RETURN TIMES DISTRIBUTION OF EXPANDING MAPS

Nicolai T A Haydn[1]

Abstract

We consider expanding systems with invariant measures that are uniformly expanding everywhere except on a small measure set and show that the limiting statistics of hitting times for zero measure sets are compound Poisson provided the limits for the cluster size distributions exist. This extends previous results from neighbourhoods around single points to neighbourhoods around zero measure sets. The assumptions require the decay of correlations to the at least polynomial and some decay condition on the non-uniformly expanding and and some regularity conditions around the limiting zero measure target set.

1. Introduction

Limiting return times distributions have been studied already by Doeblin for the Gauss map, but serious broad interest developed only around 1990 in particular with a paper by Pitskel [21] where he showed, using generating functions, that for Axiom A maps the limiting distribution at almost all points are Poissonian if the shrinking target sets are cylinder sets. He also found that at periodic points the return times also have a geometric component which gives rise to a compound Poisson distribution. A similar result was shown in [17] for the entry time using the Laplace transform and the argued to generalise to extend to Poisson distributed limiting higher return times using the weak mixing property. For harmonic maps such a results were shown in [5] using the Chen-Stein method. Later, Galves and Schmitt [12] came up with a very effective method to show exponentially distributed entry times in a symbolic setting for ψ-mixing measures and also provided error terms. This method was then carried further by Abadi and others (e.g. [1, 13, 3, 2]) where it was shown that for symbolic systems the limiting return times are almost everywhere Poisson distributed. Similar results for metric balls for non-uniformly hyperbolic systems where shown in [4, 16, 20] even providing speeds of convergence outside a small set. A closer look at periodic orbits was done in [14] where it was found that in a symbolic setting one always gets a compound Poisson distrution. Later there was found to be a dichotomy between non-periodic and periodic points where non-periodic points always have Poisson distributed hitting times and periodic points lead to compound Poisson distributions (see e.g. [10, 7, 8] and also [9]). In [15] such concepts were generalised to the situation when the limiting set is not a single point any longer but can be some arbitrary null set. In this case the limiting compound Poisson distribution can be quite general and is not any longer restricted to be Pólya-Aeppli. In that instance it was used to get limiting results on synchonised systems. Here we show

[1]Department of Mathematics, University of Southern California, Los Angeles, 90089-2532. E-mail: nhaydn@usc.edu.

a similar argument for expanding maps only which we then apply to interval maps and look in particular at the limiting hitting times distribution at the parabolic point where we recover a result of [11] where an ad hoc method was used to get the right scaling which is different from the standard Kac scaling and which follows in a natural way from our setup.

2. Expanding Maps and Limiting Distributions

Let Ω be a metric space and $T : \Omega \to \Omega$ be a measurable map and μ a T-invariant probability measure. Here we assume that T is differentiable and expanding, i.e. $|Dt(x)| \geq 1$ for all x. If $U \subset \Omega$ is a subset then put

$$Z_U^N = \sum_{j=0}^{N-1} \mathbb{1}_U \circ T^j$$

for the hit-counting function over the time interval N. In particular, if $\Gamma \subset \Omega$ has zero μ-measure and for $\rho > 0$ we denote by $B_\rho(\Gamma) = \bigcup_{x\in\Gamma} B_\rho(x)$ its ρ-neighbourhood, then for suitable orbit lengths $N(\rho)$ we want to get that the distribution of $\zeta_{B_\rho(\Gamma)}^{N(\rho)}$ converges to a non-degenerate probability distribution as $\rho \to 0$. In the classical setting Γ is usually chosen to be a simple point x and for N one then choses the Kac scaling $t/\mu(B_\rho(x))$ where > 0 is a parameter. In that case it is known that for a variety of systems the limiting distribution is a compound Poisson distribution. Here we present a general scheme for expanding maps to obtain such results which might also result in a non-standard scaling for $N(\rho)$.

Let us recall that an integer valued random variable W is compound Poisson distributed if there are i.i.d. $\mathbb{N}$-valued random variables $Y_j \geq 1$, $j = 1, 2, \ldots$, and an independent Poisson distributed random variable P so that $W = \sum_{j=1}^{P} Y_j$. The Poisson distribution P describes the distribution of what usually is referred to as clusters whose sizes are then described by the values of the random variables Y_j whose probability densities are given by $\lambda_\ell = \mathbb{P}(Y_j = \ell)$, $\ell = 1, 2, \ldots$. In particular

$$\mathbb{P}(W = k) = \sum_{\ell=1}^{k} \mathbb{P}(P = \ell)\mathbb{P}(S_\ell = k),$$

where $S_\ell = \sum_{j=1}^{\ell} Y_j$ and P is Poisson distributed with parameter t, i.e. $\mathbb{P}(P = \ell) = e^{-t}t^\ell/\ell!$, and by Wald's equation $\mathbb{E}(W) = t\mathbb{E}(Y_j)$. We say a probability measure $\tilde{\nu}$ on $\mathbb{N}_0$ is compound Poisson distributed with parameters t and λ_ℓ, $\ell = 1, 2, \ldots$ if it has the same distribution as W.

In the special case $Y_1 = 1$ and $\lambda_\ell = 0 \forall \ell \geq 2$ we recover the Poisson distribution $W = P$.

An important non-trivial compound Poisson distribution is the Pólya-Aeppli distribution which happens when the Y_j are geometrically distributed, that is $\lambda_\ell = \mathbb{P}(Y_\ell) = (1-\vartheta)\vartheta^{\ell-1}$ for $\ell = 1, 2, \ldots$, for some $\vartheta \in (0, 1)$. In this case

$$\mathbb{P}(W = k) = e^{-t} \sum_{j=1}^{k} \vartheta^{k-j}(1-\vartheta)^j \frac{s^j}{j!}\binom{k-1}{j-1}$$

and in particular $\mathbb{P}(W=0)=e^{-t}$. In the case of $p=0$ this reverts back to the straight Poisson distribution.

3. Assumptions and main result

Assume there exist $R>0$ and for every $n\in\mathbb{N}$ finitely many $y_k\in\Omega$ so that $\Omega\subset\bigcup_k B_R(y_k)$, where $B_R(y)$ is the R-disk centered at y. Denote by $\zeta_{\varphi,k}=\varphi(B_R(y_k))$ where $\varphi\in\mathscr{I}_n$ and $\mathscr{I}_n$ denotes the inverse branches of T^n. We call ζ an n-cylinder. Then there exists a constant L so that the number of overlaps $N_{\varphi,k}=|\{\zeta_{\varphi',k'}:\zeta_{\varphi,k}\cap\zeta_{\varphi',k'}\neq\varnothing,\varphi'\in\mathscr{I}_n\}|$ is bounded by L for all $\varphi\in\mathscr{I}_n$ and for all k and n. This follows from the fact that $N_{\varphi,k}$ equals $|\{k':B_R(y_k)\cap B_R(y_{k'})\neq\varnothing\}|$ which is uniformly bounded by some constant L.

Let us denote by $J_n=\frac{dT^n\mu}{d\mu}$ the Jacobian of the map T^n with respect to the measure μ. Also put $Z_U^L=\sum_{j=0}^{L-1}\mathbb{1}_U\circ T^j$.

For some $L\in\mathbb{N}$ put $Z_U^L=\sum_{j=0}^{L-1}\mathbb{1}_U\circ T^j$ for the hit counting function over a time interval of length L. Let $\Gamma\subset\Omega$ be a zero measure set and for $\rho>0$ denote by $B_\rho(\Gamma)$ its ρ-ball neighbourhood. Then we shall make the following assumptions:
(I) There exists a decay function $\mathcal{C}(k)$ so that

$$\left|\int_M G(H\circ T^k)\,d\mu-\mu(G)\mu(H)\right|\le\mathcal{C}(k)\|G\|_{Lip}\|H\|_\infty\qquad\forall k\in\mathbb{N},$$

for every $H\in L^\infty(\Omega,\mathbb{R})$ for every $G\in Lip(\Omega,\mathbb{R})$.
Then, we need some geometric assumptions:
(II) We assume that there is a set $\mathcal{G}_n\subset\{\zeta_\varphi:\varphi\in\mathscr{I}_n\}$ so that
(i) $\mu(G_n^c)\lesssim n^{-\mathfrak{g}}$ for some $\mathfrak{g}>0$, where $G_n=\bigcup_{\varphi\in\mathcal{G}_n}\zeta_\varphi$.
(ii) (Distortion) We require that $\frac{J_n(x)}{J_n(y)}=\mathcal{O}(\mathfrak{D}(n))$ for all $x,y\in\zeta$ for all $\zeta\in\mathcal{G}_n$ n-cylinders and all n, where $\mathfrak{D}$ is a non-decreasing function which below we assume to be $\mathfrak{D}(n)=\mathcal{O}(n^{\mathfrak{d}})$ for some $\mathfrak{d}\ge0$.
(iii) (Contraction) There exists a function $\delta(n)\to0$ which decays at least summably polynomially, i.e. $\delta(n)=\mathcal{O}(n^{-\mathfrak{k}})$ with $\mathfrak{k}>1$, so that $\operatorname{diam}\zeta\le\delta(n)$ for all n-cylinder $\zeta\subset\mathcal{G}_n$ and all n.
(iv)

$$\mathfrak{G}_{\rho,L}=\sum_{n=L}^{\infty}\frac{\mu(G_n^c\cap B_\rho(\Gamma)\cap T^{-n}\mathcal{V}_\rho^L)}{\mu(\mathcal{V}_\rho^L)}\longrightarrow0$$

as $\rho\to0,L\to\infty$, where $\mathcal{V}_\rho^L=\{x\in\Omega:Z_{B_\rho(\Gamma)}^L(x)\ge1\}$.
(III) (Dimension estimate) There exist $0<d_0<d_1<\infty$ such that $\rho^{d_1}\le\mu(B_\rho)\le\rho^{d_0}$.
(IV) (Annulus condition) Assume that for some $\xi\ge\beta>0$:

$$\frac{\mu(B_{\rho+r}\setminus B_{\rho-r})}{\mu(B_\rho)}=\mathcal{O}(\frac{r^\xi}{\rho^\beta})$$

for every $r<\rho$.

