



Sharp errors for point-wise Poisson approximations in mixing processes

Miguel Abadi, Nicolas Vergne

► **To cite this version:**

Miguel Abadi, Nicolas Vergne. Sharp errors for point-wise Poisson approximations in mixing processes. *Nonlinearity*, IOP Publishing, 2008, 21 (12), pp.2871-2885. 10.1088/0951-7715/21/12/008 . hal-02337199

HAL Id: hal-02337199

<https://hal-normandie-univ.archives-ouvertes.fr/hal-02337199>

Submitted on 30 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sharp errors for point-wise Poisson approximations in mixing processes

Miguel Abadi ^{*} Nicolas Vergne [†]

Abstract

We describe the statistics of the number of occurrences of a string of symbols in a stochastic process: Chosen a string A of length n , we prove that the number of visits to A up to time t , denoted by N_t , has approximately a Poisson distribution. We provide a sharp error for this approximation. Contrarily to previous works who present uniform error terms based on the total variation distance, our error is point-wise. As a byproduct we obtain that all the moments of N_t are finite. Moreover, we obtain explicit approximations for all of them. Our result holds for processes that verify the ϕ -mixing condition. The error term is explicitly expressed as function of the rate function ϕ and then easily computable.

AMS Subject class. Primary: 60F05; Secondary: 60G10, 60G55, 37A50.

Keyword: Mixing, recurrence, rare event, number of visits, Poisson distribution.

1 Introduction

This paper describes the statistics of occurrence times of a string of symbols in a mixing stochastic process with a finite alphabet. For $n \in \mathbb{N}$, we consider a fixed string of n symbols. We prove an upper bound for the difference between the law of the number of occurrences of the string in a long sequence and a Poisson law. Our result stands for ϕ -mixing processes (see definition below), with its corresponding error.

The first result about the number of visits to a fixed set is obviously the convergence of the binomial distribution to the Poisson distribution. Recently, motivated by the statistical analysis of data sources coming from different areas such as physics, biology, computer science, linguistics among other there was a major interest to generalize this convergence in various sense:

(a) dependent process;

^{*}IMECC, Universidade Estadual de Campinas, Pça Sérgio Buarque de Holanda 651 Cid. Univ. CP 6065, Cep. 13083-859, Campinas SP, Brazil. Phone +55-19 37886021 miguel@ime.unicamp.br

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

- (b) explicit rate of convergence;
- (c) different kind of observables.

There is abundant literature considering on the subject considering (a) in the dynamical systems and markovian contexts. See for instance Galves and Schmitt ([14]) and the references there in.

Probably the most used tool to attack (b) is the Chen-Stein method introduced by Chen ([11]). There is also abundant literature on this subject (see e.g. [6], [7], [8].) The principal feature of this method is that it provides only uniform bounds for the rate of convergence based on the total variation distance. As far as we know, this method was only implemented in processes that verify the Markov property. Whether it is useful in other context is an open question for us. We are aware of only one work which provides point-wise rate of convergence. Haydn and Vaienti ([15]) prove a rate of convergence using the method of factorial moments. The result holds for $(\psi - f)$ -mixing processes. The bound decreases factorially fast on k but contrary to our, it holds only for values of k that do not exceed the inverse of some (positive) power of the measure of the n -string.

Our result tends to give bring some light over (a), (b), and (c).

With respect to (b), we prove an upper bound for the rate of convergence of the number of occurrences of a fixed string to the Poisson law, namely,

$$\lim_{\mathbb{P}(A) \rightarrow 0} \mathbb{P}(N_t/\mathbb{P}(A) = k) = \frac{e^{-t} t^k}{k!} ,$$

where N_t is the number of visits of the process to the string A up to time t .

The striking point of our work tends to be the following. The error bound we obtain decreases factorially fast as a function of k for *all* values of k . This control on the tail of distribution of N_t allows us to obtain an approximation for *all the moments* of N_t by those of a Poisson random variable which are finite.

Our approach relies on a sharp result proved by Abadi ([1]) that states that for any string that does not overlap itself,

$$\mathbb{P}(N_t/\mathbb{P}(A) = 0) \approx e^{-t} .$$

A crucial point is that, if A is any string, $N_t/\mathbb{P}(A)$ could not be well approximated by a Poisson law. An example of this fact is shown in Hirata ([16]), where it is proved that for periodic points, the asymptotic limit law of $\{N_t/\mathbb{P}(A) = 1\}$ (as function of t) differs of the one-level Poisson law. When this happens, Abadi and Vergne ([5], Theorem 2) show that the law of τ_A is different from the exponential. Moreover, Theorem 24 in the same paper shows that A occurs in clumps with geometric size, which says that N_t is not Poisson distributed.

Our result is established with its own error term. This error is explicitly expressed as function of the mixing rate. As we said, it turns out that the error term depends on the overlapping properties of A . We state some basic facts about overlapping useful to prove our theorem. More on that topic can be found in [5].

With respect to (a), we establish our result under the mixing conditions. Mixing is a large family of processes. For instance, irreducible and aperiodic finite state Markov chain are known to be ψ -mixing (and then ϕ -mixing) with exponential decay. Moreover, Gibbs states which have summable variations are ψ -mixing (see [18]). They have exponential decay if they have Hölder continuous potential (see [9]). However, the ψ -mixing condition is sometimes a very condition, difficult to test. We establish our result under the more general ϕ -mixing condition. Further examples of ϕ -mixing processes can be found in [17]. The error term is explicitly expressed as a function of the mixing rate ϕ . We refer the reader to [13] and [10] for a source of examples and definitions of the several kinds of mixing processes. Those include ϕ -mixing and α -mixing with functions ϕ and α decreasing at any rate.

With respect to (c), since any observable can be constructed as a union of strings, we focus our work on them.

Our result is applied in a forthcoming paper: In [5] the authors applied the Poisson approximation to develop a method for testing hypothesis to detect strings of high or low frequency in DNA and protein sequences. This method can not work with approximation in total variation distance or any other uniform distributions distance.

This paper is organized as follows. In section 2 we establish our framework. In section 3 we collect some definitions and properties of overlapping of strings. In section 4 we state and prove the convergence of the number of occurrences to a Poisson law. This is Theorem 4.1.

