Discrete and Continuous Dynamical Systems
Vol. 45, No. 5, May 2025, pp. 1454-1491
doi:10.3934/dcds.2024135

POISSON LIMIT THEOREMS FOR SYSTEMS WITH PRODUCT STRUCTURE

Max Auer
University of Maryland, College Park, USA

(Communicated by Mark F. Demers)

Abstract. We obtain a Poisson Limit for return times to small sets for product systems. Only one factor is required to be hyperbolic, while the second factor is only required to satisfy polynomial deviation bounds for ergodic sums. In particular, the second factor can be elliptic or parabolic. As an application of our main result, several maps of the form Anosov map $\times$ another map are shown to satisfy a Poisson Limit Theorem at typical points, some even at all points. The methods can be extended to certain types of skew products, including T, T^{-1}-maps of high rank.

Part 1. Results

1. Introduction. One of the prominent limit theorems in classical probability theory is the Poisson Limit Theorem. (PLT). Due to the PLT, a variety of probabilistic models describing waiting times until unlikely events occur are well approximated by exponentially distributed variables. It has been a great discovery that many deterministic systems satisfy the same kind of limit theorems for rare events.

Limit distributions of waiting times are most classical for mixing Markov chains, where one considers returns to small cylinders, for example, see [37, Theorem A]. As remarked there, this result can be immediately generalized to systems with a Markov partition, the only caveat being that the sets are still cylinders, so geometrically not the most intuitive class. Nonetheless, waiting time limits for returns to small balls can be shown in concrete settings; for example hyperbolic toral automorphisms [12, Theorem 2.3] or more general hyperbolic maps [40, Theorem 2.8], Rychlik-maps and unimodal maps [7, Theorem 3.2 and 4.1], partially or nonuniformly hyperbolic maps [13, Theorem 8] [35, Theorem 3.3] [10, Theorem 3.3], open billiard systems [9, Theorem 1], some intermittent interval maps [11, Main Theorem], and many more. It is sometimes interesting to also ask for explicit rates of convergence, this can be shown under strong mixing properties, see [28, Theorem 2.1], [1, Theorem 7], [26, Theorem 8]. Interestingly enough, in [40, Theorem 2.8] and [8, Theorem 5.11] novel techniques have been used to obtain rates in certain billiard systems without strong mixing assumptions. We do not make any claims on completeness of the list of references given above, for a more complete picture see [32].

2020 *Mathematics Subject Classification.* Primary: 28D05, 37A50, 37C05; Secondary: 37A25, 60F05.

Key words and phrases. Ergodic theory, hyperbolic dynamical system, limit distribution, Poisson process, rare events.

The author thanks the University of Maryland for their hospitality.

Some related topics are extreme value laws [22], [23], spatiotemporal limits [36], [43] or Borel-Cantelli like Lemmas [29], [25], [16].

Similar questions can be asked for flows as well, this topic has not been studied as thoroughly as the question for maps. As shown in [34] for suspension flows, this reduces to the study for maps. Moreover, the Poisson Limit Theorem for flows can be reduced to the Poisson Limit Theorem for time 1 map with the target being the set of points that visit $B(x,r)$ within the next unit of time.

From the list above, we see that the PLT is often associated with strong mixing properties of the system. In the present work, we construct systems that are not even weakly mixing but nevertheless satisfy the PLT (a precise definition is given beneath). The systems will have a special structure $S = T \times R$, where T is hyperbolic, but R is not.

We will develop a machinery to show the PLT for such systems. This will be used to construct systems satisfying the PLT, but otherwise exhibiting properties uncharacteristic of chaotic systems - like non weak mixing, or zero entropy[1]. This suggests that the PLT is much more common than it was believed before. In fact, discovering the most flexible conditions for the validity of the PLT is a promising direction of future research.

2. **Preliminaries.**

Definition 2.1. Given a probability-preserving ergodic dynamical system $(X, \mathcal{A}, \mu, T)$ and a measurable set $A \in \mathcal{A}$, we will define the *first entry time to* A as

$$\varphi_A(x) = \min(n \geq 1 \mid T^n(x) \in A),$$

the restriction $\varphi_A|_A$ to the set A itself shall be referred to as the *first return time to* A. The *first return map* shall be denoted by $T_A(x) = T^{\varphi_A(x)}(x)$, and the *sequence of consecutive return times* by

$$\Phi_A = (\varphi_A, \varphi_A \circ T_A, \varphi_A \circ T_A^2, ...).$$

In the following, for some measurable set $A \in \mathcal{A}$ with $\mu(A) > 0$, the measure conditional on A shall be given as $\mu_A(B) = \frac{\mu(A \cap B)}{\mu(A)}$, $B \in \mathcal{A} \cap A$. The first important result in the study of φ_A, was Kac's formula, which calculates the expectation as

$$\int_A \varphi_A \, \mathrm{d}\mu_A = \frac{1}{\mu(A)}.$$

Hence, it is natural to study limits of $\mu(A)\varphi_A$ as $\mu(A) \to 0$. More explicitly let $(A_l)_{l \geq 1}$ be a sequence of rare events, that is each A_l is measurable with $\mu(A_l) \to 0$, we want to find weak limits of the form

$$\mu(A_l)\Phi_{A_l} \overset{\mu}{\Rightarrow} \Phi \quad \text{as } l \to \infty,$$

or

$$\mu(A_l)\Phi_{A_l} \overset{\mu_{A_l}}{\Rightarrow} \tilde{\Phi} \quad \text{as } l \to \infty,$$

where $\Rightarrow$ denotes convergence in distribution. In the above situation, we shall call Φ the *hitting time limit* and $\tilde{\Phi}$ the *return time limit*. An important fact is that the hitting and return time limits are intimately related, this relation was first formulated in [24, Main Theorem] (albeit only for the first marginal). The analogous relation for the entire process is shown in [42, Theorem 3.1]. For exponential returns, which is what we are concerned with, the result is as follows.

[1] This cannot be done with products, since $h(T \times R) = h(T) + h(R)$, where h denotes the metric entropy of a system. We extend our methods to skew-products of a certain form (Theorem 3.4.)

Theorem 2.2. *Let $(X, \mathcal{A}, \mu, T)$ be an ergodic probability preserving dynamical system, and let $(A_l)_{l\geq 1}$ be a sequence of rare events. Then $\Phi \overset{d}{=} \Phi_{Exp}$ is equivalent to $\tilde{\Phi} \overset{d}{=} \Phi_{Exp}$, where Φ_{Exp} is a process of iid standard exponentially distributed random variables.*

This suggests that we should expect exponential hitting and return time limits for geometrically sensible sequences of rare events.