Here and in the following we use the notation $x_n\lesssim y_n$ for $n=1,2,\dots$, to mean that there exists a constant C so that $x_n<Cy_n$ for all n. As before let $T:\Omega\circlearrowleft$ and μ a

T-invariant probability measure on Ω. For a subset $U \subset \Omega$ we put $X_i = \mathbb{1}_U \circ T^i$ and define

$$Z^L = Z^L_U = \sum_{i=0}^{L-1} X_i$$

where L is a (large) positive integer. For $\Gamma \subset \Omega$ of zero measure we put

$$\lambda_\ell = \lim_{L\to\infty} \lambda_\ell(L), \tag{1}$$

where

$$\lambda_\ell(L) = \lim_{\rho\to 0} \frac{\mathbb{P}(Z^L_{B_\rho(\Gamma)} = \ell)}{\mathbb{P}(Z^L_{B_\rho(\Gamma)} \geq 1)}.$$

Let us now formulate our main result.

Theorem 3.1. *Assume that the map $T : \Omega \to \Omega$ satisfies the assumptions (I)–(IV) where $\mathcal{C}(k)$ decays at least polynomially with power $\mathfrak{p} > (\frac{\beta}{\eta} + d_1)\frac{1+\mathfrak{d}}{d_0}$ and $\mathfrak{k} > \frac{1+\mathfrak{d}}{d_0}$. Let $\Gamma \subset \Omega$ be a zero measure set and λ_ℓ the corresponding quantity as defined in (1).*

Then

$$\lim_{L\to\infty} \lim_{\rho\to 0} \mathbb{P}(\xi^t_{B_\rho(\Gamma)} = k) = \nu(\{k\}),$$

where ν is the compound Poisson distribution for the parameters t, λ_ℓ and $N = N(L,\rho) = \frac{tL}{\mathbb{P}(Z^L_{B_\rho(\Gamma)} \geq 1)}$.

The proof of Theorem 3.1 is given in Section 6. In the following section we will express the parameters λ_ℓ in terms of the limiting return times distribution.

4. Return times anc Kac's scaling

As before let $U \subset \Omega$ be a subset of Ω so that $\mu(U) > 0$, then define the first entry/return time τ_U by $\tau_U(x) = \min\{j \geq 1 : T^j \in U\}$ which by Poincaré's recurrence theorem is finite almost everywhere. Similarly we get higher order returns by defining recursively $\tau^\ell_U(x) = \tau^{\ell-1}_U + \tau_U(T^{\tau^{\ell-1}_U}(x))$ with $\tau^1_U = \tau_U$. We also write $\tau^0_U = 0$ on U. For L some large number we then put $\alpha_\ell(L, U) = \mu_U(\tau^{\ell-1}_U \leq L < \tau^\ell_U)$ for $\ell = 1, 2, \ldots$, where μ_U is the induced measure on U given by $\mu_U(A) = \mu(A \cap U)/\mu(U), \forall A \subset \Omega$.

Now let $U_n \subset \Omega$, $n = 1, 2, \ldots$, be a nested sequence of sets and put $\Lambda = \bigcap_n U_n$. Assume the limits $\alpha_\ell(L) = \lim_{n\to\infty} \alpha_\ell(L, U_n)$, $\ell = 1, 2, \ldots$, exist for all L large enough. Similarly we put $\alpha_\ell = \lim_{L\to\infty} \alpha_\ell(L)$, for $\ell = 1, 2, \ldots$. In the special case $\ell = 1$ we get in particular $\alpha_1 = \lim_{L\to\infty} \lim_{n\to\infty} \mu_{U_n}(L < \tau_{U_n})$ which is also called the extremal index.

Lemma 4.1. [15] *Assume $\sum_\ell \ell^2 \alpha_\ell < \infty$. If $\alpha_1 > 0$ then*

$$\lambda_k = \frac{\alpha_k - \alpha_{k+1}}{\alpha_1}.$$

In particular the limit defining λ_k exists.

This lemma in particular implies that the expected length of the clusters is given by

$$\sum_{k=1}^\infty k\lambda_k = \frac{1}{\alpha_1} \sum_{k=1}^\infty k(\alpha_k - \alpha_{k+1}) = \frac{1}{\alpha_1}$$

provided α_1 is positive. Also notice that since $\lambda_k \geq 0$ one obtains that $\alpha_1 \geq \alpha_2 \geq \alpha_3 \geq \cdots$ is a decreasing sequence. Moreover $\lambda_k = \alpha_k \forall k$ only when both are geometrically distributed, i.e. $\lambda_k = \alpha_k = \alpha_1(1-\alpha_1)^k$ which results in a Polya-Aeppli compound Poisson distribution.

Let us consider the entry time $\tau_U(x)$ where $x \in \Omega$.

Lemma 4.2. *For $U \subset \Omega$ one has*

$$\mathbb{P}(\tau_U < L) = \mu(U) \sum_{j=0}^{L} \alpha_1(k, U).$$

If moreover $U_n \subset \Omega$ is a nested sequence so that $\mu(U_n) \to 0$ as $n \to \infty$ and so that the limit $\alpha_1 = \lim_{L\to\infty} \lim_{n\to\infty} \alpha_1(L, U_n)$ exists and satisfies $\alpha_1 > 0$, then

$$\lim_{L\to\infty} \lim_{n\to\infty} \frac{\mathbb{P}(\tau_{U_n} \leq L)}{L\mu(U_n)} = \alpha_1.$$

Proof. We have

$$\begin{aligned}
\mathbb{P}(\tau_U < L) &= \mathbb{P}(Z^L \geq 1) \\
&= \sum_{j=0}^{L-1} \mathbb{P}(Z^j \geq 1, T^{-j}U, \tau_U \circ T^j \geq L - j) \\
&= \sum_{j=0}^{L-1} \mathbb{P}(U, \tau_U \geq L - j) \\
&= \mu(U) \sum_{j=0}^{L} \alpha_1(k, U).
\end{aligned}$$

The second part of the statement now follows if $\alpha_1 > 0$. □

Remark 4.1. *If $\alpha_1 > 0$ then it now follows from the two previuous lemmata and its proof that*

$$\lim_{L\to\infty} \lim_{n\to\infty} \frac{\mathbb{P}(\tau^\ell_{U_n} \leq L < \tau^{\ell+1}_{U_n})}{L\mu(U_n)} = \alpha_1 \lambda_\ell$$

for $\ell = 1, 2, 3, \ldots$. In a similar way as in the previous lemma on can show for $\ell = 2, 3, \ldots$ that

$$\mathbb{P}(\tau^\ell_{U_n} \leq L) = \sum_{k=\ell}^{L} \mathbb{P}(\tau^k_{U_n} \leq L < \tau^{k+1}_{U_n}) = \mathbb{P}(\tau_{U_n} \leq L) \sum_{k=\ell}^{L} \lambda_k(L, U_n)$$

which implies as before that

$$\lim_{L\to\infty} \lim_{n\to\infty} \frac{\mathbb{P}(\tau^\ell_{U_n} \leq L)}{L\mu(U_n)} = \alpha_\ell.$$

5. The Compound Binomial Approximation

In this section we prove an approximation theorem that provides an estimate how closely the level sets of the counting function W is approximated by a compound binomial distribution which represents the independent case. As the measure of the approximating target set $B_\rho(\Gamma)$ goes to zero, the compound binomial distribution then converges to a compound Poisson distribution.

To be more precise, the following abstract approximation theorem which establishes the distance between sums of $\{0,1\}$-valued dependent random variables X_n and a random variable that has a compound Binomial distribution is used in Section 6.1 in the proof of Theorem 1 to compare the number of occurrences in a finite time interval with the number of occurrences in the same interval for a compound binomial process.

Let Y_j be $\mathbb{N}$ valued i.i.d. random variables and denote $\lambda_\ell = \mathbb{P}(Y_j = \ell)$. Let N' be a (large) positive integer, $t > 0$ a parameter and put $p = t/N'$. If Q is a binomially distributed random variable with parameters (N', p), that is $\mathbb{P}(Q = k) = \binom{N'}{k} p^k (1-p)^{N'-k}$, then $W = \sum_{i=1}^{Q} Y_i$ is compound binomially distributed. As N' goes to infinity, Q converges to a Poisson distribution with parameter t and W converges to a compound Poisson distribution with parameters t, λ_ℓ.

Let $(X_n)_{n\in\mathbb{N}}$ be a stationary $\{0,1\}$-valued process and put $Z^L = \sum_{i=0}^{L-1} X_i$ for $L \in \mathbb{N}$. Let $W_a^b = Z^b - Z^a$ for $0 \le a < b$ and $W = Z^N$. (In the following theorem we assume for simplicity's sake that N' and Δ are integers.)