2 Framework and notations

Let \mathcal{C} be a finite set. Put $\Omega = \mathcal{C}^{\mathbb{Z}}$. For each $x = (x_m)_{m \in \mathbb{Z}} \in \Omega$ and $m \in \mathbb{Z}$, let $X_m : \Omega \rightarrow \mathcal{C}$ be the m -th coordinate projection, that is $X_m(x) = x_m$. We denote by $T : \Omega \rightarrow \Omega$ 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

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

with $a_i \in \mathcal{C}$, $i = 0, \dots, n-1$.

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

For two measurables V and W , we denote as usual $\mathbb{P}_W(V) = \mathbb{P}(V|W) = \mathbb{P}(V; W) / \mathbb{P}(W)$ the conditional measure of V given W .

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

$$\phi(l) = \sup |\mathbb{P}(C | B) - \mathbb{P}(C)|,$$

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

We write $\mathbb{P}(V; W) = \mathbb{P}(V \cap W)$. We also write $V^c = \Omega \setminus V$, for the complement of V .

We use the probabilistic notation: $\{X_n^m = x_n^m\} = \{X_n = x_n, \dots, X_m = x_m\}$. For a n -string $A = \{X_0^{n-1} = x_0^{n-1}\}$ and $1 \leq w \leq n$, we define the w -string

$$A^{(w)} = \{X_{n-w}^{n-1} = x_{n-w}^{n-1}\}.$$

It belongs to the σ -algebra $\mathcal{F}_{\{n-w, \dots, n-1\}}$ and consisting of the *last* w symbols of A .

The mean of a r.v. X is 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. For brevity we put $(a \vee b) = \max\{a, b\}$ and $(a \wedge b) = \min\{a, b\}$.

3 Overlapping

In this section we describe some basic facts about overlapping of a string that are needed to establish our main result.

Definition 3.1 *Let $A \in \mathcal{C}_n$. We define the periodicity of A (with respect to T) as the number $\tau(A)$ defined as follows:*

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

Let us write $n = qp + r$, with $\tau(A) = p$, $q = \lfloor n/p \rfloor$ and $0 \leq r < p$. Thus

$$A = \left\{ X_0^{p-1} = X_p^{2p-1} = \dots = X_{(q-1)p}^{qp-1} = a_0^{p-1}; X_{qp}^{n-1} = a_0^{r-1} \right\}.$$

For instance

$$A = \left(\overbrace{\text{aaaabb}}^{\text{period}} \overbrace{\text{aaaabb}}^{\text{period}} \overbrace{\text{aaa}}^{\text{rest}} \right).$$

Thus, consider the set of overlapping positions of A :

$$\mathcal{O}(A) = \{k \in \{1, \dots, n-1\} \mid A \cap T^{-k}(A) \neq \emptyset\}.$$

Split $\mathcal{O}(A)$ in a disjoint union of $\{\tau(A), \dots, \lfloor n/\tau(A) \rfloor \tau(A)\}$ and $\mathcal{R}(A)$ where

$$\mathcal{R}(A) = \{k \in \{\lfloor n/\tau(A) \rfloor \tau(A) + 1, \dots, n-1\} \mid A \cap T^{-k}(A) \neq \emptyset\}.$$

Put $r_A = \#\mathcal{R}(A)$. The cardinal of $\mathcal{O}(A)$ is then $\sigma(A) = \lfloor n/\tau(A) \rfloor + r_A \leq n$.

4 Poisson approximation

4.1 Main result

For $1 \leq t' < t$ integers, let

$$N_{t'}^t = \sum_{i=t'}^t \mathbb{1}_{T^{-i}A}.$$

So that, $N_{t'}^t$ counts the number of occurrences of A between t' and t . For the sake of simplicity we write $N_t = N_1^t$. With some abuse of notation we also put $(-1)! = 1$.

Theorem 4.1 *Let $(X_m)_{m \in \mathbb{Z}}$ be a ϕ -mixing process. There exists a constant $C > 0$, such that for all $A \in \mathcal{C}_n$, and all non negative integer k , the following inequality holds:*

$$\left| \mathbb{P}(N_{t/\mathbb{P}(A)} = k) - \frac{e^{-t} t^k}{k!} \right| \leq C e(A) g(A, k),$$

with $e(A) = e_1(A) + e_2(A)$,

$$e_1(A) \stackrel{\text{def}}{=} \inf_{1 \leq w \leq \tau(A)} \left[(\sigma(A) + n) \mathbb{P}(A^{(w)}) + \phi((\tau(A)) - w) \right],$$

$$e_2(A) \stackrel{\text{def}}{=} \phi(n) + \inf_{n \leq \ell \leq 1/\mathbb{P}(A)} \left[\ell \mathbb{P}(A) + \frac{\phi(\ell)}{\mathbb{P}(A)} \right],$$

and

$$g(A, k) \stackrel{\text{def}}{=} \begin{cases} \frac{(2\lambda)^{k-1}}{(k-1)!} & k \notin \left\{ \frac{\lambda}{e(A)}, \dots, \frac{t}{\mathbb{P}(A)} \right\} \\ \frac{(2\lambda)^{k-1}}{\left(\frac{\lambda}{e(A)} \right)! \left(\frac{1}{e(A)} \right)^{k - \frac{1}{e(A)} - 1}} & k \in \left\{ \frac{\lambda}{e(A)}, \dots, \frac{t}{\mathbb{P}(A)} \right\} \end{cases},$$

where $\lambda \stackrel{\text{def}}{=} t \left[1 + \frac{\phi(\ell_A)}{\mathbb{P}(A)} \right]$ and ℓ_A is the ℓ that defines $e_2(A)$.

We state several remarks to better understand the error term of the theorem below the next corollaries.

In the next corollary we show how the point-wise error term given in Theorem 4.1 allows us to estimate the moments of $N_{t/\mathbb{P}(A)}$ by those of a r.v. with Poisson distribution.