In the following, let X be a C^r Riemannian manifold with $\dim(X) = d$ and assume $\mu \ll m_X$ the volume on X, with continuous density, say $\frac{\mathrm{d}\mu}{\mathrm{d}m_X} = \rho$. Most of the statements can be reformulated to hold for arbitrary invariant μ, but for the sake of simplicity, we shall keep this assumption.

Definition 2.3. (i) Let $x^* \in X$ and, for $r > 0$, denote by $B_r(x^*)$ the geodesic ball of radius r centred at x^*. We will say that T *satisfies the PLT at* x^* if

$$\mu(B_r(x^*))\Phi_{B_r(x^*)} \overset{\mu}{\Rightarrow} \Phi_{Exp} \quad \textit{as } r \to 0.$$

(ii) Let

$$PLT := \{x^* \in X \mid T \textit{ satisfies the PLT at } x^*\}.$$

If $\mu(PLT) = 1$ we say that T *satisfies the PLT almost everywhere*, and if $PLT = X$ we will say that T *satisfies the PLT everywhere*.

If T is Lipschitz-continuous along the (finite) orbit of a periodic point, then it **does not** satisfy the PLT at that point. To see this, note first that the PLT at x^* in particular implies, via Theorem 2.2, that, for each $N \geq 1$,

$$\mu_{B_r(x^*)}(\varphi_{B_r(x^*)} \leq N) \to 0 \quad \textit{as } r \to 0.$$

Now suppose x^* is a point with period p, and say $\rho(x^*) > 0$, and $|T^p(x) - T^p(y)| \leq C|x - y|$ near x^* then[2]

$$\mu_{B_r(x^*)}(\varphi_{B_r(x^*)} \leq p) \geq \mu_{B_r(x^*)}(B_{\frac{r}{C}}(x^*)) = \frac{1}{C^d} + o(1)$$

as $r \to 0$.

Situations where x^* is a periodic point are more delicate, and the limiting distribution is not exponential any more (due to immediate returns). For example, in [43, Theorem 3.3 and Theorem 10.1] this question was studied for expanding interval maps.

The main goal is to prove the PLT (almost) everywhere for some (skew-) product systems.

In the following we will consider return times in different systems - namely, we will have three different maps $T : X \to X$ (or $T_y : X \to X$) which is usually assumed hyperbolic, $R : Y \to Y$ which is parabolic or elliptic, and $S = T \times R : X \times Y \to X \times Y$ - in an attempt to keep notation simple we will (by slight abuse of notation) always

[2] Assuming $C > 1$, we have that $B_r(x^*)$ is diffeomorphic via the exponential map exp_{x^*} to a ball in $\mathbb{R}^d$. W.l.o.g. assume that $X \subset \mathbb{R}^N$

$$\begin{aligned}\mu(B_r(x^*)) &= \int_{B_r(x^*)} \rho(x)\,\mathrm{d}m_X(x) = (\rho(x^*) + o(1))m_X(B_r(x^*))\\ &= (\rho(x^*) + o(1)) \int_{(B_r(0))} \sqrt{|\det D_{\mathbf{u}}exp_{x^*}(D_{\mathbf{u}}exp_{x^*})^t}\,\mathrm{d}\lambda^d(\mathbf{u})\\ &= (\rho(x^*)\sqrt{|\det D_{\mathbf{0}}exp_{x^*}(D_{\mathbf{0}}exp_{x^*})^t} + o(1))\lambda^d(B_r(0)),\end{aligned}$$

where λ^d is the d-dimensional Lebesgue measure.

denote the return times by φ. Which map is meant will always be clear by the specified set.

3. **The PLT for (skew-)products.** In this paper, we study the PLT for systems that can be written as a product (or skew product of a special type). Therefore, let Y be another Riemannian $C^{r'}$-manifold with $\dim(Y) = d'$, and assume $R : Y \to Y$ preserves a probability measure $\nu \ll m_Y$ with continuous density. Instead of $T : X \to X$, consider now some[3] $T : X \times Y \to X$. We will prove the PLT for certain systems of the form $S(x, y) = (T(x, y), R(y))$. The case of direct products can be recovered if $T(x, y) = T(x)$ is independent of y (which will be the case for most of our examples). Denote also $T_y(x) = T(x, y)$. We will assume that T_y preserves a probability measure μ (independent of y). For measurable $A \subset X$ we introduce analogously the *consecutive fiberwise return times* as

$$\varphi_{A\times Y}(x, y) = \min(j \geq 1 \mid S^j(x, y) \in A \times Y),$$
$$\Phi_{A\times Y} = (\varphi_{A\times Y}, \varphi_{A\times Y} \circ S_{A\times Y}, \varphi_{A\times Y} \circ S^2_{A\times Y}, \dots),$$

where $S_{A\times Y} = S^{\varphi_{A\times Y}}$ is the first return map to $A \times Y$, note that we only fix a small target in the fiber.

For our purposes, it is convenient to think of y as fixed. For $n \geq 1$ denote $T^n_y(x) = T_{R^{n-1}(y)}(T_{R^{n-2}(y)}(...(T_y(x))))$, and define

$$\varphi_{A,y}(x) = \varphi^{(1)}_{A,y}(x) = \min(j \geq 1 \mid T^j_y(x) \in A),$$
$$\varphi^{(n+1)}_{A,y}(x) = \min(j \geq 1 \mid T_y^{\varphi^{(1)}_{A,y}(x)+\varphi^{(2)}_{A,y}(x)+\dots+\varphi^{(n)}_{A,y}(x)+j}(x) \in A),$$
$$\Phi_{A,y} = (\varphi^{(1)}_{A,y}, \varphi^{(2)}_{A,y}, \cdots).$$

Clearly the definitions coincide and $\Phi_{A,y}(x) = \Phi_{A\times Y}(x, y)$.

We will list here the main assumptions[4] we make in order to prove the PLT.