Theorem 5.1. [15] *Let $L \ll N$ and denote by $\tilde{\nu}$ be the compound binomial distribution measure where the binomial part has values $p = \mathbb{P}(Z^L \ge 1)$ and $N' = N/L$ and the compound part has probabilities $\lambda_\ell = \mathbb{P}(Z^L = \ell)/p$.*

Then there exists a constant C_1, independent of L and $\Delta < L$, such that

$$|\,\mathbb{P}(W = k) - \tilde{\nu}(\{k\})\,| \le C_1(N'(\mathcal{R}_1 + \mathcal{R}_2) + \Delta\mathbb{P}(Z^L \ge 1)),$$

where

$$\mathcal{R}_1 = \sup_{\substack{0<\Delta<M\le N' \\ 0<q<N'-\Delta-1/2}} \left|\sum_{u=1}^{q-1}\Big(\mathbb{P}\big(Z^L = u \wedge W_{\Delta L}^{ML} = q-u\big) - \mathbb{P}(Z^L = u)\mathbb{P}\big(W_{\Delta L}^{ML} = q-u\big)\Big)\right|$$

$$\mathcal{R}_2 = \sum_{n=2}^{\Delta} \mathbb{P}(Z^L \ge 1 \wedge W_{nL}^{(n+1)L} \ge 1).$$

6. Proof of Theorem 1

In this section we bound the quantities in the assumption of Theorem 5.1 in the usual way by making a distinction between short interactions, i.e. those that are limited by a gap of length Δ, and long interactions which constitute the principal part. The near independence of long interactions is expressed bv the decay of correlations and gives rise to the error term $\mathcal{R}_1$. The short interactions are estimated by $\mathcal{R}_2$ and use the assumptions on limited distortion, the fact that 'cylinders' are pull-backs of uniformly sized balls and the positivity of the local dimension.

6.1. **Compound binomial approximation of the return times distribution.** To prove Theorem 1 we will employ the approximation theorem from Section 5 where we put $U = B_\rho(\Gamma)$. Let $X_i = \mathbb{1}_U \circ T^{i-1}$ and for the blocking argument let L be an integer and put as before $V_a^b = \sum_{j=a}^b Z_j$, where the $Z_j = \sum_{i=jL}^{(j+1)L-1} X_i$ are stationary random variables. Then we put $N = [tL/\mathbb{P}_\mu(Z_0 \geq 1)]$, where t is a positive parameter (from now on we omit the integer part brackets $[\cdot]$). Then for any $2 \leq \Delta \leq N' = N/L = t/\mathbb{P}_\mu(Z_0 \geq 1)]$ (for simplicity's sake we assume N is a multiple of L)

$$\left|\mathbb{P}(V_0^{N'} = k) - \tilde{\nu}(\{k\})\right| \leq C_3(N'(\mathcal{R}_1 + \mathcal{R}_2) + \Delta\mu(Z_0 \geq 1)), \tag{2}$$

where

$$\mathcal{R}_1 = \sup_{\substack{0<\Delta<M\leq N' \\ 0<q<N'-\Delta-1/2}} \left|\sum_{u=1}^{q-1}\big(\mathbb{P}\big(Z_0 = u \wedge V_\Delta^M = q-u\big) - \mathbb{P}(Z_0 = u)\mathbb{P}\big(V_\Delta^M = q-u\big)\big)\right|$$

$$\mathcal{R}_2 = \sum_{j=1}^{\Delta} \mathbb{P}(Z_1 \geq 1 \wedge Z_j \geq 1),$$

and $\tilde{\nu}$ is the compound binomial distribution with parameters $p = \mathbb{P}(Z_j \geq 1)$ and distribution $\frac{t}{p}\mathbb{P}(Z_j = k)$. Notice that $\mathbb{P}(V_0^{N'} = k) = 0$ for $k > N$ and also $\tilde{\nu}(\{k\}) = \mathbb{P}(\tilde{V}_0^{N'} = k) = 0$ for $k > N$.

We now proceed to estimate the error between the distribution of S and a compound binomial based on Theorem 5.1.

6.2. **Estimating $\mathcal{R}_1$.** Let us fix ρ for the moment and put $U = B_\rho(\Gamma)$. Fix q and u and we want to estimate the quantity

$$\mathcal{R}_1(q, u) = \big|\mathbb{P}\big(Z_0 = u, V_\Delta^M = q-u\big) - \mathbb{P}(Z_0 = u)\mathbb{P}\big(V_\Delta^M = q-u\big)\big|$$

In order to use the decay of correlations (II) to obtain an estimate for $\mathcal{R}_1(q, u)$ we approximate $\mathbb{1}_{Z_0=u}$ by Lipschitz functions from above and below as follows. Let $r > 0$ be small ($r \ll \rho$) and put $U''(r) = B_r(U)$ for the outer approximation of U and $U'(r) = (B_r(U^c))^c$ for the inner approximation. We then consider the set $\mathcal{U} = \{Z_0 = u\}$ which is a disjoint union of sets

$$\bigcap_{j=1}^{u} T^{-v_j}U \cap \bigcap_{i\in[0,L)\setminus\{v_j:j\}} T^{-i}U^c$$

where $0 \leq v_1 < v_2 < \cdots < v_u \leq L-1$ the u entry times vary over all possibilities. Similarly we get its outer approximation $\mathcal{U}''(r)$ and its inner approximation $\mathcal{U}'(r)$ by using $U''(r)$ and $U'(r)$ respectively. We now consider Lipschitz continuous functions approximating $\mathbb{1}_\mathcal{U}$ as follows

$$\phi_r(x) = \begin{cases} 1 & \text{on } \mathcal{U} \\ 0 & \text{outside } \mathcal{U}''(r) \end{cases} \qquad \text{and} \qquad \hat{\phi}_r(x) = \begin{cases} 1 & \text{on } \mathcal{U}'(r) \\ 0 & \text{outside } \mathcal{U} \end{cases}$$

with both linear in between. The Lipschitz norms of both ϕ_r and $\hat{\phi}_r$ are bounded by a^L/r where $a = \sup_{x\in\mathcal{G}} |DT(x)|$. By design $\hat{\phi}_r \le \mathbb{1}_{Z_0=u} \le \phi_r$. Moreover let us note that since

$$\mathcal{U}''(r) \setminus \mathcal{U}'(r) \subset \bigcup_{i=0}^{L-1} T^{-i}(B_{\rho+r}(\Gamma) \setminus B_{\rho-r}(\Gamma))$$

one has by Assumption (IV)

$$\mu(\mathcal{U}''(r) \setminus \mathcal{U}'(r)) \lesssim L\frac{r^\xi}{\rho^\beta}.$$

In order to use the decay of correlation let us approximate as follows:

$$\begin{aligned}\mathbb{P}\big(Z_0 = u, V_\Delta^M = q-u\big) &- \mathbb{P}(Z_0 = u)\mathbb{P}\big(V_\Delta^M = q-u\big)\\ &\le \int_M \phi_r \,\cdot \mathbb{1}_{V_\Delta^M = q-u}\, d\mu - \int_M \mathbb{1}_{Z_0=u}\, d\mu \int_M \mathbb{1}_{V_\Delta^M=q-u}\, d\mu\\ &= \mathbb{X} + \mathbb{Y}\end{aligned}$$

where

$$\mathbb{X} = \left(\int_M \phi_r\, d\mu - \int_M \mathbb{1}_{Z_0=u}\, d\mu\right) \int_M \mathbb{1}_{V_\Delta^M=q-u}\, d\mu$$
$$\mathbb{Y} = \int_M \phi_r\, (\mathbb{1}_{V_\Delta^M=q-u})\, d\mu - \int_M \phi_r\, d\mu \int_M \mathbb{1}_{V_\Delta^M=q-u}\, d\mu.$$

The two terms $\mathbb{X}$ and $\mathbb{Y}$ are estimated separately. The first term is readily estimated by:

$$\mathbb{X} \le \mathbb{P}(V_\Delta^M = q-u) \int_M (\phi_r - \mathbb{1}_{Z_0=u})\, d\mu \le \mu(\mathcal{U}''(r) \setminus \mathcal{U}(r)) \lesssim L\frac{r^\xi}{\rho^\beta}\mu(\mathcal{U}) = L\rho^{\xi w-\beta}\mu(\mathcal{U})$$

if we put $r = \rho^w$ for some $w > 1$. In order to estimate the second term $\mathbb{Y}$ we use the decay of correlations. Using assumption (II), we can estimate as follows:

$$\begin{aligned}\mathbb{Y} &= \left|\int_M \phi_r\, T^{-\Delta}(\mathbb{1}_{V_0^{M-\Delta}=q-u})\, d\mu - \int_M \phi_r\, d\mu \int_M \mathbb{1}_{V_0^{M-\Delta}=q-u}\, d\mu\right|\\ &\le \mathcal{C}(\Delta)\|\phi_r\|_{Lip}\end{aligned}$$

as $\|\mathbb{1}_{V_0^{M-\Delta}=q-u}\|_{\mathscr{L}^\infty} = 1$. Hence

$$\begin{aligned}\mu(\mathcal{U} \cap T^{-\Delta}\{V_0^{M-\Delta} = q-u\}) &- \mu(\mathcal{U})\,\mathbb{P}(V_0^{M-\Delta} = q-u)\\ &\lesssim a^L\frac{\mathcal{C}(\Delta)}{r} + L\rho^{\xi w-\beta}\mu(\mathcal{U})\end{aligned}$$

Since we get similar estimates using the inner approximation $\hat{\phi}_r$ we obtain

$$\mathcal{R}_1 \lesssim a^L\frac{\mathcal{C}(\Delta)}{\rho^w} + L\rho^{\xi w-\beta}\mathbb{P}(Z_0 \ge 1)$$

using the fact that $\mathcal{U} = \{Z_0 = u\} \subset \{Z_0 \ge 1\}$ for any $u \ge 1$.