Corollary 4.1 *Let $(X_m)_{m \in \mathbb{Z}}$ be a ϕ -mixing process with summable sequence ϕ . Let $\beta > 0$. Let Z be a r.v. with Poisson distribution of parameter $t > 0$. Then*

$$\left| \mathbb{E} \left(N_{t/\mathbb{P}(A)}^\beta \right) - \mathbb{E} (Z^\beta) \right| \leq C_{t,\beta} e(A),$$

where $C_{t,\beta}$ is a constant that just depends on t and β .

Corollary 4.2 *Let $(X_m)_{m \in \mathbb{Z}}$ be a ϕ -mixing process with summable sequence ϕ . Let Z be a r.v. with Poisson distribution of parameter $t > 0$. Then*

$$\sup_{K \subseteq \mathbb{N}} \left| \mathbb{P}(N_{t/\mathbb{P}(A)} \in K) - \mathbb{P}(Z \in K) \right| \leq C_t e(A),$$

where C_t is a constant that just depends on t .

Remark 4.1 Clearly $e(A)$ is the uniform error term and $g(A, k)$ is the error factor that provides the control on the tail of distribution.

Remark 4.2 $e_1(A)$ is the error that arises from the short correlations of the process while $e_2(A)$ is the error that arises from long ones.

Remark 4.3 $\mathbb{P}(A_n) \leq Ce^{-cn}$ (see [1]). $\phi(n)$ goes to zero by hypothesis. Therefore $e_1(A)$ is small if $\tau(A)$ is large enough to chose a w between 1 and $\tau(A)$ such that Ce^{-cw} and $\phi(\tau(A) - w)$ are small.

Remark 4.4 Take a sequence of n -strings A_n with n diverging. $e_1(A) \rightarrow 0$ if $\tau(A_n)$ also diverges with n faster than $\ln n$ (since $\mathbb{P}(A_n)$ decays exponentially fast).

Remark 4.5 $e_2(A) \rightarrow 0$ as $n \rightarrow \infty$ if the sequence $\phi(\ell)$ is summable.

Remark 4.6 Collet et al. ([12]) proved that for exponentially ψ -mixing processes there exist positive constants C and c such that

$$\mathbb{P}(A \in \mathcal{C}_n ; \tau(A) \leq n/3) \leq Ce^{-cn} .$$

Abadi ([1]) extended the above inequality to ϕ -mixing processes when $n/3$ is replaced with some $s \in (0, 1)$. Abadi and Vaienti ([3]) proved the above inequality for ψ -mixing processes for any value of s (with $c = c(s)$.) This shows that Theorem 4.1 holds for typical (in the sense of $\tau(A)$) strings. Taking limit on the length of the strings along infinite sequences, we get that the Poisson limit law holds almost everywhere.

Remark 4.7 When $\tau(A)$ is not large enough, the return time is better approximated by a mixture of a Dirac measure at the origin and an exponential law as shown by Abadi and Vergne ([5], Theorem 2). Therefore, the numbers of occurrences of the string can not be Poisson distributed.

Remark 4.8 When $e_2(A)$ is small, so is $\phi(\ell)/\mathbb{P}(A)$. Therefore λ is just the parameter of the Poisson law with a small correction factor $1 + \phi(\ell)/\mathbb{P}(A)$. Thus $\lambda/e(A)$ is a large number (smaller or equal to $t/\mathbb{P}(A)$.)

For $k \leq \lambda/e(A)$ or $k \geq t/\mathbb{P}(A)$ we get that $g(A, k)$ decays factorially fast. For k in the strip $\lambda/e(A)$ to $t/\mathbb{P}(A)$ we do not get $k!$ but something that we could call "truncated factorial": just get $(1/e(A))!$ times $k - (1/e(A))$ factors $1/e(A)$.

4.2 Examples

Example 4.1 Suppose that $(X_m)_{m \in \mathbb{Z}}$ are i.i.d. r.v. Then the process is ϕ -mixing with sequence $\phi(l) = 0$ for all $l \in \mathbb{N}$. Then $\ell_A = n$ and $e_2(A) = n\mathbb{P}(A)$. Further, take $w = \tau_A$. Thus $e_1(A) \leq 2n\mathbb{P}(A^{\tau_A})$. Thus $e(A) \leq 3n\mathbb{P}(A^{\tau_A})$. Here $\mathbb{P}(A^{\tau_A})$ is the probability of the part of the string A that does not overlap A . In particular, if A does not overlap itself, then $e(A) \leq 3n\mathbb{P}(A)$.

Example 4.2 Suppose that $(X_m)_{m \in \mathbb{Z}}$ is an irreducible and aperiodic finite state Markov chain. Then a classical theorem of Markov chains said that the process is ϕ -mixing and there are positive constants C and M such that

$$\phi(l) \leq Ce^{-Ml} \quad \text{for all } l \in \mathbb{N} .$$

We recall that the measure of n -cylinders decays exponentially fast on n . Thus, take for instance $\ell_A = K_1 n$ with K_1 a positive constant large enough to make $\phi(K_1 n)/\mathbb{P}(A)$ small. Assume that $\tau(A) = K_2 n$. Take $w = \tau(A)/2$. Thus $e_1(A)$ and $e_2(A)$ decay exponentially fast on n . In particular, if A does not overlap itself, (as is typically the case, see Remark 4.6) then $\sigma(A) = 0$. Thus

$$e_1(A) = \inf_{1 \leq w \leq n} \{n\mathbb{P}(A^{(w)}) + \phi(n-w)\} ,$$

and

$$e_2(A) = K_1 n \mathbb{P}(A) + \frac{\phi(K_1 n)}{\mathbb{P}(A)} + \phi(n) .$$

Example 4.3 Suppose that $(X_m)_{m \in \mathbb{Z}}$ is ϕ -mixing with polynomial sequence ϕ such that $\phi(l) = l^{-\kappa}$ for some $\kappa > 1$. Then $l = \mathbb{P}(A)^{-2/(\kappa+1)}$. Thus, the first term in $e_2(A)$ is $\mathbb{P}(A)^{(\kappa-1)/(\kappa+1)}$ which decays exponentially fast on n . So, $e_2(A)$ is of order $n^{-\kappa}$. Assume A does not overlap itself, namely $\tau(A) = n$. Take $w = K_1 n$. Thus the first term of $e_1(A)$ is exponential and $e_1(A)$ is of order $(K_2 n)^{-\kappa}$ given by the second term.

