_A (\tau_A > t) - I\!P_A (\tau_A > \bar{t}\ ) \right| \\
=\ & I\!P_A \left(\tau_A > t\ ;\ \tau_A^{[\bar{t}]} \leq \bar{t} - t \right) \\
\leq\ & I\!P_A \left(\tau_A > (k-1) f\ ;\ \tau_A^{[\bar{t}]} \leq \Delta \right) \\
\leq\ & \left(I\!P (\tau_A > f - \Delta) + \phi(\Delta) \right)^{k-2} (\Delta I\!P(A) + \phi(\Delta)) \ .
\end{aligned}$$

For the third term, first note that $e^{-\zeta_A I\!P(A)\bar{t}} = e^{-\zeta_A [k+(p/q)]/2}$. Yet, by stationarity and (22)

$$\left| I\!P (\tau_A > f - 2\Delta) - e^{-\zeta_A} \right| \leq C\epsilon(A) \ .$$

Therefore, the MVT implies that the third term is bounded by

$$\begin{aligned}
& C\epsilon(A)\ I\!P(A)\bar{t}\ \left(I\!P (\tau_A > f - 2\Delta) \vee e^{-\zeta_A} \right)^{I\!P(A)\bar{t} - 1} \\
\leq\ & C\epsilon(A)\ I\!P(A) t\ e^{-(\zeta_A - \epsilon(A)) I\!P(A) t} \ .
\end{aligned}$$

The upper bound for the fourth term is obtained similarly by the MVT and the fact that $|t - \bar{t}| \leq \Delta$. Finally, the second term is bounded as in the first part of the proof. To end the proof we notice that

$$I\!P (\tau_A > f - \Delta) \leq I\!P (\tau_A > f - \Delta) + \phi(\Delta) = I\!P (\tau_A > f - 2\Delta) \ .$$

The equality follows since $\phi(\Delta) = I\!P \left(\tau_A \leq \Delta; \tau_A^{[\Delta]} > f - 2\Delta \right)$ (see [2] p. 250.) Therefore

$$I\!P (\tau_A > f - 2\Delta)^{k-2} \leq C e^{-(\zeta_A - \epsilon(A)) I\!P(A) t} \ .$$

This ends the proof of the theorem. $\square$

Proof of Corollary 7 Rewrite (6) as

$$\left|\mathbb{P}_A(\mathbb{P}(A)\tau_A > t) - 1\!\!1_{\{t<\mathbb{P}(A)\tau(A)\}} - 1\!\!1_{\{t\geq\mathbb{P}(A)\tau(A)\}}\zeta_A e^{-\zeta_A(t-\mathbb{P}(A)\tau(A))}\right|$$
$$\leq C_1\epsilon(A)f(A,t/\mathbb{P}(A)) \ . \tag{23}$$

Let $Y = Y_1 + Y_2$ where

$$\mathbb{P}(Y_1 > t) = 1\!\!1_{\{t<\mathbb{P}(A)\tau(A)\}}$$

and

$$\mathbb{P}(Y_2 > t) = 1\!\!1_{\{t\geq\mathbb{P}(A)\tau(A)\}}\zeta_A e^{-\zeta_A(t-\mathbb{P}(A)\tau(A))}.$$

Integrating (23) we get

$$\begin{aligned}
&\left|\mathbb{E}\left((\mathbb{P}(A)\tau_A)^\beta\right) - \mathbb{E}\left(Y^\beta\right)\right| \\
= &\left|\int_{\mathbb{P}(A)}^\infty \beta t^{\beta-1}\left(\mathbb{P}\left(\mathbb{P}(A)\tau_A > t\right) - \mathbb{P}\left(Y > t\right)\right)\right| \\
\leq &\int_{\mathbb{P}(A)}^\infty \beta t^{\beta-1}\left|\mathbb{P}\left(\mathbb{P}(A)\tau_A > t\right) - \mathbb{P}\left(Y > t\right)\right| \\
\leq &\ C_1\epsilon(A)\int_{\mathbb{P}(A)}^\infty \beta t^{\beta-1}f(A,t/\mathbb{P}(A))dt \ .
\end{aligned} \tag{24}$$

Now we procide to compute $\mathbb{E}\left(Y^\beta\right) = \int_{\mathbb{P}(A)}^\infty \beta t^{\beta-1}\mathbb{P}(Y > t)$.