In the exponential case when $\delta(n) = \mathcal{O}(\vartheta^n)$, for some $\vartheta < 1$, we choose $\Delta = \mathfrak{t}|\log\rho|$ for some $\mathfrak{t} > w/|\,\vartheta\,|$ and obtain the estimate

$$\mathcal{R}_1 \lesssim a^L\rho^{\mathfrak{t}|\,\vartheta\,|-w} + L\rho^{\xi w-\beta}\mathbb{P}(Z_0 \ge 1).$$

6.3. **Estimating the terms $\mathcal{R}_2$.** To estimate the contributions made by short return times we use, as it is typically done, expansiveness and the bounds on distortion. In order to estimate the measure of $U \cap T^{-j}U$ for some positive j define

$$\mathscr{C}_j(U) = \{\zeta_{\varphi,j} : \zeta_{\varphi,j} \cap U \neq \varnothing, \varphi \in \mathscr{I}_j\}$$

for the cluster of j-cylinders that cover the set U. If we put $\mathcal{V} = \{Z_0 \geq 1\}$, then

$$\begin{aligned} \mu(T^{-j}\mathcal{V} \cap U \cap G_j) &\leq \sum_{\zeta \in \mathscr{C}_j(U) \cap \mathcal{G}_j} \frac{\mu(T^{-j}\mathcal{V} \cap \zeta)}{\mu(\zeta)} \mu(\zeta) \\ &\lesssim \sum_{\zeta \in \mathscr{C}_j(U) \cap \mathcal{G}_j} \mathfrak{D}(j) \frac{\mu(\mathcal{V} \cap T^j\zeta)}{\mu(T^j\zeta)} \mu(\zeta) \end{aligned}$$

Since the sets $\zeta_{\varphi,k}$ are φ-pre-images of R-balls, the denominator is uniformly bounded from below because $\mu(T^j\zeta) = \mu(B_R(y_k))$ Thus, by assumption (I),

$$\mu(T^{-j}\mathcal{V} \cap U \cap G_j) \lesssim \mathfrak{D}(j)\mu(\mathcal{V}) \sum_{\zeta \in \mathscr{C}_j(U) \cap \mathcal{G}_j} \mu(\zeta) \lesssim \mathfrak{D}(j)\mu(\mathcal{V})\, L\, \mu\left(\bigcup_{\zeta \in \mathscr{C}_j(U) \cap \mathcal{G}_j} \zeta\right)$$

Now, since diam $\zeta \leq \delta(j) \lesssim j^{-\mathfrak{k}}$ for $\zeta \in \mathcal{G}_n$, in the polynomial case, one has

$$\bigcup_{\zeta \in \mathscr{C}_j(U) \cap \mathcal{G}_j} \zeta \subset B_{\delta(j)}(U).$$

Since by assumption $\mu(B_{\delta(j)}(U)) = \mathcal{O}((\delta(j) + \rho)^{d_0})$ we get in the polynomial case

$$\mu(T^{-j}\mathcal{V} \cap U \cap G_j) \lesssim \mathfrak{D}(j)\mu(\mathcal{V})(\delta(j)^{d_0} + \rho^{d_0}) \lesssim \mathfrak{D}(j)\mu(\mathcal{V})(j^{-\mathfrak{k}d_0} + \rho^{d_0}).$$

On the set G_j^c we use Assumption (II-iv) and obtain

$$\mu(T^{-j}\mathcal{V} \cap U \cap G_j^c) \leq \mu(\mathcal{V}) \frac{\mu(U \cap T^{-j}\mathcal{V} \cap G_j^c)}{\mu(\mathcal{V})}$$

For the estimate of $\mathcal{R}_2$ there are two cases to consider, namely (I) if $j \geq 2$ where we have a gap of length L to give us some dacay, and (II) when $j = 1$ in which case we have to open a gap to achieve some decay.

(I) If $j \geq 2$ then we use the decomposition

$$\{Z_0 \geq 1, Z_j \geq 1\} = \mathcal{V} \cap T^{-jL}\mathcal{V} = T^{-jL}\mathcal{V} \cap \bigcup_{k=0}^{L-1} T^{-k}U$$

and obtain

$$\mathbb{P}(Z_0 \geq 1, Z_j \geq 1) \leq \sum_{k=0}^{L-1} \mu(T^{-k}U \cap T^{-jL}\mathcal{V}) = \sum_{u=(j-1)L}^{jL-1} \mu(U \cap T^{-u}\mathcal{V}).$$

Consequently

$$\sum_{j=2}^{\Delta} \mathbb{P}(Z_0 \geq 1 \wedge Z_j \geq 1) \quad \leq \quad \sum_{u=L}^{\Delta L-1} \mu(U \cap T^{-u}\mathcal{V})$$

$$\begin{aligned} &\lesssim \mu(\mathcal{V})\sum_{u=L}^{\Delta L-1}\left(\mathfrak{D}(u)(u^{-\mathfrak{k}d_0}+\rho^{d_0})+\frac{\mu(U\cap T^{-u}\mathcal{V}\cap G_u^c)}{\mu(\mathcal{V})}\right)\\ &\lesssim \mu(\mathcal{V})\big(L^{-\sigma}+(L\Delta)^{1+\mathfrak{d}}\rho^{d_0}+\mathfrak{G}_L\big)\end{aligned}$$

since $\mathfrak{D}(u)=\mathcal{O}(u^{\mathfrak{d}})$, provided $\sigma=\mathfrak{k}d_0-\mathfrak{d}-1$ is larger than 0, where

$$\mathfrak{G}_L=\sum_{u=L}^{\infty}\frac{\mu(U\cap T^{-u}\mathcal{V}\cap G_u^c)}{\mu(\mathcal{V})}$$

goes to zero as $L\to\infty$ by Assumption (II-iv).

(II) If $j=1$ let $\alpha=\frac{1}{1+\sigma}$ and put $Z_0'=\sum_{i=L-L'+1}^{L-1}X_i$ and $Z_0''=Z_0-Z_0'$, where $L'=L^\alpha$. Then

$$\mathbb{P}(Z_0\ge1,Z_1\ge1)\le\mathbb{P}(Z_0''\ge1,Z_1\ge1)+\mathbb{P}(Z_0'\ge1),$$

where $\mathbb{P}(Z_0'\ge1)=\mu(\mathcal{V}')$, where $\mathcal{V}'=\{Z_0'\ge1\}$. Similar to the case (I) above we have now a gap of length L' which allows us to estimate in the same way

$$\begin{aligned}\mathbb{P}(Z_0''\ge1,Z_1\ge1)&\le\sum_{u=L'}^{L-1}\mu(U\cap T^{-u}\mathcal{V})\\ &\lesssim\mu(\mathcal{V})\big(L'^{-\sigma}+L^{1+\mathfrak{d}}\rho^{d_0}+\mathfrak{G}_{L'}\big)\end{aligned}$$

we conclude that

$$\mathbb{P}(Z_0\ge1,Z_1\ge1)\lesssim\mu(\mathcal{V})(L^{-\sigma\alpha}+L^{1+\mathfrak{d}}\rho^{d_0}+\mathfrak{G}_{L'})+\mu(\mathcal{V}').$$

Finally if we combine steps (I) and (II) then the entire error term can be estimated by

$$\begin{aligned}N'\mathcal{R}_2&\le N'\sum_{j=1}^{\Delta}\mathbb{P}(Z_0\ge1,Z_j\ge1)\\ &\lesssim N'\mu(\mathcal{V})(L^{-\alpha\sigma}+(L\Delta)^{1+\mathfrak{d}}\rho^{d_0}+\mathfrak{G}_{L^\alpha})+N'\mu(\mathcal{V}')\\ &\lesssim t\big(L^{-\alpha\sigma}+(L\Delta)^{1+\mathfrak{d}}\rho^{d_0}+\mathfrak{G}_{L^\alpha}\big)+N'\mu(\mathcal{V}')\\ &\lesssim L^{-\sigma\alpha}+L^{1+\mathfrak{d}}\rho^{v'}+\mathfrak{G}_{L^\alpha}+\frac{\mu(\mathcal{V}')}{\mu(\mathcal{V})}\end{aligned}$$

assuming $v'=d_0-v(1+\mathfrak{d})>0$ (as $\Delta=\rho^{-v}$), as $N'=t/\mu(\mathcal{V})$.

In the exponential case when $\delta(n)=\mathcal{O}(\vartheta^n)$, for some $\vartheta<1$, then we put $\Delta=\mathfrak{t}|\log\rho|$, for a suitable $\mathfrak{t}$, and obtain

$$\begin{aligned}\sum_{j=2}^{\Delta}\mathbb{P}(Z_0\ge1\wedge Z_j\ge1)&\lesssim\mu(\mathcal{V})\sum_{u=L}^{\Delta L-1}\left(u^{\mathfrak{d}}(\vartheta^{ud_0}+\rho^{d_0})+\frac{\mu(U\cap T^{-u}\mathcal{V}\cap G_u^c)}{\mu(\mathcal{V})}\right)\\ &\lesssim\mu(\mathcal{V})\Big(\tilde\vartheta^L+L^{1+\mathfrak{d}}\rho^{\tilde d_0}+\mathfrak{G}_L\Big)\end{aligned}$$

for any $\tilde\vartheta\in(\vartheta,1)$ and any positive $\tilde d_0<d_0$. This yields

$$N'\mathcal{R}_2\lesssim\tilde\vartheta^{L^\alpha}+L^{1+\mathfrak{d}}\rho^{\tilde d_0}+\mathfrak{G}_{L^\alpha}+\frac{\mu(\mathcal{V}')}{\mu(\mathcal{V})}.$$