4.3 Preparatory results

The next lemma says that the occurrence of two copies of A very close have small probability.

Lemma 4.1 Let $(X_m)_{m \in \mathbb{Z}}$ be a ϕ -mixing process. Then, for all $A \in \mathcal{C}_n$ the following inequalities hold:

•

$$\mathbb{P}_A \left(\bigcup_{j=1}^{2n-1} T^{-j} A \right) \leq e_1(A) ,$$

• for all $\ell \geq 2n$

$$\mathbb{P}_A \left(\bigcup_{j=2n}^{\ell} T^{-j} A \right) \leq \ell \mathbb{P}(A) + \phi(n) .$$

Proof By the overlapping properties of A one has

$$A \cap \bigcup_{j=1}^{2n-1} T^{-j} A = A \cap \left(\bigcup_{j \in \mathcal{O}(A)} \bigcup_{j=n}^{2n-1} T^{-j} A \right) .$$

Now since $T^{-j}A \subseteq T^{-j}A^{(w)}$ for any $1 \leq w \leq \tau(A)$, the first part of the lemma follows using the ϕ -mixing property with $B = A$ and

$$C = \bigcup_{j \in \mathcal{O}(A)} \bigcup_{j=n}^{2n-1} T^{-j}A^{(w)} .$$

Namely

$$\mathbb{P}_A \left(\bigcup_{j=1}^{2n-1} T^{-j}A \right) \leq \mathbb{P} \left(\bigcup_{j \in \mathcal{O}(A)} \bigcup_{j=n}^{2n-1} T^{-j}A^{(w)} \right) + \phi(\tau(A) - w) .$$

The first statement of the lemma follows since the cardinal of the union is $\sigma(A) + n$. The cardinal of the union in the second statement of the lemma is $\ell - n + 1$. The second part of the lemma follows using the ϕ -mixing property as in the previous case. \square

Definition 4.1 Given $A \in \mathcal{C}_n$, and $j \in \mathbb{N}$, we define the j -occurrence time of A as the r.v. $\tau_A^{(j)} : \Omega \rightarrow \mathbb{N} \cup \{\infty\}$, defined as follows: For any $x \in \Omega$, define $\tau_A^{(1)}(x) = \inf\{k \geq 1 : T^k(x) \in A\}$ and for $j \geq 2$

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

The next proposition says that the measure of all the configurations where there are no two occurrences of A very close, is close to the product measure.

Proposition 4.1 Let $(X_m)_{m \in \mathbb{Z}}$ be a ϕ -mixing process. Then, for all $A \in \mathcal{C}_n$, all $0 \leq t_1 < t_2 < \dots < t_k \leq t$, and all $k \in \mathbb{N}$ for which

$$\min_{2 \leq j \leq k} \{t_j - t_{j-1}\} > 2(\ell_A + n) ,$$

(ℓ_A defined in Theorem 4.1) the following inequality holds:

$$\begin{aligned} & \left| \mathbb{P} \left(\bigcap_{j=1}^k \tau_A^{(j)} = t_j ; \tau_A^{(k+1)} > t \right) - \mathbb{P}(A)^k \prod_{j=1}^{k+1} \mathbb{P}(t_j - t_{j-1} - 2(\ell_A + n)) \right| \\ & \leq 5k (\mathbb{P}(A) + \phi(\ell_A))^k e(A) . \end{aligned}$$

Proof We prove the proposition by induction on k . For shorthand notation put $\bar{\ell}_A = 2(\ell_A + n)$, $\Delta_1 = t_1$, $\Delta_{k+1} = t - t_k$, $\Delta_i = t_i - t_{i-1}$ and $\mathcal{P}_i = \mathbb{P}(\tau_A > \Delta_i - \bar{\ell}_A)$; $i = 1, \dots, k+1$.

For $k = 1$, the triangle inequality gives

$$\left| \mathbb{P}(\tau_A = t_1 ; \tau_A^{(2)} > t) - \mathbb{P}(A) \prod_{j=1}^2 \mathcal{P}_j \right| \tag{4.1}$$

$$\leq \left| \mathbb{P}(\tau_A = t_1; \tau_A^{(2)} > t) - \mathbb{P}(\tau_A = t_1; N_{t_1 + \ell_A + n}^t = 0) \right| \quad (4.2)$$

$$+ \left| \mathbb{P}(\tau_A = t_1; N_{t_1 + \ell_A + n}^t = 0) - \mathbb{P}(\tau_A = t_1) \mathcal{P}_2 \right| \quad (4.3)$$

$$+ \left| \mathbb{P}(A; \tau_A > t_1 - 1) - \mathbb{P}(A; N_{n + \ell_A}^{t_1 - 1} = 0) \right| \mathcal{P}_2 \quad (4.4)$$

$$+ \left| \mathbb{P}(A; N_{n + \ell_A}^{t_1 - 1} = 0) \mathcal{P}_2 - \mathbb{P}(A) \prod_{j=1}^2 \mathcal{P}_j \right|. \quad (4.5)$$

In (4.4) we used that by stationarity $\mathbb{P}(\tau_A = t) = \mathbb{P}(A; \tau_A > t - 1)$. Term (4.2) is equal to

$$\begin{aligned} & \mathbb{P} \left(\tau_A = t_1; \bigcup_{i=t_1+1}^{t_1 + \ell_A + n - 1} T^{-i} A; N_{t_1 + \ell_A + n}^t = 0 \right) \\ & \leq \mathbb{P} \left(T^{-t_1} A; \bigcup_{i=t_1+1}^{t_1 + \ell_A + n - 1} T^{-i} A \right) \\ & = \mathbb{P} \left(A; \bigcup_{i=1}^{\ell_A + n - 1} T^{-i} A \right). \end{aligned} \quad (4.6)$$

We divide the above union in those sets with $1 \leq i < 2n$, and $2n \leq i \leq \ell_A + n$. Lemma 4.1 implies

$$\mathbb{P} \left(A; \bigcup_{i=1}^{2n-1} T^{-i} A \right) \leq \mathbb{P}(A) e_1(A).$$

and,

$$\mathbb{P} \left(A; \bigcup_{i=2n+1}^{\ell_A + n} T^{-i} A \right) \leq \mathbb{P}(A) (\ell_A \mathbb{P}(A) + \phi(n)).$$

Term (4.3) is bounded using the mixing property by $\phi(\ell_A) \mathbb{P}(A)$. Analogous computations are used to bound terms (4.4) and (4.5). This shows that (4.1) is bounded by $2\epsilon(A) \mathbb{P}(A)$.