- (MEM) We will say T is *multiple exponentially mixing* there are constants $r > 0$, $C > 1$ and $\gamma > 0$ such that, for almost all $y \in Y$,
$$\begin{aligned} &\left| \int_X \prod_{j=0}^{n-1} f_j \circ T_y^{k_j} \,\mathrm{d}\mu - \prod_{j=0}^{n-1} \int_X f_j \,\mathrm{d}\mu \right| \\ &\qquad \leq C e^{-\gamma \min_{0\leq j_1<j_2\leq n-1} |k_{j_1} - k_{j_2}|} \prod_{j=0}^{n-1} ||f_j||_{C^r}, \end{aligned} \tag{1}$$
for $n \geq 1$, $f_0, ..., f_{n-1} \in C^r$ and $0 \leq k_0 \leq ... \leq k_{n-1}$.
- (EE) There are $r' > 0$ and $\delta < 1$ such that
$$\left|\left| \sum_{n=0}^{N-1} f \circ R^n - N \int f \,\mathrm{d}\nu \right|\right|_{L^2(\nu)} \leq C||f||_{C^{r'}} N^\delta \quad \forall f \in C^{r'},\ \forall N \geq 1. \tag{2}$$
- (LR(y^*)) There is a $c > 0$ such that, for $r > 0$ and ν-a.e $y \in B_r(y^*)$, we have
$$\varphi_{B_r(y^*)}(y) \geq c|\log(r)|. \tag{3}$$

[3] $X \times Y$ is considered as a Riemannian manifold with the natural (Euclidean) product metric $d((x, y), (x', y')) = \sqrt{d(x, x')^2 + d(y, y')^2}$. Analogously, one could consider different metrics, e.g. the box metric $d((x, x'), (y', y)) = \max(d(x, y), d(x', y'))$, where many of the proofs become easier. However, we will use here the Euclidean product metric as the most natural choice.

[4] We often only assume a subset of these, most commonly (MEM), (EE), and (BR(x^*, y^*)). But we will always state the current assumptions.

- (SLR(y^*)) There is a $\psi : (0,\infty) \to (0,\infty)$ with $|\log(r)| = o(\psi(r))$ as $r \to 0$ such that, for $r > 0$ and ν-a.e $y \in B_r(y^*)$, we have

$$\varphi_{B_r(y^*)}(y) \geq \psi(r). \tag{4}$$

- (LR'(x^*)) For ν-a.e $y \in Y$ there is a $c = c_y > 0$ such that, for $r > 0$ and μ-a.e $x \in B_r(x^*)$, we have

$$\varphi_{B_r(x^*),y}(x) \geq c|\log(r)|. \tag{5}$$

- (NSR(x^*)) There is a $\xi : (0,\infty) \to (0,\infty)$ with $|\log(r)| = o(\xi(r))$ as $r \to 0$ such that, for ν-a.e $y \in Y$, we have

$$\mu_{B_r(x^*)}(\varphi_{B_r(x^*),y} \leq \xi(r)) \to 0 \quad \textit{as } r \to 0. \tag{6}$$

- (BR(x^*, y^*)) One of the following is satisfied
  - (SLR(y^*)),
  - (NSR(x^*)) AND (LR(y^*)),
  - or (NSR(x^*)) AND (LR'(x^*)).

Colloquially, we will also refer to (MEM) as multiple exponential mixing, and to (EE) as the Quantitative Ergodic Theorem or effective ergodicity. Both are standard assumptions and have been studied for many classes of systems.

Conditions (LR), (LR'), (SLR), and (NSR) all are concerned with the fact that points in a small ball B cannot return to B too quickly. Sometimes in literature, the center x^* or y^* is referred to as a slowly recurrent point. For technical reasons, we need to distinguish different versions of slow recurrence, (SLR) being the strongest. The abbreviations (LR), (SLR), (NSR), and (BR) stand for 'large returns', 'strong large returns', 'no short returns', and 'big returns' respectively.

Remark 3.1. (i) In the case $T(x,y) = G_{\tau(y)}(x)$, where G is a flow satisfying (a continuous version of) (MEM)[5] and τ is bounded, the condition (NSR(x^*)) is satisfied at almost every x^*. Indeed, it was shown in [16, Lemma 4.13], albeit for maps instead of flows, that condition (NSR(x^*)) is satisfied[6] for G at almost every x^*. Since τ is bounded, T also has this property.

(ii) It is shown in [3, Lemma 5] that, for a map of positive entropy, condition (LR) is satisfied at almost every point (In fact (3) is satisfied for all $y \in B_r(y^*)$). This remains true for maps of the form $T(x,y) = G_{\tau(y)}(x)$, (in this case (LR') is satisfied) for bounded τ, where G has positive entropy.

(iii) Considering the previous remarks, it may seem unnecessary to state condition (SLR). Note however that none of the conditions can be satisfied at periodic points, and the maps we want to use for T will have plenty of periodic points. (SLR) will be useful to show the PLT **everywhere**, if we can choose R without periodic points.

---

[5] It is in fact enough to assume exponential mixing.

[6] It is shown that, for every fixed $A, K > 0$, we have

$$\mu_{B_r(x^*)}(B_r(x^*) \cap G^{-n}B_r(x^*)) \leq |\log(r)|^{-A} \quad \forall n \leq K|\log(r)|. \tag{7}$$

For $A > 1$, summing over $n \in [1, K|\log(r)|]$ yields

$$\mu_{B_r(x^*)}(\varphi_{B_r(x^*),G} \leq K|\log(r)|) \to 0 \quad \textit{as } r \to 0.$$

Since this is true for all $K > 0$, we can easily replace K by some $K(r) \nearrow \infty$ growing slowly enough. This is a routine argument which is left to the reader.

(iv) In most of the examples (see §4) we will have

$$\left\|\sum_{j=0}^{n-1} f\circ R^j - n\int f\,\mathrm{d}\nu\right\|_{L^2(\nu)} \le C\|f\|_{H^{r'}} n^\delta \quad \forall f\in H^{r'},\ \forall n\ge 1.$$

Since $C^{r'}\subset H^{r'}$ and $\|f\|_{H^{r'}}\le \|f\|_{C^{r'}}$ for $f\in C^{r'}$, this implies condition (EE).

Theorem 3.2. *Assume that $S(x,y)=(T(x,y),R(y))$ satisfies conditions (MEM), (EE), and (BR(x^*,y^*)) for some $(x^*,y^*)\in X\times Y$. If*

$$d > \frac{r'+d'\delta}{1-\delta} \tag{8}$$

then S satisfies the PLT at (x^*,y^*).

Corollary 3.3. *If $T(x,y)=T(x)$ preserves a smooth measure and satisfies (MEM), and R satisfies (EE), then $S=T\times R$ satisfies the PLT almost everywhere.*

If T preserves a smooth measure, then, by [17], T is Bernoulli, in particular, it has positive entropy. (NSR(x^*)) and (LR'(x^*)) are satisfied almost everywhere by Remark 3.1.