6.4. **The total error.** For the total error we now put $r = \rho^w$ and as above $\Delta = \rho^{-v}$ where $v < d_0$ since $\Delta \ll N$ and $N \geq \rho^{-d_0}$. Moreover $L' = L^\alpha$ for $\alpha = 1/(1+\sigma)$ and in the polynomial case when $\mathcal{C}(\Delta) = \mathcal{O}(\Delta^{-p}) = \mathcal{O}(\rho^{pv})$ we get

$$\begin{aligned}
&|\mathbb{P}(W=k) - \tilde{\nu}(\{k\})| \\
&\quad\lesssim\; N'\left(a^L \frac{\mathcal{C}(\Delta)}{\rho^w} + L\rho^{\xi w-\beta}\mu(\mathcal{U})\right) + L^{-\alpha\sigma} + L^{1+\mathfrak{d}}\rho^{v'} + \mathfrak{G}_{L^\alpha} + \frac{\mu(\mathcal{V}')}{\mu(\mathcal{V})} + \Delta\mu(\tau_U \leq L) \\
&\quad\lesssim\; \frac{a^L}{L}\rho^{v\mathfrak{p}-w-d_1} + \rho^{w\xi-\beta} + L^{1+\mathfrak{d}}\rho^{v'} + L^{-\alpha\sigma} + \mathfrak{G}_{L^\alpha} + \frac{\mu(\mathcal{V}')}{\mu(\mathcal{V})} + L\rho^{d_0-v}
\end{aligned}$$

as $N'\mu(U) = \frac{s}{L}$, $s = N'\mathbb{P}(Z_0 \geq 1)$, $N = t/\mathbb{P}(Z_0 \geq 1) \gtrsim \rho^{d_1}$ and $\Delta\mu(\tau_U \leq L) \lesssim \rho^{-v+}\rho^{d_0}$. When $\rho \to 0$ then $\mu(U) \to 0$ and in order to get convergence we require $v\mathfrak{p} - w - d_1 > 0$, $w\xi - \beta > 0$ and $v' = d_0 - v(1+\mathfrak{d}) > 0$. This can be achieved if $w > \beta/\xi$ is sufficiently close to β/ξ and $\mathfrak{p} > \left(\frac{\beta}{\xi} + d_1\right)\frac{1+\mathfrak{d}}{d_0}$ in the case when $\mathcal{C}$ decays polynomially with power $\mathfrak{p}$, i.e. $\mathcal{C}(k) \sim k^{-\mathfrak{p}}$. These choices also satisfy $v < d_0$. Since we also must have $\sigma = \mathfrak{k}d_0 - \mathfrak{d} - 1 > 0$ this is satisfied if $\mathfrak{k} > \frac{\mathfrak{d}+1}{d_0}$.

In the exponential case ($\operatorname{diam}\zeta = \mathcal{O}(\vartheta^n)$ for n cylinders ζ and $\mathcal{C}(\Delta) \sim \vartheta^\Delta$) we obtain with $\Delta = s|\log\rho|$ for s large enough

$$|\mathbb{P}(W=k) - \tilde{\nu}(\{k\})| \lesssim a^L\rho^{s|\log\vartheta|-w-d_1} + \rho^{ws-\beta} + L^{1+\mathfrak{d}}|\log\rho|^{1+\mathfrak{d}} + \tilde{\vartheta}^{L^\alpha} + \frac{\mu(\mathcal{V}')}{\mu(\mathcal{V})} + \mathfrak{G}_{L^\alpha} + \Delta\mu(\tau_U \leq L),$$

for some $\tilde{\vartheta} \in (\vartheta, 1)$. Note that in this case we only need to have $d_0 > 0$.

6.5. **Convergence to the compound Poisson distribution.** For $t > 0$ and any $L \in \mathbb{N}$ large we take $N' = t/\mathbb{P}(Z^L \geq 1)$ In the double limit first ρ goes to zero and then we let L go to infinity. Denote by $\tilde{\nu}_{L,\rho}$ be the compound binomial distribution with the paramters $p = \mathbb{P}(Z^L \geq 1)$ and $N' = t/p$. Then, as $\rho \to 0$ the compound binomial distribution $\tilde{\nu}_{L,\rho}$ converges to the compound Poisson distribution $\tilde{\nu}_L$ for the parameters $t\lambda_\ell(L)$. Thus for every L:

$$\mathbb{P}(W=k) \longrightarrow \tilde{\nu}_L(\{k\}) + \mathcal{O}(L^{-\alpha\sigma} + \mathfrak{G}_{L^\alpha}).$$

Now let $L \to \infty$. Then $\lambda_\ell(L) \to \lambda_\ell$ for all $\ell = 1, 2, \ldots$ and $\tilde{\nu}_L$ converges to the compound Poisson distribution ν for the parameters $\lambda_\ell = \lim_{L\to\infty}\lambda_\ell(L)$. Finally we obtain

$$\mathbb{P}(W=k) \longrightarrow \nu(\{k\})$$

as $\rho \to 0$. This concludes the proof of Theorem 3.1. □

7. Examples

7.1. C^2 **interval maps.** Let $T : I \to I$ is piecewise expanding on the interval I. We assume that T is piecewise C^2 with uniformly bounded C^2-norms. For $\varphi \in \mathscr{I}_n$ ($\mathscr{I}_n$ are, as before, the set of inverse branches of T^n) we denote by $\zeta_\varphi = \varphi(I)$ the n-cylinder associated with φ. If, as before,

$$\delta(n) = \max_{\varphi\in\mathscr{I}_n}|\zeta_\varphi|$$

then $\delta(n)$ decays exponentially fast since we assume that $|T'(x)| > c_0$ uniformly in x for a constant $c_0 > 1$. That is $\delta(n) \leq c_0^{-n}$ (his corresponds to the case when $\mathfrak{k} = \infty$).

Moreover one has decay of the annealed correlation function (I) and also the decay of the quenched correlation functions (II). In fact, the decay function $\lambda(n)$ decays exponentially fast to zero (this is equivalent to $\mathfrak{p} = \infty$).

Since the measure μ is absolutely continuous with respect to Lebesgue measure and a density h which is bounded and bounded away from 0, Condition (V) is satisfied with any values $d_0 < 1 < d_1$ arbitrarily close to 1. Condition (III) follows from the uniform boundedness of second order derivatives which implies that $\mathfrak{d} = 0$. Let us notice that $\mathcal{G}_j^c = \varnothing$ for all j and that the contraction rate of the maximal size of n-cylinders $\delta(n)$ is exponential. The Annulus Condition (VI) is satisfied with $\xi = \beta = 1$.

By [15], Lemma 4, x is a periodic point if and only if the $\tau_\rho(x) = \inf\{j \geq 1 : T^j B_\rho(x) \cap B_\rho(x) \neq \varnothing\}$. Note that $\tau_\rho(x) \leq \tau_{\rho'}(x)$ if $\rho > \rho' > 0$ and if $\rho \to 0$ then $\tau_\rho(x)$ converges to the minimal period of m of x if x is a periodic point and to ∞ if x is not periodic. If x is periodic with minimal period m, then $\lambda_\ell = (1-\vartheta)\vartheta^{\ell-1}$, where

$$\vartheta = \lim_{\rho \to 0} \frac{\mu(B_\rho(x) \cap T^\ell B_\rho(x))}{\mu(B_\rho(x))} = |DT^m(x)|^{-1}$$

is the Pitskel value of x. If x is not periodic then we find that $\lambda_1 = 1$ and $\lambda_\ell = 0$ for all $\ell \geq 2$ which follows from the fact that $\tau_\rho(x)$ then diverges to ∞.

We can therefore invoke Theorem 3.1 and obtain the following result:

Theorem 7.1. *Let the map T be piecewise C^2 with uniformly bounded C^2 derivative and uniformly expanding. Then*

$$W_{x,\rho} = \sum_{j=0}^{N-1} \mathbb{1}_{B_\rho(x)} \circ T^j$$

with respect to the measure μ is:
(i) Poisson(t) if x is non-periodic,
(ii) Pólya-Aeppli if x is periodic with minimal period m and the Pitskel value $\vartheta = |DT^m(x)|^{-1}$, $\lambda_\ell = (1-\vartheta)\vartheta^{\ell-1}$,
where $N = N_\rho(x) = \frac{t}{\mu(B_\rho(x))}$ (Kac scaling).

7.2. **Parabolic interval maps.** Let us consider the Pomeau-Manneville map which for a parameter $\alpha \in (0,1)$ is given by

$$T(x) = \begin{cases} x + 2^\alpha x^{1+\alpha} & \text{if } x \in [0, \frac{1}{2}) \\ 2x - 1 & \text{if } x \in [\frac{1}{2}, 1] \end{cases}$$

and has a parabolic point at $x = 0$, i.e. $T'(0) = 1$. Otherwise $T'(x) > 1$ for all $x \neq 0$. These maps have a neutral (parabolic) fixed point at $x = 0$ and are otherwise expanding. It is known that T has an invariant absolutely continuous probability measure with density $h(x)$, where $h(x) \sim x^{-\alpha}$ for x close to 0.

There exists a constant C so that

$$\left| \int \psi(\phi \circ T^n)\, d\mu - \int \psi\, d\mu \int \phi\, d\mu \right| \leq C \|\psi\|_{Lip} \|\phi\|_\infty \frac{1}{n^\gamma},$$

where $\gamma = \frac{1}{\alpha} - 1$.

Consequently, for the purposes of Theorem 3.1, Assumption (I) is satisfied with $\lambda(n) = \mathcal{O}(n^{-\mathfrak{p}})$ with $\mathfrak{p} = \frac{1}{\alpha} - 1$. Clearly the dimensions of μ and μ^ω are equal to one and Assumption (V) is satisfied with any $d_0 < 1 < d_1$ arbitrarily close to 1 for any point x away from the parabolic point at 0. Assumption (VI) is satisfied with $\xi = \beta = 1$.

Let us denote by ψ_0 for the parabolic inverse branch of T and denote by ψ_1 the other inverse branch $\psi_1(x) = \frac{1}{2} + \frac{x}{2}$. Then ψ_0^n the (unique) inverse branch of T^n which contains the parabolic point 0, then one has that $|\psi_0^n(I)| = \mathcal{O}(n^{-1/\alpha})$, where $\psi_0^n(I) = [0, a_n]$, $a_n = \psi_0^n(1) \sim n^{-\frac{1}{\alpha}}$.