Now let us suppose that the proposition holds for $k - 1$ and let us prove it for k . We use a triangle inequality where the terms involved are defined below. We briefly comment the idea behind each term. For brevity denote for each non negative i , $\mathcal{S}_i = \{\tau_A^{(i)} = t_i\}$. Thus we have

$$\left| \mathbb{P} \left(\bigcap_{j=1}^k \mathcal{S}_j; \tau_A^{(k+1)} > t \right) - \mathbb{P}(A)^k \prod_{j=1}^{k+1} \mathcal{P}_j \right| \leq I + II + III + IV + V.$$

In I we open a gap of length $\ell_A + n$ at the left of the k -th occurrence of A , namely, between coordinates $t_k - (\ell_A + n)$ and $t_k - 1$.

$$\begin{aligned}
I &\stackrel{def}{=} \left| \mathbb{P} \left(\bigcap_{j=1}^k \mathcal{S}_j; \tau_A^{(k+1)} > t \right) - \mathbb{P} \left(\bigcap_{j=1}^{k-1} \mathcal{S}_j; N_{t_{k-1}+1}^{t_k - (\ell_A + n)} = 0; T^{-t_k} A; N_{t_{k+1}}^t = 0 \right) \right| \\
&= \mathbb{P} \left(\bigcap_{j=1}^{k-1} \mathcal{S}_j; N_{t_{k-1}+1}^{t_k - (\ell_A + n)} = 0; \bigcup_{i=t_k - (\ell_A + n) + 1}^{t_k - 1} T^{-i} A; T^{-t_k} A; N_{t_{k+1}}^t = 0 \right) \quad (4.7) \\
&\leq \mathbb{P} \left(\bigcap_{j=1}^{k-1} T^{-t_j} A; \bigcup_{i=t_k - (\ell_A + n) + 1}^{t_k - 1} T^{-i} A; T^{-t_k} A \right).
\end{aligned}$$

As with (4.6) we split the above union in sets with $t_k - (\ell_A + n) + 1 \leq i \leq t_k - 2n$, $t_k - 2n + 1 \leq i \leq t_k - 1$. We recall that by hypothesis $\Delta_i > \ell_A$ for all $i = 1, \dots, k$. As in Lemma 4.1 we have for $t_k - (\ell_A + n) + 1 \leq i \leq t_k - 2n$

$$\begin{aligned}
&\mathbb{P} \left(\bigcap_{j=1}^{k-1} T^{-t_j} A; \bigcup_{i=t_k - (\ell_A + n) + 1}^{t_k - 2n} T^{-i} A; T^{-t_k} A \right) \\
&\leq \mathbb{P} \left(\bigcap_{j=1}^{k-1} T^{-t_j} A; \bigcup_{i=t_k - (\ell_A + n) + 1}^{t_k - 2n} T^{-i} A \right) (\mathbb{P}(A) + \phi(n)).
\end{aligned}$$

By the ϕ -mixing property over the left most factor in the right hand side of the above inequality, we get that it is bounded by

$$\mathbb{P} \left(\bigcap_{j=1}^{k-1} T^{-t_j} A \right) (\ell_A \mathbb{P}(A) + \phi(\ell_A)).$$

Iterating this procedure we get

$$\mathbb{P} \left(\bigcap_{j=1}^{k-1} T^{-t_j} A \right) \leq (\mathbb{P}(A) + \phi(\ell_A))^{k-1}.$$

Similarly, for $t_k - 2n + 1 \leq i \leq t_k - 1$

$$\mathbb{P} \left(\bigcap_{j=1}^{k-1} T^{-t_j} A; \bigcup_{i=t_k - 2n + 1}^{t_k - 1} T^{-i} A; T^{-t_k} A \right) \leq (\mathbb{P}(A) + \phi(\ell_A))^k e_1(A).$$

In II we apply the ϕ -mixing property to factorize the probability in the right hand side of the modulus in I . Then we iterated the ϕ -mixing property

to obtain the last inequality.

$$\begin{aligned}
II &\stackrel{def}{=} \left| \mathbb{P} \left(\left(\bigcap_{j=1}^{k-1} \mathcal{S}_j ; N_{t_{k-1}+1}^{t_k - (\ell_A + n)} = 0 \right) ; (T^{-t_k} A ; N_{t_k+1}^t = 0) \right) - \right. \\
&\quad \left. - \mathbb{P} \left(\bigcap_{j=1}^{k-1} \mathcal{S}_j ; N_{t_{k-1}+1}^{t_k - (\ell_A + n)} = 0 \right) \mathbb{P} (A ; N_1^{t-t_k} = 0) \right| \\
&\leq \mathbb{P} \left(\bigcap_{j=1}^{k-1} \mathcal{S}_j ; N_{t_{k-1}+1}^{t_k - (\ell_A + n)} = 0 \right) \phi(\ell_A) \\
&\leq \mathbb{P} \left(\bigcap_{j=1}^{k-1} T^{-t_j} A \right) \phi(\ell_A) \\
&\leq (\mathbb{P}(A) + \phi(\ell_A))^k \frac{\phi(\ell_A)}{\mathbb{P}(A)}.
\end{aligned}$$