For some applications it will be useful to choose $T(x,y)=G_{\tau(y)}(x)$, where $\int_Y\tau\,\mathrm{d}\nu=0$. However, in this case, T will not satisfy condition (MEM). Fortunately, we can apply similar techniques if ergodic averages of τ grow faster than logarithmically. More explicitly denote $\tau_n=\sum_{j=0}^{n-1}\tau\circ R^j$, assume there is a $\zeta:\mathbb{N}\to(0,\infty)$ with $\log(n)=o(\zeta(n))$ and a $\kappa>0$ such that

$$\nu(|\tau_n|<\zeta(n))\le O(n^{-\kappa}). \tag{BA}$$

Theorem 3.4. *Assume that $S(x,y)=(T(x,y),R(y))$, where $T(x,y)=G_{\tau(y)}(x)$, satisfies conditions (MEM) with G instead of T. Suppose that R satisfies (EE), and τ satisfies (BA). Let $x^*\in X$, $y^*\in Y$. If there is a $\delta_2>0$ such that for small enough $\rho>0$ we have*

$$\varphi_{B_\rho(y^*)}\ge\rho^{-\delta_2}\quad \text{on } B_\rho(y^*), \tag{9}$$

and

$$d>\frac{r'+d'\delta}{1-\delta}\quad \text{and}\quad \kappa>\frac{d'}{\delta_2}. \tag{10}$$

then S satisfies the PLT at (x^*,y^*).

Remark 3.5. Let us remark here, that hitting times for skew product, say $S: X\times Y\to X\times Y$ with $S(x,y)=(T(x,y),R(y))$, have previously been investigated by other authors, for example, [27] and [38]. The main differences between [27] and our results are;

- In [27] the system is viewed from the standpoint of random dynamics, therefore the relevant target sets are of the form $B_\rho(x^*)\times Y$. In contrast, we focus on geometric balls $B_\rho(x^*,y^*)$.
- In [27], R is a full shift. This is needed to prove a "no short return" property akin to (NSR(x^*)), which for us, is one of the assumptions. This allows different choices of R, namely, for us, R need not be hyperbolic or even mixing. This is the main novelty of our approach.

4. **Examples.** The definitions of the maps in Examples 4.1, 4.5, and Lemma 4.4 are given in §11. For most of the examples we present, the choice of R is more interesting than the choice of T, mostly because (MEM) implies chaotic behavior and so the PLT in that setting is not surprising. We will thus not focus too much on T for this section. We only present some examples here, there are many others one can verify using Theorem 3.2.

Example 4.1. Let T be a map satisfying (MEM) on a manifold of sufficiently high[7] dimension then

(i) if R is a Diophantine rotation, then $T \times R$ satisfies the PLT everywhere;
(ii) if R is the time 1 map of a horocycle flow on $\Gamma \backslash SL_2(\mathbb{R})$ where Γ is a cocompact lattice, then $T \times R$ satisfies the PLT everywhere;
(iii) if R is a skew-shift, then $T \times R$ satisfies the PLT everywhere.

Remark 4.2. (i) In §11 we will show that the map R from example 4.1(i)-(iii) satisfies (EE) and (SLR(y^*)) for every y^*. The conclusion then follows from Theorem 3.2.
(ii) The PLT **almost everywhere** can be shown more readily. By Corollary 3.3, we just have to check (EE) for the map R, which holds for a big class of maps, examples will be given in §11.5.

Example 4.3. At this point, let us point out that, while T and R act on manifolds and preserve smooth measure, T and R themselves need not be smooth maps (not even continuous). For example, if $T : (0, 1] \to (0, 1]$ is the Gauss map (or a more general mixing, expanding interval map as in [39])

$$T(x) = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor,$$

then T preserves the density $\rho(x) = \frac{1}{\log 2 \cdot (1+x)}$. It is standard to show that T is multiple exponentially mixing, in fact, a fortiori, the Perron-Frobenius transfer operator has a spectral gap. Now if R_α is an irrational rotation on $\mathbb{T}$, and α is of bounded type, then, as shall be demonstrated in §11, Theorem 3.2 applies to show that $T \times R_\alpha$ satisfies the PLT everywhere.

Theorem 3.4 can be used to construct T, T^{-1} transformations of zero entropy that satisfy the PLT. All that remains to do is to construct a τ satisfying (BA), this can be done with the construction given in [14, Proposition 3.9].

Lemma 4.4. *Let $R_\alpha : \mathbb{T}^{d'} \to \mathbb{T}^{d'}$ be a Diophantine rotation, i.e*

$$|\langle k, \alpha\rangle - l| > C|k|^{-\lambda} \quad \forall k \in \mathbb{Z}^{d'}, k \neq 0, l \in \mathbb{Z}, \tag{D}$$

for some $\lambda \geq d'$. For $\frac{n}{2} < \rho < d'$ there is a $d \geq 1$ and a function $\tau \in C^\rho(\mathbb{T}^{d'}, \mathbb{R}^d)$ such that $\nu(\tau) = 0$, while

$$\nu(||\tau_n|| < \log^2(n)) = o(n^{-5}).$$

Note that in order to apply Theorem 3.4 we can always make d as big as we want.

Example 4.5. Let $R = R_\alpha$ be a Diophantine rotation with $d' = 2$ and $\lambda = 2$, τ be the function from Lemma 4.4, and let G be the Weyl Chamber flow on $SL(d, \mathbb{R})/\Gamma$, where Γ is a uniform lattice. If $d > 2$ then $S(x, y) = (G_{\tau(y)}(x), R_\alpha(y))$ satisfies the PLT everywhere.

[7] Sufficient bounds are given in §11.

5. **The delayed PLT.** The main step in the proof will be to show a generalised version of the PLT (for fiberwise returns), along a subsequence, this is what we will call a delayed PLT.

This 'delayed PLT' in itself is of independent interest, so let us make a more general statement.