One has $(\psi_0^k)'(1) = 1/(T^k)'(a_k)$ and since $T'(s) = 1 + (1+\alpha)2^\alpha x^\alpha$ we obtain

$$\begin{aligned}
(T^k)'(a_k) &= \prod_{\ell=1}^{k} (1 + (1+\alpha)2^\alpha a_\ell^\alpha) \\
&\sim \exp \sum_{\ell=1}^{k} (1+\alpha)2^\alpha a^\alpha \frac{1}{\ell} \\
&\sim \exp\left((1+\alpha)(2\alpha)^\alpha \log k\right) \\
&= k^c,
\end{aligned}$$

where $c = \frac{1+\alpha}{(2\alpha)^\alpha} = \frac{1+\alpha}{\alpha} = 1 + \frac{1}{\alpha}$.

Now for $j \in \mathbb{N}$, let us put $\zeta = \psi_{\vec{i}}(I)$ for the j-cylinder given by $\vec{i} = (i_1, i_2, \dots, i_j) \in \{0,1\}^j$, where $\psi_{\vec{i}} = \psi_{i_j} \circ \cdots \psi_{i_2} \circ \circ \psi_{i_1}$. The ζ then describe all possible j-cylinders as $\vec{i}$ ranges over all possibilities. For some $\beta < 1$ to be determined later, let us put $q = j^\beta$ and define the set of 'bad' j-cylinders $\mathcal{G}_j^c$ by

$$\mathcal{G}_j^c = \left\{ \zeta = \psi_{\vec{i}}(I) : \vec{i} \in \{0,1\}^j, i_1 = i_2 = \cdots = i_q = 0 \right\}$$

and put $G_j^c = \bigcup_{\zeta \in \mathcal{G}_j^c} \zeta$. Then

$$\mu(G_j^c) \le \mu(U_q) \lesssim \frac{1}{q^\gamma}.$$

Now let n_0 be so that $x \in U_{n_0} \setminus U_{n_0 - 1}$, then $\zeta \cap B_\rho(x) \neq \varnothing$ if $\zeta = \psi_{\vec{i}}(I)$, for $i_{j-n_0} = i_{j-n_0+1} = \cdots = i_j = 0$.

7.2.1. *Away from the parabolic point.* From now on we let $\Gamma = \{x\}$ for $x \in I$, $x \neq 0$. In order to get better estimates of the term $\mathcal{R}^2$ we will, unlike done earlier, stratify the set $\mathcal{G}_j^n$ as follows. If $\zeta \in \mathcal{G}_j$ then there exists $s < q$ so that $\zeta = \psi_{\vec{i}}(I)$, where $i_1 = i_2 = \cdots = i_s = 0$ and $i_{s+1} = 1$. In this case, if moreover $\zeta \cap B_\rho(x) \neq \varnothing$, then

$$\operatorname{diam}(\zeta) \lesssim \operatorname{diam}(U_s) \frac{1}{(j - s - n_0)^c} \frac{1}{n_0^c}.$$

Or, since $\operatorname{diam}(U_s) \le a_s \lesssim s^{-\frac{1}{\alpha}}$ we get

$$\delta_s(j) \lesssim \frac{1}{s^{\frac{1}{\alpha}}} \frac{1}{((j - s - n_0)n_0)^c}$$

which is the diameter estimate for those j-cylinders for which have the given value for s. In a similar way we can estimate the distortion of those cylinders by

$$\mathfrak{D}_s(j) \lesssim \text{distortion}(T_\alpha^s|_{U_s}) \lesssim s^c.$$

Since the dimension of the measures μ on the unit interval I is equal to 1, we get the following refined estimate of the error term $\mathcal{R}_2$ where the gap is again denoted by Δ. For the contribution of the short returns we get (as $q \ll j$)

$$\begin{aligned}
\mathcal{R}_{2,\text{good}} &\lesssim \sum_{j=L'}^{\Delta}\sum_{s=0}^{q-1}\delta_s(j)\mathfrak{D}_s(j)\\
&\lesssim \sum_{j=L'}^{\Delta}\sum_{s=0}^{q-1}\frac{s^{1+\frac{1}{\alpha}}}{s^{\frac{1}{\alpha}}((j-s-n_0)n_0)^c}\\
&\lesssim \sum_{j=L'}^{\Delta}\frac{q^2}{j^c}\\
&\lesssim \frac{1}{L'^{\frac{1}{\alpha}-2\beta}}
\end{aligned}$$

provided $\frac{1}{\alpha}-2\beta$ is positive, i.e. $\beta < \frac{1}{2\alpha}$ and $L' < L$. Therefore

$$\mathcal{R}_2 \lesssim \frac{1}{L'^{\frac{1}{\alpha}-2\beta}} + \mathfrak{G}_{L'}.$$

The estimates of the other terms $\mathcal{R}_1$ and $\mathcal{R}_3$ proceed unchanged. As before we approximate the characteristic function $\mathbb{1}_{B_\rho(x)}$ from the outside by a Lipschitz continuous function ϕ which smoothens on an annulus of thickness ρ^w for a $w > 1$ and in a similar way with a Lipschitz function $\hat{\phi}$ from the inside. Then $\|\phi\|_{\text{Lip}}, \|\hat{\phi}\|_{\text{Lip}} \lesssim \rho^{-w}$ and gives for the correlation term:

$$\mathcal{R}_{1,\text{corr}} \lesssim a^L \rho^{-w}\frac{1}{\Delta^{\frac{1}{\alpha}-1}} \lesssim \rho^{v(\frac{1}{\alpha}-1)-w},$$

where we used the decay of correlation and where we put, as before $\Delta = \rho^{-v}$ for a $v < 1$. This tells us that the condition on the parameters is $v(\frac{1}{\alpha}-1)-1-w > 0$. For the annulus term we get as before

$$\mathcal{R}_{1,\text{ann}} \lesssim \rho^w \lesssim \rho^w \mathbb{P}(Z^L \geq 1)$$

which goes to zero as w is larger than 1. Then $\mathcal{R}_1 = \mathcal{R}_{1,\text{corr}} + \mathcal{R}_{1,\text{ann}}$. The third error term is then

$$\mathcal{R}_3 \lesssim \sum_{k=0}^{N}\sum_{j=1}^{\Delta}\mu(\mathcal{V})^2 = N\Delta\mu(\mathcal{V})^2 \lesssim \Delta\mu(\mathcal{V})$$

which goes to zero as $v < 1$. The total error now is

$$\mathcal{R} \lesssim N'(\mathcal{R}_1 + \mathcal{R}_2) + \mathcal{R}_3.$$

Since $\mathfrak{g} = \beta\gamma$ which implies that $\mu(G_j^c) \le j^{-\beta\gamma}$ and define $N_L = \sum_{j=L}^\infty \mathbb{1}_{G_j^c}$ for which we get

$$\mu(N_L) = \sum_{j=L}^\infty \mu(G_j^c) \lesssim \sum_{j=L}^\infty \frac{1}{j^{\beta\gamma}} \lesssim \frac{1}{L^{\beta\gamma-1}}.$$

The Hardy-Littlewood maximal function is then

$$HN_L(x) = \sup_{\rho>0} \frac{1}{\mu(B_\rho(x))} \int_{B_\rho(x)} N_L \, d\mu = \sup_{\rho>0} \sum_{j=L}^\infty \frac{\mu(G_j^c \cap B_\rho(x))}{\mu(B_\rho(x))}$$

for which we get by the maximal inequality for $\varepsilon > 0$:

$$\mathbb{P}_\mu(HN_L > \varepsilon) \lesssim \frac{1}{\varepsilon} \int_I N_L \, d\mu \lesssim \frac{1}{\varepsilon L^{\beta\gamma-1}} \qquad \forall \rho > 0.$$

If we put $\varepsilon = L^{-(\beta\gamma-1)/2}$ then

$$\mathbb{P}_\mu\left(\sum_{j=L}^\infty \frac{\mu(G_j^c \cap B_\rho(x))}{\mu(B_\rho(x))} > \frac{1}{L^{(\beta\gamma-1)/2}} \right) \lesssim \frac{1}{L^{(\beta\gamma-1)/2}}.$$

Since $N_L \ge 1$ if $N_L \ne 0$ we get that $\mathbb{P}_\mu(N_L \ne 0) \lesssim L^{-(\beta\gamma-1)}$ decays to zero as $L \to \infty$ if $\gamma > 1$. By [6] Theorem 3.18 we now obtain

$$\mathfrak{G}_{L,\rho}(x) = \sum_{j=L}^\infty \frac{\mu(G_j^c \cap B_\rho(x))}{\mu(B_\rho(x))} \longrightarrow N_L = 0$$

for all $x \in \{N_L = 0\}$ as $\rho \to 0$. Since $\mathbb{P}(N_L \ne 0)$ is summable if $\beta(\frac{1}{\alpha} - 1) - 1 > 1$ and $\beta < 1$ can be arbitrarily close to 1, we see that if $\alpha < \frac{1}{3}$ by the Borel-Cantelli theorem

$$\lim_{L\to\infty} \lim_{\rho\to 0} \mathfrak{G}_{L,\rho}(x) = 0$$

for almost every $x \in I$, i.e. μ-almost every x lies in $\liminf_{L,\rho}\{N_L = 0\}$.

Let us note that if $x \in \liminf_{L,\rho}\{N_L = 0\}$ is periodic with minimal period m then we get that the limiting distribution is Pólya-Aeppli with parameter $\vartheta = |DT^m(x)|^{-1}$ and the Kac. scaling $N = \frac{t}{\mathbb{P}(Z^L \ge 1)/L} \sim \frac{\vartheta t}{\mu(B_\rho)}$. If $x \in \liminf_{L,\rho}\{N_L = 0\}$ is non-periodic then the limiting distribution is Poisson(t) with the scaling $N = \frac{t}{\mathbb{P}(Z^L \ge 1)/L} \sim \frac{t}{\mu(B_\rho)}$.