Since $\mathbb{P}(Y_1 > t)$ and $\mathbb{P}(Y_2 > t)$ have disjoint support, one has $\mathbb{E}(Y^\beta) = \mathbb{E}(Y_1^\beta) + \mathbb{E}(Y_2^\beta)$. On one hand $\mathbb{E}(Y_1^\beta) = (\mathbb{P}(A)\tau(A))^\beta$. On the other hand

$$\begin{aligned}
\mathbb{E}(Y_2^\beta) &= \int_{\mathbb{P}(A)\tau_A}^\infty \beta t^{\beta-1}\zeta_A e^{-\zeta_A(t-\mathbb{P}(A)\tau_A)}dt \\
&= \zeta_A\ e^{\zeta_A\mathbb{P}(A)\tau_A}\left(\int_0^\infty - \int_0^{\mathbb{P}(A)\tau_A}\right)\beta t^{\beta-1}e^{-\zeta_A t}dt \ .
\end{aligned}$$

Yet, since $\zeta_A\mathbb{P}(A)\tau_A \leq \mathbb{P}(A)n$ we have and $\mathbb{P}(A)$ decays exponentially fast we have $e^{\zeta_A\mathbb{P}(A)\tau_A} - 1 \leq C\mathbb{P}(A)n$. Further, the first integral is $\Gamma(\beta+1)/\zeta_A^\beta$. The second one is bounded by $(\mathbb{P}(A)\tau_A)^\beta$. We conclude that

$$\left|\mathbb{E}\left(Y^\beta\right) - \mathbb{E}\left(Y_2^\beta\right)\right| \leq Cn\mathbb{P}(A) + 2(n\mathbb{P}(A))^\beta) \leq C(n\mathbb{P}(A))^{(\beta\wedge 1)} \ .$$

Similar computations give

$$\int_{\mathbb{P}(A)}^\infty \beta t^{\beta-1}f(A,t/\mathbb{P}(A))dt \leq \frac{\beta}{\beta+1}\frac{\Gamma(\beta+2)}{(\zeta_A-\epsilon(A))^{\beta+1}} \leq \frac{\beta e^{2\epsilon(A)(\beta+1)/\zeta_A}}{\zeta_A^2}\frac{\Gamma(\beta+1)}{\zeta_A^{\beta-1}} \ .$$

In the last inequality we used $x \leq 2(1-e^{-x})$ for small enough $x > 0$. This ends the proof of the corollary. □

Proof of Corollary 11. $(a) \Leftrightarrow (d)$. It follows directly from Theorem 2.

$(b) \Rightarrow (a), (c)$. It follows by Theorem 2 and Theorem 1 in [2]

$(a) \Rightarrow (b)$ and $(c) \Rightarrow (b)$. They follow by Theorem 2, Theorem 1 in [2] and (22). The corollary is proven. $\square$

5 Sojourn time

Definition 19 *Let $A \in \mathcal{C}_n$. We define the* sojourn time *on the set A as the r.v. $S_A : \Omega \to I\!N \cup \{\infty\}$*

$$S_A(x) = \sup \left\{ k \in I\!N | x \in A \cap T^{-j\tau(A)}A \ ; \ \forall j = 1, \ldots, k \right\} \ ,$$

and $S_A(x) = 0$ if the supremum is taken over the empty set.

Definition 20 *Given $A \in \mathcal{C}_n$, we define the sequence of probabilities $(p_i(A))_{i \in I\!N}$ as follows:*

$$p_i(A) \stackrel{def}{=} I\!P \left(A \mid \bigcap_{j=1}^{i} T^{j\tau(A)}A \right) \ .$$

By stationarity $p_1(A) = 1 - \zeta_A$. Whenever A is fixed we will avoid the dependence on A in $p_i(A)$ writing just p_i.