Definition 5.1. Let $(X, \mathcal{A}, \mu, T)$ be a probability-preserving ergodic dynamical system and $\alpha = (\alpha^{(n)})_{n\geq 1}$ be a sequence of natural numbers, we will refer to α as the *delay sequence*, and denote $\tilde{\alpha}^{(n)} = \sum_{j=1}^{n} \alpha^{(n)}$. For measurable $A \subset X$ we define the *delayed consecutive return times to A along α* as

$$\begin{aligned} &\varphi_{A,\alpha}(x) := \varphi_{A,\alpha}^{(1)}(x) := \min(j \geq 1 \mid T^{\tilde{\alpha}^{(j)}}(x) \in A) \\ &\varphi_{A,\alpha}^{(n+1)}(x) := \min(j \geq 1 \mid T^{\tilde{\alpha}^{(\varphi_{A,\alpha}^{(1)}(x)+\varphi_{A,\alpha}^{(2)}(x)+\cdots+\varphi_{A,\alpha}^{(n)}(x)+j)}}(x) \in A) \\ &\Phi_{A,\alpha} := (\varphi_{A,\alpha}^{(1)}, \varphi_{A,\alpha}^{(2)}, \ldots). \end{aligned} \tag{11}$$

Like for classical return times, we will consider delayed return times for different systems. In an attempt to keep notation simple, we will not specify the underlying map, only the target set A.

The main example in this paper will be $\alpha^{(n)} = \varphi_B(R_B^n(y))$ for some $y \in Y$ and $B \subset Y$[8], however other choices are of interest, for example $\alpha^{(n)} = g(n)$, where g is a polynomial, or $\alpha^{(n)} = p_n$, where p_n is the nth prime.

Given a rare sequence $(A_l)_{l\geq 1}$, we will distinguish between two cases

1) the delay sequence is fixed in l,
2) the delay sequence is allowed to vary with l.

Definition 5.2. (i) Let $x^* \in X$ and, for $r > 0$, denote by $B_r(x^*)$ the geodesic ball of radius r centred at x^*. Let $\alpha = (\alpha^{(n)})_{n\geq 1}$ be a sequence of natural numbers. We will say that T *satisfies the delayed PLT along α at x^** if

$$\mu(B_r(x^*))\Phi_{B_r(x^*),\alpha} \overset{\mu}{\Rightarrow} \Phi_{Exp} \quad \textit{as } r \to 0.$$

(ii) Let $PLT(\alpha) = \{x^* \in X \mid T \text{ satisfies the delayed PLT along } \alpha \text{ at } x^*\}$.
If $x^* \in PLT(\alpha)$ for all sequences of natural numbers α, then we say that T *satisfies the delayed PLT at x^**.

The analogous definition for varying α is

Definition 5.3. (i) Let $x^* \in X$ and, for $r > 0$, denote by $B_r(x^*)$ the geodesic ball of radius r centred at x^*. Let $\alpha = ((\alpha_r^{(n)})_{n\geq 1})_{r>0}$ be a collection of sequences of natural numbers. We will say that T *satisfies the varying delayed PLT along α at x^** if

$$\mu(B_r(x^*))\Phi_{B_r(x^*),\alpha_r} \overset{\mu}{\Rightarrow} \Phi_{Exp} \quad \textit{as } r \to 0.$$

(ii) Let $PLT(\alpha) = \{x^* \in X \mid T \text{ satisfies the varying delayed PLT along } \alpha \text{ at } x^*\}$. If $x^* \in PLT(\alpha)$ for all sequences of natural numbers α, then we say that T *satisfies the delayed PLT at x^**.

In case 1) the main result is a straightforward modification of Theorems 3.2 and 3.4.

---

[8]As in §7, see especially (13).

Theorem 5.4. *Let α be a sequence of positive integers. Assume the conditions of Theorem 3.2 or 3.4 hold replacing (EE) in both cases by*

$$\left|\left|\sum_{j=0}^{n-1} f \circ R^{\tilde{\alpha}^{(j)}} - n \int f \, d\nu \right|\right|_{L^2(\nu)} \leq C||f||_{C^{r'}} n^\delta \quad \forall f \in C^{r'}, \, \forall n \geq 1.$$

Then S satisfies the delayed PLT along α at (x^*, y^*).

The proof is analogous to the proof of Theorem 3.2 resp. 3.4 (with Φ_{B_l} replaced by $\Phi_{B_l,\alpha}$ and Lemma 7.1 replaced by Lemma 12.1). Therefore, no detailed proof will be given [9].

In case 2) the main result is a special case of Proposition 8.4(III), however it is worth stating by itself.

Theorem 5.5. *Suppose T satisfies conditions (MEM), (SLR(x^*)) and (NSR(x^*)), then T satisfies the varying delayed PLT at x^*.*

Part 2. Proofs

6. **Cumulative return times.** Most of the statements (and proofs) below are much more convenient to state in terms of cumulative return times. For a measurable set $A \subset X$ (or $B \subset Y, C \subset X \times Y$) the *sequence of cumulative return times* to A is given by[10]

$$\sigma_A^{(n)} = \sum_{j=0}^{n-1} \varphi_A \circ T_A^j, \quad n \geq 1$$

$$\Sigma_A = (\sigma_A^{(1)}, \sigma_A^{(2)}, \dots),$$

and similar notation for delayed returns.

When studying distributional convergence of Φ_A, one can equivalently study for distributional convergence of Σ_A.

Indeed, let $\iota : [0, \infty)^{\mathbb{N}} \to [0, \infty)^{\mathbb{N}}$ be the map

$$\iota(x_1, x_2, x_3, ...) = (x_1, x_1 + x_2, x_1 + x_2 + x_3, \dots).$$

Since ι is a homeomorphism, standard theory shows that

$$\mu(A_l)\Phi_{A_l} \overset{\mu}{\Rightarrow} \Phi \;\; \textit{if and only if} \;\; \mu(A_l)\Sigma_{A_l} \overset{\mu}{\Rightarrow} \iota(\Phi), \tag{12}$$

where we use the obvious extension of ι to $[0, \infty]^{\mathbb{N}}$. Denote $\Sigma_{Exp} = \iota(\Phi_{Exp})$.

7. **Idea of the proof.** Let us first outline the strategy of proving Theorem 3.2.

Say we want to show the PLT for the geodesic balls (Q_l) converging to (x^*, y^*), then, for every $\varepsilon > 0$, we find suitable rectangles[11] approximating Q_l in the sense that $\bigcup_{k=1}^K A_l^{(k)} \times B_l^{(k)} \subset Q_l$ and

$$(\mu \times \nu)_{Q_l} \left(Q_l \setminus \bigcup_{k=1}^K A_l^{(k)} \times B_l^{(k)} \right) < \varepsilon.$$

---

[9] Note however that there is no relation of delayed hitting times to delayed return times as in Theorem 2.2.

[10] Recall from Definition 2.1 that $\varphi_A(x) = \min(n \geq 1 \mid T^n(x) \in A)$ is the first entry/return time to A, and $T_A = T^{\varphi_A}$ is the first return map.