7.2.2. *Asymptotics at the parabolic point.* If $x = 0$ we obtain a different scaling since in this case $\alpha_1 = 0$. Instead of $\rho \to 0$ we shall use the neighbourhoods $U_n = [0, a_n]$. Denote by $A_n = U_n \setminus U_{n+1} = (a_{n+1}, a_n]$ and $V_{n,L} = \bigcup_{j=n}^{n+L-1} A_j = U_n \setminus U_{n+L}$ for $L \in \mathbb{N}$. For given n let as before $Z^N = \sum_{j=0}^{N-1} \mathbb{1}_{U_n} \circ T^j$ be the hit counting function where we assume for simplicity's sake that $N = rL$ for some $r, L \in \mathbb{N}$. If $k < K$ (for some $K < L$) we obtain

$$\begin{aligned} \mathbb{P}(Z^N = k) &= \mathbb{P}\left(\bigcap_{\vec{k} \in \mathscr{K}_r(k)} \bigcap_{j=1}^r T^{-(j-1)L}\{Z^L = k_j\} \right) \\ &= \mathbb{P}\left(\bigcap_{\vec{k} \in \mathscr{K}_r(k)} \bigcap_{j=1}^r T^{-(j-1)L}\{\hat{Z}^L = k_j\} \right) \end{aligned}$$

$$= \mathbb{P}\left(\sum_{j=0}^{N-1} \hat{Z}^L \circ T^{jL} = k\right),$$

where $\mathscr{K}_r(k) = \{\vec{k} \in \mathbb{N}_0^r : \sum_{j=1}^r k_j = k\}$ and $\hat{Z}^L = \sum_{j=1}^{L-1} \mathbb{1}_{V_{n,K}} \circ T^j$. This is because if $Z_{U_n}^N = k < K$ then for the iterates one has $T^j x \notin U_{n+K}$ for $j = 0, \ldots, N$. Let $K = L^\eta$ for some $\eta \in (0,1)$. Then $K \to \infty$ as $L \to \infty$.

Lemma 7.1. *For all* $\alpha \in (0,1)$*:*

$$\frac{\mathbb{P}(\hat{Z}^L \geq 1)}{\mu(V_{n,L+K-1})} \longrightarrow 1$$

as $n \to \infty$*. In particular there is a constant* C *so that*

$$C^{-1} n^{-\frac{1}{\alpha}} \leq \mathbb{P}(\hat{Z}^L \geq 1) \leq C n^{-\frac{1}{\alpha}}.$$

Proof. We do the following decomposition as $\psi_0^i V_{n,K} = V_{n+i,K}$:

$$\begin{aligned}
\{\hat{Z}^L \geq 1\} &= \bigcup_{\ell=0}^{L-1} T^{-\ell} V_{n,K} \\
&= \bigcup_{\ell=0}^{L-1} \left(V_{n+\ell,K} \cup \bigcup_{i=0}^{\ell-1} T^{-(\ell-i-1)} \psi_1 V_{n+i,K} \right) \\
&= V_{n,K+L-1} \cup \mathcal{E}
\end{aligned}$$

where we get the error term ($k = \ell - i$)

$$\mathcal{E} = \bigcup_{\ell=0}^{L-1} \bigcup_{i=0}^{\ell-1} T^{-(\ell-i-1)} \psi_1 V_{n+i,K} = \bigcup_{k=0}^{L-1} T^{-(k-1)} \bigcup_{i=0}^{L-k} \psi_1 V_{n+i,K} = \bigcup_{k=0}^{L-1} T^{-(k-1)} \psi_1 V_{n,K+L-k}$$

which can be estimated as follows:

$$\mu(\mathcal{E}) \lesssim \sum_{k=0}^{L-1} \frac{K+L-k}{n^{\frac{1}{\alpha}+1}} \lesssim \frac{L(K+L)}{n^{\frac{1}{\alpha}+1}}.$$

Since

$$\frac{\mu(V_{n,K+L-1})}{(K+L-1) n^{-\frac{1}{\alpha}}}$$

is bounded and bounded away from 0 uniformly in n we get the desired estimates. ∎

Lemma 7.2. *If* $\alpha, \beta \in (0,1)$ *then*

$$\mathfrak{G}_{L,n} = \sum_{j=L}^{\infty} \frac{\mu(V_{n,K} \cap T^{-j}\{\hat{Z}^L \geq 1\} \cap G_j^c)}{\mathbb{P}(\hat{Z}^L \geq 1)} \longrightarrow 0$$

as $n \to \infty$*.*

Proof. In order to verify Assumption (II-iv) we write

$$V_{n,K} \cap T^{-j}\{\hat{Z}^L \geq 1\} \cap G_j^c = V_{n,K} \cap T^{-j} \bigcup_{\ell=0}^{L-1} T^{-\ell} V_{n,K} \cap T^{-q} U_q = V_{n,K} \cap T^{-(j-q)} \hat{\mathcal{V}}$$

(as $G_j^c = T^{-q}U_q$), where

$$\hat{\mathcal{V}} = \bigcup_{\ell=0}^{L-1}\left(V_{n+\ell+q,K} \cup \bigcup_{i=0}^{\ell+q-1-i} T^{-(\ell+q-1-i)}\psi_1 V_{n+i,K} \cap U_q\right)$$

Since

$$T^{-(\ell+q-1-i)}\psi_1 V_{n+i,K} \cap U_q \subset A_{\ell+q-1-i} \cap U_q = \varnothing$$

if $\ell + q - i - 1 < q$, or, if $i > \ell - 1$, we can put

$$\hat{\mathcal{V}}' = \bigcup_{\ell=0}^{L-1}\bigcup_{i=0}^{\ell-1} T^{-(\ell+q-1-i)}\psi_1 V_{n+i,K} \cap U_q$$

which implies $\hat{\mathcal{V}}' \subset V_{q,L-1}$ and to write

$$\hat{\mathcal{V}} = \bigcup_{\ell=0}^{L-1} V_{n+\ell+q,K} \cup \hat{\mathcal{V}}' = V_{n+q,K+L-1} \cup \hat{\mathcal{V}}'.$$

Note that since

$$T^{-(\ell+q-1-i)}\psi_1 V_{n+i,K} \cap U_q = \bigcup_{k=q}^{\ell+q-i-1} \psi_0^k \psi_1 T^{-(\ell+q-i-2-k)}\psi_1 V_{n+i,K}$$

(the term $k = \ell + q - 1 - i$ is when one of the maps ψ_1 is deleted) we get for the Lebesgue measure m

$$\begin{aligned} m\big(T^{-(\ell+q-1-i)}\psi_1 V_{n+i,K} \cap U_q\big) &\lesssim \sum_{k=q}^{\ell+q-i-1} m\big(\psi_0^k \psi_1 T^{-(\ell+q-i-2-k)}\psi_1 V_{n+i,K}\big) \\ &\lesssim \sum_{k=q}^{\ell+q-i-1} \frac{1}{k^{\frac{1}{\alpha}+1}} m(V_{n+i,K}) \\ &\lesssim \frac{K}{(n+i)^{\frac{1}{\alpha}+1} q^{\frac{1}{\alpha}}} \end{aligned}$$

as $D\psi_0^k|_{V_0} \lesssim k^{-(\frac{1}{\alpha}+1)}$. Consequently

$$m(\hat{\mathcal{V}}') \lesssim \frac{KL^2}{n^{\frac{1}{\alpha}+1} q^{\frac{1}{\alpha}}}$$

and therefore

$$\begin{aligned} m(\hat{\mathcal{V}}) &\lesssim m(V_{n+q,K+L-1}) + m(\hat{\mathcal{V}}') \\ &\lesssim \frac{K+L}{(n+q)^{\frac{1}{\alpha}+1}} + \frac{KL^2}{n^{\frac{1}{\alpha}+1} q^{\frac{1}{\alpha}}} \\ &\lesssim \frac{KL^2}{(n+q)^{\frac{1}{\alpha}+1}}. \end{aligned}$$

Since $h|_{V_{n,K}} \lesssim n$ one has

$$\mu(V_{n,K} \cap T^{-j}\{\hat{Z}^L \geq 1\} \cap G_j^c) \lesssim n \cdot m(V_{n,K} \cap T^{-(j-q)}\hat{\mathcal{V}})$$

and therefore

$$V_{n,K}\cap T^{-(j-q)}\hat{\mathcal{V}}=(V_{n,K}\cap\psi_0^{j-q}\hat{\mathcal{V}})\cup\bigcup_{i=0}^{j-q-1}\bigcup_{k=q}^{j-q-i-1}\psi_0^k\psi_1T^{-(j+q-i-2-k)}\psi_1\psi_0^i\hat{\mathcal{V}}\cap V_{n,K}.$$

Since $\psi_0^k\psi_1T^{-(j+q-i-2-k)}\psi_1\psi_0^i\hat{\mathcal{V}}\subset A_k$ we must have $n\le k<n+K$ which implies in particular that $j-q-i-2\ge n$ and thus $j\ge n+1$ in the second term. For $k\in[n,n+K)$ this leads to

$$\begin{aligned}m(\psi_0^k\psi_1T^{-(j+q-i-2-k)}\psi_1\psi_0^i\hat{\mathcal{V}}\cap V_{n,K}) &\lesssim \frac{1}{k^{\frac{1}{\alpha}+1}}\left(\frac{q}{q+i}\right)^{\frac{1}{\alpha}+1}m(\hat{\mathcal{V}})\\ &\lesssim \frac{1}{n^{\frac{1}{\alpha}+1}}\left(\frac{q}{q+i}\right)^{\frac{1}{\alpha}+1}\frac{KL^2}{(n+q)^{\frac{1}{\alpha}+1}}\end{aligned}$$

as $D\psi_0^i|_{V_{n,K}}\lesssim\left(\frac{q}{q+i}\right)^{\frac{1}{\alpha}+1}$. Since $h|_{V_{n,K}}\lesssim n$ this leads to

$$\begin{aligned}&\mu\left(\bigcup_{i=0}^{j-q-1}\bigcup_{k=q}^{j-q-i-1}\psi_0^k\psi_1T^{-(j+q-i-2-k)}\psi_1\psi_0^i\hat{\mathcal{V}}\cap V_{n,K}\right)\\ &\qquad\lesssim n\sum_{i=0}^{j-q-1}K\frac{1}{n^{\frac{1}{\alpha}+1}}\left(\frac{q}{q+i}\right)^{\frac{1}{\alpha}+1}\frac{KL^2}{(n+q)^{\frac{1}{\alpha}+1}}\\ &\qquad\lesssim \frac{K^2L^2}{n^{\frac{1}{\alpha}}}\frac{q^{\frac{1}{\alpha}+1}}{q^{\frac{1}{\alpha}}}\frac{1}{(n+q)^{\frac{1}{\alpha}+1}}\\ &\qquad= \frac{K^2L^2q}{n^{\frac{1}{\alpha}}(n+q)^{\frac{1}{\alpha}+1}}.\end{aligned}$$