In *III* we “fill-up” the gap we opened in *I*

$$\begin{aligned}
III &\stackrel{def}{=} \left| \mathbb{P} \left(\bigcap_{j=1}^{k-1} \mathcal{S}_j ; N_{t_{k-1}+1}^{t_k - (\ell_A + n)} = 0 \right) - \mathbb{P} \left(\bigcap_{j=1}^{k-1} \mathcal{S}_j ; N_{t_{k-1}+1}^{t_k - 1} = 0 \right) \right| \times \\
&\quad \times \mathbb{P} (A ; N_1^{t-t_k} = 0) \\
&\leq \mathbb{P} \left(\bigcap_{j=1}^{k-1} \mathcal{S}_j ; N_{t_{k-1}+1}^{t_k - (\ell_A + n)} = 0 ; \bigcup_{i=t_k - (\ell_A + n) + 1}^{t_k - 1} T^{-i} A \right) \mathbb{P}(A) \\
&\leq \mathbb{P} \left(\bigcap_{j=1}^{k-1} T^{-t_j} A ; \bigcup_{i=t_k - (\ell_A + n) + 1}^{t_k - 1} T^{-i} A \right) \mathbb{P}(A) \\
&\leq (\mathbb{P}(A) + \phi(\ell_A))^k 2^{\ell_A} \mathbb{P}(A).
\end{aligned}$$

In *IV* we use the inductive hypothesis

$$\begin{aligned}
IV &\stackrel{def}{=} \left| \mathbb{P} \left(\bigcap_{j=1}^{k-1} \mathcal{S}_j ; N_{t_{k-1}+1}^{t_k - 1} = 0 \right) - \mathbb{P}(A)^{k-1} \prod_{j=1}^k \mathcal{P}_j \right| \mathbb{P} (A ; N_1^{t-t_k} = 0) \\
&\leq C(k-1) (\mathbb{P}(A) + \phi(\ell_A))^{k-1} e(A) \mathbb{P}(A).
\end{aligned}$$

In *V* we use that the proposition is already proved for $k = 1$ to get

$$V \stackrel{def}{=} \mathbb{P}(A)^{k-1} \prod_{j=1}^k \mathcal{P}_j \left| \mathbb{P} (A ; N_1^{t-t_k} = 0) - \mathbb{P}(A) \mathcal{P}_{k+1} \right|$$

$$\leq \mathbb{P}(A)^k 2e(A) .$$

Summing the bounds above we end the proof of the proposition. \square

4.4 Proof of Theorem 4.1 and Corollary 4.1.

Proof of Theorem 4.1. Take $t \in \mathbb{N}$. Let us write for the sake of simplicity $N = N_t$. For $k = 0$ note that $\mathbb{P}(N = 0) = \mathbb{P}(\tau_A^{(1)} > t)$. By Theorem 1 in Abadi ([2]) one has

$$|\mathbb{P}(\tau_A^{(1)} > t) - e^{-\xi_A \mathbb{P}(A)t}| \leq e(A)(\mathbb{P}(A)t \vee 1)e^{-\xi_A \mathbb{P}(A)t} , \quad (4.8)$$

with a certain $\xi_A > 0$. Moreover, it follows in the proof of Theorem 2 in Abadi and Vergne ([4]) that $|\xi_A - \zeta_A| \leq e_1(A)$ where $\zeta_A = \mathbb{P}_A(\tau_A^{(1)} > \tau(A))$. Finally $|\zeta_A - 1| = \mathbb{P}_A(\zeta_A = \tau(A)) \leq e_1(A)$ by Lemma 4.1. This concludes the proof for $k = 0$.

For $k > t$ we have that $\mathbb{P}(N = k) = 0$. Then

$$\begin{aligned} \left| \mathbb{P}(N = k) - \frac{e^{-t\mathbb{P}(A)}(t\mathbb{P}(A))^k}{k!} \right| &= \frac{e^{-t\mathbb{P}(A)}(t\mathbb{P}(A))^k}{k!} \\ &\leq \frac{(t\mathbb{P}(A))^{k-1}}{(k-1)!} \mathbb{P}(A) . \end{aligned}$$

To conclude just note that $\mathbb{P}(A) \leq e(A)$.

Let us consider now k with $1 \leq k \leq t$. The idea of the proof is the following: Consider a realization $x = (x_m)_{m \in \mathbb{Z}}$ of the process $(X_m)_{m \in \mathbb{Z}}$ such that the sequence (x_1, \dots, x_t) contains exactly k occurrences of A . These occurrences can appear in clusters or isolated one from each other. We prove that realizations with isolated A 's give the approximation to the Poisson law and realizations with clustered A 's have small measure. We now formalize this idea. Given $1 \leq t_1 < \dots < t_k \leq t$, let us define the following measurable set:

$$\mathcal{T}(t_1, \dots, t_k) = \bigcap_{j=1}^k \left\{ \tau_A^{(j)} = t_j \right\} \cap \left\{ \tau_A^{(k+1)} > t \right\} .$$

As in Proposition 4.1 we put $\Delta_j = t_j - t_{j-1}$, for $j = 2, \dots, k$. Put also $\Delta_1 = t_1$ and $\Delta_{k+1} = t - t_k$. Define the minimum distance between two consecutive occurrences of A by

$$I(\mathcal{T}(t_1, \dots, t_k)) = \min \{ \Delta_j \mid 2 \leq j \leq k \} .$$

As before put $\bar{\ell}_A = 2(\ell_A + n)$. Let us divide $\{N = k\}$ in two sets

$$B_k = \bigcup_{I(\mathcal{T}(t_1, \dots, t_k)) < \bar{\ell}_A} \mathcal{T}(t_1, \dots, t_k) \quad \text{and} \quad G_k = \bigcup_{I(\mathcal{T}(t_1, \dots, t_k)) \geq \bar{\ell}_A} \mathcal{T}(t_1, \dots, t_k) .$$

Since $\{N = k\} = B_k \cup G_k$, disjoint union, we have

$$\begin{aligned} & \left| \mathbb{P}(N = k) - \frac{e^{-t\mathbb{P}(A)} t^k \mathbb{P}(A)^k}{k!} \right| \\ & \leq \mathbb{P}(B_k) + \left| \mathbb{P}(G_k) - \frac{e^{-t\mathbb{P}(A)} t^k \mathbb{P}(A)^k}{k!} \right|. \end{aligned} \quad (4.9)$$

We will prove that both quantities in the right hand side of (4.9) are small.

Proof: configurations with clusters have small measure.