Example 21 *For a i.i.d. Bernoulli process with parameter $0 < \theta = I\!P(X_i = 1) = 1 - I\!P(X_i = 0)$, and for the n-string $A = \{X_0^{n-1} = 1\}$, we have that $p_i = 1 - \zeta_A = \theta$ for all $i \in I\!N$.*

Example 22 *Let $\{X_m\}_{m \in Z\!\!Z}$ be a irreducible and aperiodic finite state Markov chain. For $A = \{X_0^{n-1} = a_0^{n-1}\} \in \mathcal{C}_n$, the sequence $(p_i)_{i \in I\!N}$ is constant. More precisely, by the Markovian property and for all $i \in I\!N$*

$$\begin{aligned} p_i &= I\!P \left(X_{n-\tau(A)}^{n-1} = a_{n-\tau(A)}^{n-1} | X_{\tau(A)-1} = a_{\tau(A)-1} \right) \\ &= \prod_{j=\tau(A)}^{n-1} I\!P \left(X_j = a_j | X_{j-1} = a_{j-1} \right) \ . \end{aligned}$$

In the following theorem we assume that $(p_i(A))_{i \in I\!N}$ converges with velocity $d_i = d_i(A)$. Namely, there is a real number $\rho(A) \in (0, 1)$ such that

$$|p_i(A) - \rho(A)| \leq d_i \qquad \text{for all } i \in I\!N, \tag{25}$$

where d_i is a sequence decreasing to zero.

Remark 23 *In the previous two examples, the sequence $(p_i(A))_{i \in I\!N}$ not just converges but even is constant, so $d_i = 0$ for all $i \in I\!N$.*

Theorem 24 *Let* $\{X_m\}_{m\in Z}$ *be a stationary process. Let* $A \in \mathcal{C}_n$*. Assume that (25) holds. Then, there is a constant* $c \in [0,1)$*, such that the following inequalities hold for all* $k \in I\!N$*:*

$$\left|I\!P_A\left(S_A = k\right) - \left(1-\rho(A)\right)\,\rho(A)^k\right| \le c^k \sum_{i=1}^{k+1} d_i \le c^k (k+1) d_1 \ .$$

Remark 25 *If* $\{A_n\}_{n\in I\!N}$ *is a sequence of strings for which* $\tau(A_n) \le sn$ *for all* n *and fixed* $0 < s < 1$*, and* $d_1(A_n)$ *goes to zero as* $n \to \infty$*, then the* β*-moments of* $(S_{A_n})_{n\in I\!N}$ *converge to* $I\!E(Y^\beta)$ *with velocity* $d_1 = d_1(A_n)$*. (Observe that we don't require the strings* A_n *being decreasing, namely* $A_{n+1} \subset A_n$ *for all* $n \in I\!N$*).*

Corollary 26 *Let* Y *be a r.v. with geometric distribution with parameter* $\rho(A)$*. Let* β *a positive integer. Then*

$$\left|I\!E_A\left(S_A^\beta\right) - I\!E(Y^\beta)\right| \le C_\beta d_1 \ ,$$

where C_β *is a constant that just depends on* β*.*

Lemma 27 *Let* $(l_i)_{i\in I\!N}$ *be a sequence of real numbers such that* $0 < l_i < 1$*, for all* $i \in I\!N$*. Let* $0 \le l < 1$ *be such that* $|l_i - l| \le d_i$ *for all* $i \in I\!N$ *with* $d_i \to 0$*. Then, there is a constant* $c \in [0,1)$*, such that the following inequalities hold for all* $k \in I\!N$*:*

$$\left|\prod_{i=1}^k l_i - l^k\right| \le c^{k-1}\sum_{i=1}^k d_i \le k\ c^{k-1} d_1 \ .$$