For the other term we get

$$\mu(V_{n,K}\cap\psi_0^{j-q}\hat{\mathcal{V}})\lesssim n\cdot m(V_{n,K}\cap\psi_0^{j-q}\hat{\mathcal{V}})\lesssim n\Big(m(V_{n,K}\cap V_{n+j,K+L-1})+m(V_{n,K}\cap\psi_0^{j-q}\hat{\mathcal{V}}')\Big).$$

Since for $j\ge L$ one has $V_{n,K}\cap V_{n+j,K+L-1}=\varnothing$ and since $\hat{\mathcal{V}}'\subset V_{q,L-1}$ one gets $\psi_0^{j-q}\hat{\mathcal{V}}'\subset V_{j,L-1}$ and therefore $V_{n,K}\cap\psi_0^{j-q}\hat{\mathcal{V}}'\neq\varnothing$ only if $n-L\le j<n+K$ we obtain

$$m(V_{n,K}\cap\psi_0^{j-q}\hat{\mathcal{V}}')\lesssim\frac{K+L}{n^{\frac{1}{\alpha}+1}n^{\beta\frac{1}{\alpha}}}$$

as $q=j^\beta$.

Finally we arrive at

$$\mu(V_{n,K}\cap T^{-j}\{\hat{Z}^L\ge1\}\cap G_j^c)\lesssim\frac{k^2L^2j^\beta}{n^{\frac{1}{\alpha}}(n+j^\beta)^{\frac{1}{\alpha}+1}}\chi_{[n,\infty)}(j)+\frac{k+L}{n^{\frac{1}{\alpha}}n^{\frac{\beta}{\alpha}}}\chi_{[n-L,n+K)}(j)$$

and since $\mathbb{P}(\hat{Z}^L\ge1)\gtrsim LKn^{-\frac{1}{\alpha}}$ this gives us

$$\mathfrak{G}_{L,n}\lesssim\sum_{j=L}^{\infty}\frac{j^\beta}{(n+j^\beta)^{\frac{1}{\alpha}+1}}\chi_{[n,\infty)}(j)+\frac{1}{n^{\frac{\beta}{\alpha}}}\longrightarrow0$$

as $n \to \infty$. □

The estimates of the $\mathcal{R}$ terms are much the same as for non-parabolic points:

$$\mathcal{R}_{1,\mathrm{corr}} \lesssim \frac{a^L}{\rho^w}\frac{1}{\Delta^\gamma} \lesssim a^L \rho^{-w} N^{-v\gamma},$$

where we put $\Delta = N^v$ for some $v < 1$. Similarly

$$\mathcal{R}_{1,\mathrm{ann}} \lesssim \rho^w \mu(\mathcal{V}) h(a_n) \lesssim n\rho^w \mu(\mathcal{V}).$$

No change in the $\mathcal{R}_3$ term as we still have $\mathcal{R}_3 \lesssim \mu(\mathcal{V})^{1-v}$. Similarly we have

$$\mathcal{R}_2 \lesssim \frac{1}{L'^{\frac{1}{\alpha}-2\beta}} + \mathfrak{G}_{L',n}.$$

For the coefficients of the limiting compound Poisson distribution we get if $\ell \geq 2$:

$$\lambda_\ell = \lim_{L\to\infty}\lim_{n\to\infty} \frac{\mathbb{P}(\hat{Z}^L = \ell)}{\mathbb{P}(\hat{Z}^L \geq 1)} = \lim_{L\to\infty}\lim_{n\to\infty} \frac{\mu(V_{n,\ell})}{\mu(V_{n,L-K-1})}(1 + o(1)) = 0$$

which implies that we get in the limit that $\lim_{L\to\infty}\lim_{n\to\infty} \frac{\mathbb{P}(Z^N=0)}{\mathbb{P}(Z^L\geq 1)} \longrightarrow e^t$, that is the appropriately rescaled entry times converge to an exponential distribution.

Notice that for the parabolic point we have the non-Kac scaling $N = \frac{t}{\mathbb{P}(\hat{Z}^L\geq 1)/L} \sim \frac{tL}{\mu(V_{n,L})} \sim tn^{\frac{1}{\alpha}}$ rather than the standard Kac scaling which would require N to have the value $\frac{t}{\mu(U_n)} \sim tn^{\frac{1}{\alpha}-1}$ which grows much more slowly.

An application of Theorem 5.1 then leads to the following result.

Theorem 7.2. *Let $T : I \circlearrowleft$ be as described above, where the map T is the parabolic map T_α. Assume $0 < \alpha < \frac{1}{3}$. Denote by μ the absolutely continuous invariant measure, then for all $t > 0$ the counting function*

$$W_{x,\rho} = \sum_{j=0}^{N-1} \mathbb{1}_{B_\rho(x)} \circ T^j$$

converges in distribution to Poisson(t) for Lebesgue almost every $x \in [0, 1]$ and for $x = 0$, where for:
$x \neq 0$: $N = \frac{t}{\mathbb{P}(Z^L_{B_\rho(x)}\geq 1)/L}$
$x = 0$: $N = \frac{t}{\mu(V_{n,L+K})/L}$ where $K < L$ (e.g. $K = L^\eta$ for some $\eta < 1$) and a double limit $n \to \infty$ and then $L \to \infty$.

References

[1] M Abadi: Exponential Approximation for Hitting Times in Mixing Stochastic Processes; Mathematical Physics Electronic Journal **7** (2001).

[2] M Abadi and B Saussol: Hitting and returning into rare events for all alpha-mixing processes; Stoch. Proc. Appl. **121** (2011), 314–323.

[3] M Abadi and N Vergne: Sharp errors for point-wise Poisson approximations in mixing processes; Nonlinearity **21** (2008), 2871–2885.

[4] Chazottes, J-R. and P. Collet. "Poisson approximation for the number of visits to balls in non-uniformly hyperbolic dynamical systems." Ergodic Theory and Dynamical Systems 33.1 (2013): 49–80.

[5] M Denker, M Gordin and A Sharova. "A Poisson limit theorem for toral automorphisms." Illinois Journal of Mathematics 48.1 (2004): 1–20.
[6] G Folland. *Real Analysis*, Wiley 1999.
[7] J M Freitas, N Haydn and M Nicol: Convergence of rare event point processes to the Poisson process for planar billiards, Nonlinearity; **27(7)** (2014), 1669–1687
[8] A C M Freitas, J M Freitas, M Magalhães: Convergence of Marked Point Processes of Excesses for Dynamical Systems; J. Eur. Math. Soc. (JEMS), **20(9)** (2018), 2131–2179.
[9] D Faranda, A C Moreira Freitas, J Milhazes Freitas, M Holland, T Kuna, V Lucarini, M Nicol, M Todd, S Vaienti; *Extremes and Recurrence in Dynamical Systems*, Wiley, New York, 2016.
[10] A C M Freitas, J M Freitas, M Todd: The compound Poisson limit ruling periodic extreme behaviour of non-uniformly hyperbolic dynamics; Comm. Math. Phys. **321(2)** (2013), 483–52.
[11] A C Freitas, J M Freitas, M Todd, S Vaienti: Rare events for the Manneville-Pomeau map; 2018
[12] A Galves and B Schmitt: Inequalities for hitting time in mixing dynamical systems; Random Comput. Dynam. **5** (1997), 319–331.
[13] N T A Haydn and Y Psiloyenis: Return times distribution for Markov towers with decay of correlations, 2010, preprint available at `http://arxiv.org/abs/1010.1029`.
[14] N Haydn and S Vaienti: The distribution of return times near periodic orbits; Probability Theory and Related Fields **144** (2009), 517–542.
[15] N Haydn and S Vaienti. "Limiting entry and return times distribution for arbitrary null sets." Communications in Mathematical Physics **378** (2020): 149–184.
[16] N T A Haydn and K Wasilewska: Limiting distribution and error terms for the number of visits to balls in non-uniformly hyperbolic dynamical systems; Disc. Cont. Dynam. Syst. **36(5)** (2016), 2585–2611.
[17] M Hirata: Poisson law for Axiom A diffeomorphisms; Ergod. Th. & Dynam. Syst. **13** (1993), 533–556.
[18] M Kač: On the notion of recurrence in discrete stochastic processes; Bull. Amer. Math. Soc. **53** (1947), 1002–1010.
[19] Xin Li and Helder Vilarinho: Almost sure mixing rates for non-uniformly expanding maps; 2014, preprint available at `http://arxiv.org/abs/1501.00010`.
[20] F Pène, B Saussol, : Poisson law for some nonuniformly hyperbolic dynamical systems with polynomial rate of mixing; Ergodic Theory Dynam. Systems, **36(8)** (2016), 2602–2626.
[21] B Pitskel: Poisson law for Markov chains; Ergod. Th. & Dynam. Syst. **11** (1991), 501–513.
[22] J. Rousseau: Hitting time statistics for observations of dynamical systems, Nonlinearity **27** (2014) 2377–2392.