[11] Using the exponential map, this is a simple exercise in $\mathbb{R}^{d+d'}$. Here we use continuity of the densities of μ resp. ν!

Denote $Q'_l := \bigcup_{k=1}^K A_l^{(k)} \times B_l^{(k)}$. We will use Proposition 8.4 to show the PLT for Q'_l, i.e

$$(\mu \times \nu)(Q'_l)\Sigma_{Q'_l} \overset{(\mu\times\nu)}{\Rightarrow} \Sigma_{Exp} \quad \textit{as } l \to \infty.$$

Letting $\varepsilon \to 0$, the next Lemma will help us conclude that

$$(\mu \times \nu)(Q_l)\Sigma_{Q_l} \overset{(\mu\times\nu)}{\Rightarrow} \Sigma_{Exp} \quad \textit{as } l \to \infty.$$

The critical approximation here is the following Lemma - see [42, Theorem 4.4] for a qualitative version, and the following statement can be easily deduced from the proof - for convenience we give a proof of this quantitative version in §12.

Lemma 7.1. *There is a metric D defined on the space of probability measures on $[0,\infty]^{\mathbb{N}}$ and modelling weak convergence of measures*[12] *such that*

$$D\left(law_{(\mu\times\nu)_Q}((\mu \times \nu)(Q)\Phi_Q), law_{(\mu\times\nu)_{Q'}}((\mu \times \nu)(Q')\Phi_{Q'}\right) \le 4(\mu \times \nu)_Q(Q \setminus Q'),$$

for measurable $Q' \subset Q \subset X \times Y$.

For returns to rectangles, say $A_l \times B_l$, for fixed $y \in Y$, we can ignore all the times j where $R^j(y) \notin B_l$. Hence, denoting $\alpha_l(y) = \Phi_{B_l}(y)$, we first show that for ν-a.e $y \in Y$

$$\mu(A_l)\Sigma_{A_l,\alpha_l(y)} \overset{\mu}{\Rightarrow} \Sigma_{Exp} \quad \textit{as } l \to \infty. \tag{13}$$

To show this, the idea is that, due to assumption (BR), the times[13] $\tilde{\alpha}_l^{(j)}(y)$ are sufficiently far apart to use (MEM), we apply Proposition 8.4. Since (13) is now true for ν-a.e $y \in Y$, we also have

$$\mu(A_l)\Sigma_{A_l,\Phi_{B_l}} \overset{\mu\times\nu}{\Rightarrow} \Sigma_{Exp} \quad \textit{as } l \to \infty, \tag{14}$$

where $\Sigma_{A_l,\Phi_{B_l}}(x,y) = \Sigma_{A_l,\Phi_{B_l}(y),y}(x)$.

These are not quite the returns that we wanted to consider. However, notice that, in every step, we skip exactly φ_{B_l} steps, thus we have the following relation

$$\sigma_{A_l\times B_l}(x,y) = \sum_{j=0}^{\sigma_{A_l,\Phi_{B_l}(y),y}(x)-1} \varphi_{B_l} \circ R^j_{B_l}(y). \tag{15}$$

We now use (EE) to control the ergodic sums of φ_{B_l} and show that

$$\mu(A_l)\nu(B_l)\Sigma_{A_l\times B_l} \overset{\mu\times\nu}{\Rightarrow} \Sigma_{Exp} \quad \textit{as } l \to \infty.$$

The same idea works for unions of rectangles $\bigcup_{k=1}^K A_l^{(k)} \times B_l^{(k)}$, however, in (13), we will have to slightly modify the definition of return times, which shall be done at the beginning of §8.2.

8. **PLT along varying subsequences.**

8.1. **Approximation.**

In our work, we often need to apply the mixing condition (MEM), and the quantitative ergodicity (EE), for indicator functions, hence we have to approximate them by functions in C^r resp. $C^{r'}$.

Definition 8.1. Let M be a C^r manifold with dimension $dim(M) = d$, λ be a measure on M.

---

[12] In the sense that $\lambda_l \Rightarrow \lambda$ if and only if $D(\lambda_l, \lambda) \to 0$.

[13] It can be recalled from Definition 5.1 that $\tilde{\alpha}_l^{(n)} = \sum_{j=1}^n \alpha_l^{(j)} = \sum_{j=0}^{n-1} \varphi_{B_l} \circ R^j$.

Let $B_\alpha \subset M$ be measurable subsets for α in some index set. We say that $\{B_\alpha\}$ is *regularly approximable in* C^r if there is a constant $\mathcal{B} > 0$ such that, for each α and for $0 < \varepsilon < \frac{1}{10}\lambda(B_\alpha)^{\frac{1}{d}}$, there are $\overline{h}, \underline{h} \in C^r$ with $\underline{h} \leq 1_{B_\alpha} \leq \overline{h}$ and for $h \in \{\overline{h}, \underline{h}\}$

$$\lambda(1_{B_\alpha} \neq h) \leq \lambda(B_\alpha)^{\frac{d-1}{d}}\varepsilon, \quad \textit{while} \quad ||h||_{C^r} \leq \mathcal{B}\varepsilon^{-r}. \tag{16}$$

We call the least such constant $\mathcal{B} > 0$ the *approximant of* B and denote it by $\mathrm{app}(\{B_\alpha\})$.

Lemma 8.2. *Let M, r, d and λ be as in Definition 8.1. Assume in addition that λ is absolutely continuous w.r.t volume with bounded density. Suppose that $\bigcup_\alpha B_\alpha$ is relatively compact, and there is an open $U \supset \overline{\bigcup_\alpha B_\alpha} \supset \bigcup_\alpha B_\alpha$ and a C^r-diffeomorphism $\iota : U \to V$ for some open set $V \subset \mathbb{R}^d$ such that each $\iota(B_\alpha)$ is a ball. Then $\{B_\alpha\}$ is regularly approximable in C^r.*

Proof. (i) We may assume that $M = \mathbb{R}^d$ and λ is Lebesgue measure, otherwise we pick up another constant, which can be absorbed into C. Furthermore, we can assume that all B_α are balls centered at the origin. In the following fix α and, dropping the α from our notation, let $B = B_t(0)$. We have $\lambda(B) = C_1 t^d$, where C_1 is the volume of the d-dimensional unit ball.