Proof

$$\begin{aligned}
\left|\prod_{i=1}^k l_i - l^k\right| &= \left|\prod_{i=1}^k l_i - \prod_{i=1}^{k-1} l_i l + \prod_{i=1}^{k-1} l_i l - \prod_{i=1}^{k-2} l_i l^2 + \prod_{i=1}^{k-2} l_i l^2 - \ldots - l^k\right| \\
&\le \sum_{i=1}^k \left(\prod_{j=1}^{k-i} l_j\right) |l_{k-i+1} - l|\ l^{i-1} \le c^{k-1} \sum_{i=1}^k d_i \\
&\le k\ c^{k-1} d_1 \ ,
\end{aligned}$$

where $c = \max\left(l_0, l\right)$. $\square$

Proof of Theorem 24 For $k=0$, we just note that $I\!P_A\left(S_A = 0\right) = 1 - p_1$ and $|1 - p_1 - (1-\rho(A))| \le d_1$. Suppose $k \ge 1$. Therefore

$$\begin{aligned}
& I\!P_A\left(S_A = k\right) \\
=\ & I\!P_A\left(\bigcap_{j=0}^k T^{-j\tau(A)} A \ \ ; \ \ T^{-(k+1)\tau(A)} A^c\right)
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{P}\left(T^{-(k+1)\tau(A)}A^c|\bigcap_{j=0}^{k}T^{-j\tau(A)}A\right)\prod_{i=1}^{k}\mathbb{P}\left(T^{-i\tau(A)}A|\bigcap_{j=0}^{i-1}T^{-j\tau(A)}A\right)\\
&= (1-p_{k+1})\prod_{i=1}^{k}p_i \ .
\end{aligned}
$$

Third equality follows by stationarity. Lemma 27 ends the proof of the theorem.
□

Proof of Corollary 26 We use the inequality

$$
\left|\mathbb{E}\left(X^\beta\right)-\mathbb{E}\left(Y^\beta\right)\right| \le \sum_{k\ge 0}k^\beta\left|\mathbb{P}\left(X=k\right)-\mathbb{P}\left(Y=k\right)\right| \ ,
$$

which holds for any pair of positive r.v. X, Y. We apply the above inequality with $X = S_A$ and Y geometrically distributed with parameter $\rho(A)$.

The exponential decay of the error term in Theorem 24 ends the proof of the corollary. □

Acknowledgements MA thanks Capes for support during part of this work. The authors thank P. Ferrari and A. Galves for useful discussions.

References

[1] M. Abadi. Exponential approximation for hitting times in mixing processes. *Math. Phys. Elec. J.* **7**, 2, (2001).

[2] M. Abadi. Sharp error terms and necessary conditions for exponential hitting times in mixing processes. *Ann. Probab.* **32**, 1A, (2004) 243-264.

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

[4] M. Abadi, N. Haydn and S. Vaienti. Statistics properties of repetition times. *Preprint.* (2005).

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

[6] J. R. Chazottes. Hitting and returning to non-rare events in mixing dynamical systems. *Nonlinearity* **16**, (2003) 1017-1034.

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

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

[9] W. Doeblin. Remarques sur la thorie mtrique des fractions continues. *Compositio Math.* **7**, (1940) 353-371.

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

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

[12] N. Haydn and S. Vaienti. The limiting distribution and error terms for return times of dynamical systems. *Discrete Contin. Dyn. Syst.***10**, 3, (2004) 589–616.

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

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

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

[16] D. Ornstein and B. Weiss. Entropy and data compression schemes. *IEEE Trans. Inform. Theory* **39**, 1, (1993) 78-83.

[17] P. Walters. Ruelle's operator theorem and g-measures. *Trans. Amer. Math. Soc.* **214**, (1975) 375-387.

[18] A. Wyner and J. Ziv. Some asymptotic properties of the entropy of a stationary ergodic data source with applications to data compression. *IEEE Trans. Inform. Theory* **35**, 6, (1989) 1250–1258.