We will prove an upper bound for $\mathbb{P}(B_k)$. Let us start computing how many clusters there are in a given $\mathcal{T}(t_1, \dots, t_k)$ with

$$C(\mathcal{T}(t_1, \dots, t_k)) = \sum_{j=2}^k \mathbb{1}_{\{\Delta_j > \bar{\ell}_A\}} + 1.$$

Suppose that $C(\mathcal{T}(t_1, \dots, t_k)) = 1$ and fix the position t_1 . Each occurrence inside the unique cluster (with the exception of the most left one which is fixed at t_1) can appear at distance d of the previous one, with $d \in \mathcal{O}(A)$ or $n \leq d \leq \bar{\ell}_A$. Firstly note that

$$\mathcal{T}(t_1, t_2, \dots, t_k) \subseteq \bigcap_{j=1}^k T^{-t_j} A.$$

Then

$$\bigcup_{\substack{i=2, \dots, k \\ t_i = t_{i-1}, \dots, t_{i-1} + \bar{\ell}_A}} \mathcal{T}(t_1, t_2, \dots, t_k) \subseteq \bigcup_{\substack{i=2, \dots, k \\ t_i = t_{i-1}, \dots, t_{i-1} + \bar{\ell}_A}} \bigcap_{j=1}^k T^{-t_j} A.$$

Therefore, the iterative argument of the ϕ -mixing property used to bound (4.7) leads to the bound

$$\begin{aligned} & \mathbb{P} \left(\bigcup_{\substack{i=2, \dots, k \\ t_i = t_{i-1}, \dots, t_{i-1} + \bar{\ell}_A}} \bigcap_{j=1}^k T^{-t_j} A \right) \\ & \leq \mathbb{P}(A) (e_1(A) + \bar{\ell}_A \mathbb{P}(A) + \phi(n))^{k-1} \\ & \leq \mathbb{P}(A) e(A)^{k-1}. \end{aligned} \quad (4.10)$$

Suppose now that $C(\mathcal{T}(t_1, \dots, t_k)) = i$. Assume also that the most left occurrence of these i clusters occurs at $1 \leq t(1) < \dots < t(i) \leq t$ fixed. By the same argument used in (4.10), we have the inequalities

$$\mathbb{P} \left(\bigcup_{\{t_1, \dots, t_k\} \setminus \{t(1), \dots, t(i)\}} \mathcal{T}(t_1, \dots, t_k) \right)$$

$$\begin{aligned}
&\leq \mathbb{P}(A) (\mathbb{P}(A) + \phi(\ell_A))^{i-1} e(A)^{k-i} \\
&\leq (\mathbb{P}(A) + \phi(\ell_A))^i e(A)^{k-i} .
\end{aligned} \tag{4.11}$$

To obtain an upper bound for $\mathbb{P}(B_k)$ we must sum the above bound over all $\mathcal{T}(t_1, \dots, t_k)$ such that $C(\mathcal{T}(t_1, \dots, t_k)) = i$ with i that runs from 1 to $k-1$.

Fixed $C(\mathcal{T}(t_1, \dots, t_k)) = i$, the locations of the most left occurrences of A of each one of the i clusters can be chosen at most in $\binom{t}{i}$ many ways.

The cardinality of each one of the i clusters can be arranged in $\binom{k-1}{i-1}$ many ways. (This corresponds to break the interval $(1/2, k+1/2)$ in i intervals at points chosen from $\{1+1/2, \dots, k-1/2\}$.)

Collecting these information and (4.11) we have that $\mathbb{P}(B_k)$ is bounded by

$$\sum_{i=1}^{k-1} \binom{t}{i} \binom{k-1}{i-1} (\mathbb{P}(A) + \phi(\ell_A))^i e(A)^{k-i} \leq e(A)^k \max_{1 \leq i \leq k-1} \left\{ \frac{\gamma^i}{i!} \right\} \sum_{i=1}^{k-1} \binom{k-1}{i-1},$$

where $\gamma = t\mathbb{P}(A) [1 + \phi(\ell_A)/\mathbb{P}(A)] / e(A)$. The maximum in the above expression is reached at $(k-1 \wedge \gamma)$. The most right sum is bounded by 2^{k-1} . Therefore we have

$$\mathbb{P}(B_k) \leq e(A) \cdot \begin{cases} \frac{(2\gamma e(A))^{k-1}}{(k-1)!} & k-1 < \gamma \\ \frac{2^{k-1} (\gamma e(A))^\gamma}{\gamma! \left(\frac{1}{e(A)}\right)^{k-\gamma-1}} & k \geq \gamma \end{cases} .$$

This ends the proof of the bound for $\mathbb{P}(B_k)$.

Proof: A 's isolated provide the Poisson limit law.

We can bound the most right term on the right-hand side of (4.9) by the following triangular inequality:

$$\sum_{\mathcal{T}(t_1, \dots, t_k) \in G_k} \left| \mathbb{P} \left(\bigcap_{j=1}^k \tau_A^{(j)} = t_j ; \tau_A^{(k+1)} > t \right) - \mathbb{P}(A)^k \prod_{j=1}^{k+1} \mathcal{P}_j \right| \tag{4.12}$$

$$+ \mathbb{P}(A)^k \sum_{\mathcal{T}(t_1, \dots, t_k) \in G_k} \left| \prod_{j=1}^{k+1} \mathcal{P}_j - \prod_{j=1}^{k+1} e^{-(\Delta_j - \bar{\ell}_A) \mathbb{P}(A)} \right| \tag{4.13}$$

$$+ \mathbb{P}(A)^k \#G_k \left| e^{-(t-(k+1)\bar{\ell}_A) \mathbb{P}(A)} - e^{-t\mathbb{P}(A)} \right| \tag{4.14}$$

$$+ \left| \frac{\#G_k k!}{t^k} - 1 \right| \frac{e^{-t\mathbb{P}(A)} t^k \mathbb{P}(A)^k}{k!} . \tag{4.15}$$

By a simple combinatorial argument we get the bounds

$$\frac{(t - k(n + \bar{\ell}_A))^k}{k!} \leq \binom{t - k(n + \bar{\ell}_A - 1) - 1}{k} \leq \#G_k \leq \binom{t}{k} \leq t^k / k! . \tag{4.16}$$