(ii) Let $\theta : \mathbb{R} \to [0,1]$ be a smooth function with $\theta(x) = 1$ if $x < 0$, and $\theta(x) = 0$ if $x > 1$. For $t > \varepsilon > 0$ consider

$$\hat{\theta}(x) = \theta(\varepsilon^{-1}(x - t))$$

then $\hat{\theta}$ is still smooth and $||\hat{\theta}||_{C^r} = \varepsilon^{-r}||\theta||_{C^r}$. Consider $\overline{h} : \mathbb{R}^d \to [0,1]$ given by $\overline{h}(x) = \hat{\theta}(|x|)$, then

- $\overline{h}$ is smooth, away from the origin because it is the composition of smooth functions, and near the origin, it is constant 1,
- $\overline{h}(x) = 1$ if $|x| < t$ and $\overline{h}(x) = 0$ if $|x| > t + \varepsilon$,
- and $||\overline{h}||_{C^r} \leq C_3\varepsilon^{-r}$ where $C_3 = rC_2||\theta||_{C^r}$ and C_2 is the C^r norm of the smooth function $x \mapsto |x|$ on $\{t \leq |x| \leq t + \varepsilon\}$.

Furthermore, we have

$$\begin{aligned} \lambda(\overline{h} \neq 1_B) &= \lambda(t \leq |x| \leq t + \varepsilon) = C_1((t+\varepsilon)^d - t^d) \\ &\leq C_1 d t^{d-1}\varepsilon \leq dC_1^{\frac{1}{d}}\lambda(B)^{\frac{d-1}{d}}\varepsilon, \end{aligned}$$

all the constants can be absorbed in the constant C from the claim, the constant only depends on r, d.

For $\underline{h}$ repeat the calculations with $\hat{\theta} = \theta(\varepsilon^{-1}(x - (t - \varepsilon)))$ instead. □

8.2. **Proof of the PLT along varying subsequences.** For our purposes it will not be enough to consider the delayed PLT for a single rare sequence $(A_l)_{l\geq 1}$, rather let $\mathcal{K} \geq 1$, and $A_l^{(1)}, ..., A_l^{(\mathcal{K})}$ be subsets of X such that $\{A_l^{(k)}\}$ is regularly approximable in C^r. Assume that there are $\omega^{(1)}, ..., \omega^{(\mathcal{K})} > 0$ and $r_l \to 0$ such that

$$\mu(A_l^{(k)}) = \omega^{(k)} r_l^d + o(r_l^d). \tag{17}$$

Given $\kappa_l^{(j)} \in \{1, ..., \mathcal{K}\}$, for $l, j \geq 1$, define the *cumulative return times* by

$$\begin{aligned} &\sigma^{(1)}_{\kappa_l,\alpha_l,y}(x) = \min(j \geq 1 \mid T_y^{\tilde{\alpha}_l^{(j)}}(x) \in A_l^{(\kappa_l^{(j)})}), \\ &\sigma^{(n+1)}_{\kappa_l,\alpha_l,y}(x) = \min(j \geq \tau^{(n)}_{\kappa_l,\alpha_l,y}(x) + 1 \mid T_y^{\tilde{\alpha}_l^{(j)}}(x) \in A_l^{(\kappa_l^{(j)})}), \\ &\Sigma_{\kappa_l,\alpha_l,y} = (\sigma^{(1)}_{\kappa_l,\alpha_l,y}, \sigma^{(2)}_{\kappa_l,\alpha_l,y}, ...). \end{aligned}$$

Denote the frequency with which $A_l^{(\kappa_l)} = A_l^{(k)}$ by

$$p_{l,t}^{(k)} := \frac{1}{t}\#\{j = 1, ..., t_l \mid \kappa_l^{(j)} = k\}.$$

For now suppose that there are positive constants $\theta^{(k)} > 0$ such that, for all $k = 1, ..., \mathcal{K}$, and for all $t_l \nearrow \infty$ with $t_l = O(r_l^{-d})$,

$$p_{l,t_l}^{(k)} \to \frac{\theta^{(k)}}{\sum_{j=1}^{\mathcal{K}} \theta^{(j)}} =: p^{(k)} \quad \textit{as } l \to \infty. \tag{18}$$

Later on, when we prove PLT in product systems, we will choose κ and θ in a specific way[14] and (18) will be satisfied by Lemma 9.4.

The main estimate of mixing rates for regularly approximable sets is the following.

Lemma 8.3. *Suppose T satisfies (MEM), and let $y \in Y$ be such that (1) is satisfied. Let $m \geq 1$, $A^{(1)}, ..., A^{(k)} \subset X$ be regularly approximable in C^r, and $1 \leq n_1 < ... < n_k$, then*

$$\left|\mu\left(\bigcap_{i=1}^k T_y^{-n_i} A^{(i)}\right) - \prod_{i=1}^k \mu(A^{(i)})\right| \leq K \max_{i=1,...,k} \mu(A^{(i)})^{\frac{d-1}{d}\frac{kr}{kr+1}} e^{-\frac{\gamma p}{kr+1}},$$

where $p = \min_{i=1,...,k-1} |n_{i+1} - n_i|$ and the constant $K > 0$ only depends on k and app$(\{A^{(1)}, ..., A^{(k)}\})$.

Proof. Let $C = \text{app}(\{A^{(1)}, ..., A^{(k)}\})$. By Lemma 8.2, for every $\varepsilon > 0$, there are $h^{(i)} \in C^r$ such that $0 \leq h^{(i)} \leq 1_{A^{(i)}}$ and

$$\mu(1_{A^{(i)}} \neq h^{(i)}) \leq \mu(A^{(i)})^{\frac{d-1}{d}} \varepsilon, \quad \textit{while } ||h^{(i)}||_{C^r} < C\varepsilon^{-r}.$$

We estimate

$$\begin{aligned}
\left|\mu\left(\bigcap_{i=1}^k T^{-n_i} A^{(i)}\right) - \prod_{i=1}^k \mu(A^{(i)})\right| &\leq \left|\mu\left(\bigcap_{i=1}^k T^{-n_i} A^{(i)}\right) - \int_X \prod_{i=1}^k h^{(i)} \circ T^{n_i}\, \mathrm{d}\mu\right| \\
&+ \left|\int_X \prod_{i=1}^k h^{(i)} \circ T^{n_i}\, \mathrm{d}\mu - \prod_{i=1}^k \int_X h^{(i)} \circ T^{n_i}\, \mathrm{d}\mu\right| \\
&+ \left|\prod_{i=1}^k \int_X h^{(i)} \circ T^{n_i}\, \mathrm{d}\mu - \prod_{i=1}^k \mu(A^{(i)})\right| \\
&\leq 4k \max_{i=1,...,k} \mu(A^{(i)})^{\frac{d-1}{d}} \varepsilon + C^k e^{-\gamma p} \varepsilon^{-kr}.
\end{aligned}$$