Moreover, the leading term in (4.12) is bounded using Proposition 4.1. Thus (4.12) is bounded by

$$5 \frac{t^k}{(k-1)!} (\mathbb{P}(A) + \phi(\ell_A))^k e(A) .$$

The difference between the leading factors in (4.13) is bounded as follows: again by (4.8)

$$|\mathcal{P}_j - e^{-\xi_A \mathbb{P}(A)(\Delta_j - \bar{\ell}_A)}| \leq C e_1(A) .$$

As stated at the beginning of the proof one has $|\xi_A - 1| \leq e_1(A)$. Therefore (4.13) is bounded by

$$\frac{t^k}{k!} \mathbb{P}(A)^k (k+1) \max_{1 \leq j \leq k+1} |\mathcal{P}_j - e^{-(\Delta_j - \bar{\ell}_A) \mathbb{P}(A)}| \leq \frac{k+1}{k} \frac{(t \mathbb{P}(A))^k}{(k-1)!} C e_1(A) .$$

(4.14) is bounded using the Mean Value Theorem by

$$\frac{t^k \mathbb{P}(A)^k}{k!} (k+1) \bar{\ell}_A \mathbb{P}(A) \leq \frac{k+1}{k} \frac{(t \mathbb{P}(A))^k}{(k-1)!} 4 \ell_A \mathbb{P}(A) .$$

The left hand side of (4.16) and the Mean Value Theorem provide a bound for the difference below

$$\left| \frac{\#G_k}{t^k} - 1 \right| \leq \left| \frac{(t - k(n + \bar{\ell}_A))^k}{t^k} - 1 \right| \leq \frac{k}{t} \frac{k(n + \bar{\ell}_A)}{t} \leq k .$$

So, (4.15) is bounded by

$$\frac{(t \mathbb{P}(A))^k}{(k-1)!} 4 \ell_A \mathbb{P}(A) .$$

Summing the bounds obtained for (4.12), (4.13), (4.14) and (4.15) we get the desired bound for the difference in the right hand term of inequality (4.9). The exchange of variables $\tilde{t} = t \mathbb{P}(A)$ ends the proof of the theorem. \square

Proof of Corollary 4.1. By definition

$$\begin{aligned} \left| \mathbb{E} \left(N_{t/\mathbb{P}(A)}^\beta \right) - \mathbb{E}(Z^\beta) \right| &= \left| \sum_{k \geq 0} k^\beta \mathbb{P}(N_{t/\mathbb{P}(A)} = k) - \sum_{k \geq 0} k^\beta \frac{e^{-t} t^k}{k!} \right| \\ &\leq \sum_{k \geq 0} k^\beta \left| \mathbb{P}(N_{t/\mathbb{P}(A)} = k) - \frac{e^{-t} t^k}{k!} \right| . \end{aligned}$$

Since for ϕ summable

$$\sum_{k=0}^{\infty} k^\beta g(A, k) \leq C_{t,\beta} < \infty ,$$

the corollary follows. \square

Proof of Corollary 4.2. This follows by the above corollary with $\beta = 1$. \square

Acknowledgments MA is partially supported by CNPq grant 308250/2006-0. The authors are beneficiaries of Capes, Brasil - Cofecub, France grant. The authors thank P. Ferrari, A. Galves, B. Prum and P. Shields for useful discussions.

References

- [1] Abadi, M. (2001). Exponential approximation for hitting times in mixing processes. *Math. Phys. Elec. J.* **7**, 2.
- [2] Abadi, M. (2001). Exponential approximation for hitting times in mixing processes. *Ann. Probab.* **32**, 1A, 243-264.
- [3] Abadi, M., and Vaienti, S. (2007). Large deviation for short recurrence. *To appear in Discrete and Continuous Dynamical Systems*.
- [4] Abadi, M., and Vergne, N. (2006). Sharp error terms for return time statistics under mixing conditions. *Submitted. Technical report Imecc-Unicamp*.
- [5] Abadi, M., and Vergne, N. (2006). Poisson approximation for search of rare words in DNA sequences. *Submitted. Technical report Imecc-Unicamp*.
- [6] Arratia, R., Goldstein, L., and Gordon, L. (1989). Two moments suffice for Poisson approximation: the Chen-Stein method. *Ann. Probab.* **17**, 9 - 25.
- [7] Arratia, R., Goldstein, L., and Gordon, L. (1990). Poisson approximation and the Chen-Stein method. *Stat. Science*, **5** 4. 403 - 434.
- [8] Barbour, A.D., Holst, L., and Janson, S. (1992). Poisson approximation. *Oxford, University Press, New York*.
- [9] Bowen, R. (1975). Equilibrium states and the ergodic theory of Anosov diffeomorphisms. *Lecture Notes in Math*, **470**. Springer-Verlag, New York.
- [10] Bradley, R. C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probability Surveys* **2**, 107-144.
- [11] Chen, L. H. Y. (1975). Poisson approximation for dependent trials. *Ann. Prob.* **3**, 534-545.
- [12] Collet, P., Galves, A., and Schmitt, B. (1999). Repetition times for gibbsian sources. *Nonlinearity* **12**, 1225-1237.

- [13] Doukhan, D. (1995). Mixing. Properties and examples. *Lecture Notes in Statistics* **85**, Springer-Verlag.
- [14] Galves, A., and Schmitt, B. (1997). Inequalities for hitting times in mixing dynamical systems. *Random Comput. Dyn.* **5**, 337-348.
- [15] Haydn, N., and Vaienti, S. (2004). The limiting distribution and error terms for return times of dynamical systems. *Discrete Contin. Dyn. Syst.***10**, 3, 589–616.
- [16] Hirata, M. (1993). Poisson law for Axiom A diffeomorphism. *Ergod. Th. Dyn. Sys.* **13**, 533-556.
- [17] Liverani, C., Saussol, B., and Vaienti, S. (1998). Conformal measures and decay of correlations for covering weighted systems. *Ergod. Theor. dynam. Sys.* **18** 1399-420.
- [18] Walters, P. (1975). Ruelle's operator theorem and g -measures. *Trans. Amer. Math. Soc.* **214**, 375-387.