This bound is optimised for

$$\varepsilon^* = \left(\frac{C^k r}{4k} \max_{i=1,...,k} \mu(A^{(i)})^{-\frac{d-1}{d}} e^{-\gamma p}\right)^{\frac{1}{kr+1}},$$

so

$$\begin{aligned}
\left|\mu\left(\bigcap_{i=1}^k T^{-n_i} A^{(i)}\right) - \prod_{i=1}^k \mu(A^{(i)})\right| &\leq \hat{C} \max_{i=1,...,k} \mu(A^{(i)})^{\frac{d-1}{d}\left(1-\frac{1}{kr+1}\right)} e^{-p\frac{,\gamma}{kr+1}} \\
&+ \bar{C} \max_{i=1,...,k} \mu(A^{(i)})^{\frac{kr}{kr+1}\frac{d-1}{d}} e^{-\gamma p\left(1-\frac{kr}{kr+1}\right)}
\end{aligned}$$

[14] Say we want to prove a PLT for the system $T \times R$ and sets of the form $B_r(x^*) \times B_r(y^*)$ then we choose $\mathcal{K} = 1$ and $\theta > 0$ such that $\nu(B_r(y^*)) = \theta r^{d'} + o(r^{d'})$.

$$\leq K \max_{i=1,\ldots,k} \mu(A^{(i)})^{\frac{d-1}{d}\frac{kr}{kr+1}} e^{-\frac{\gamma}{kr+1}p},$$

where the constants $\hat{C}, \bar{C}, K > 0$ only depend on k and on C. □

In the following Proposition, we shall demonstrate exponential limit distributions for fiberwise return times (compare (13)). The proof is standard and uses multiple exponential mixing and the method of moments, however, due to having multiple sets, the notations are quite cumbersome. We will present a proof for the convenience of the reader.

Proposition 8.4. *Suppose that T satisfies (MEM), $(A_l^{(1)}), \ldots, (A_l^{(\mathcal{K})})$ are sequences of rare events with $\mu(A_l^{(k)}) = \omega^{(k)} r_l^d + o(r_l^d)$, for $\mathcal{K} \geq 1$, κ_l satisfy (18) for some $p^{(k)} > 0$ with $\sum_{k=1}^{\mathcal{K}} p^{(k)} = 1$, and let $\alpha_l = (\alpha_l^{(n)})_{n\geq 1}$ be sequences of natural numbers. Denote $A_l = \bigcup_{k=1}^{\mathcal{K}} A_l^{(k)}$ and suppose that either*

(I) α_l grows faster than $|\log \mu(A_l)|$ in the sense that

$$|\log \mu(A_l)| = o(\min_{n\geq 2} |\alpha_l^{(n)}|), \tag{19}$$

(II) short returns to A_l are rare in the sense that

$$\mu_{A_l}(\varphi_{A_l} \leq a_l) \to 0 \quad \text{as } l \to \infty, \tag{20}$$

for some sequence $(a_l)_{l\geq 1}$ with $|\log(\mu(A_l))| = o(a_l)$, and α_l grows at least as fast $|\log \mu(A_l)|$ in the sense that

$$|\log \mu(A_l)| = O(\min_{n\geq 2} |\alpha_l^{(n)}|), \tag{21}$$

(III) or short returns to A_l are rare in the sense that

$$\mu_{A_l}(\varphi_{A_l} \leq a_l) \to 0 \quad \text{as } l \to \infty, \tag{22}$$

for some sequence $(a_l)_{l\geq 1}$ with $|\log(\mu(A_l))| = o(a_l)$, and returns are at least logarithmically large, i. e there exists $c > 0$ such that

$$\varphi_{A_l}(x) \geq c \log(\mu(A_l)) \quad \mu - a.e\ x \in A_l. \tag{23}$$

Then for ν-a.e $y \in Y$

$$\Omega r_l^d \Sigma_{\kappa_l,\alpha_l,y} \stackrel{\mu}{\Rightarrow} \Sigma_{Exp} \quad \text{as } l \to \infty,$$

where $\Omega = \sum_{k=1}^{\mathcal{K}} \omega^{(k)} p^{(k)}$.

Proof. Fix y as in (MEM) and denote $T^n = T_y^n$.

(i) Taking a subsequence if necessary, we may assume that there is a $[0,\infty]$-valued process $\Sigma = (\sigma^{(1)}, \sigma^{(2)}, \ldots)$ such that

$$\Omega r_l^d \Sigma_{\kappa_l,\alpha_l} \stackrel{\mu}{\Rightarrow} \Sigma \quad \text{as } l \to \infty.$$

For $J \geq 1$ and $0 < t_1 < \ldots < t_J$ we show

$$\mathbb{P}\left(\sigma^{(j)} \leq t_j,\ j = 1, \ldots, J\right) = \mathbb{P}\left(\sigma_{Exp}^{(j)} \leq t_j,\ j = 1, \ldots, J\right). \tag{24}$$

The trick is to look at the "dual" object

$$S_{\kappa_l,\alpha_l}^{(n)} = \sum_{j=1}^{n} 1_{A_l^{\kappa_l^{(j)}}} \circ T^{\tilde{\alpha}_l^{(j)}}.$$

The important relation here is the following

$$S_{\kappa,\alpha_l}^{(n)} \geq N \iff \sigma_{\kappa,\alpha_l}^{(N)} \leq n. \tag{25}$$

Indeed, evaluating both sides at some $x \in X$, the left side says that there are at least N different times $1 \le j_1 < ... < j_N \le n$ such that $T^{\tilde{\alpha}_l^{(j)}}(x) \in A_l^{\kappa_l^{(j)}}$. The right-hand side expresses that, if $1 \le j_1 < ... < j_N$ are the first N times that $T^{\tilde{\alpha}_l^{(j)}}(x) \in A_l^{\kappa_l^{(j)}}$, then $j_1 + \sum_{i=2}^{N}(j_i - j_{i-1}) \le n$. Let $(P_t)_{t\ge0}$ be a Poisson process, such that Σ_{Exp} are the cumulative waiting times of (P_t), i.e. (P_t) and Σ_{Exp} are related by

$$P_t \ge N \iff \sigma_{Exp}^{(N)} \le t.$$

The right side of (24) is equal to

$$\mathbb{P}\left(\sigma_{Exp}^{(k)} \le t_k,\ k = 1, ..., J\right) = \mathbb{P}(P_{t_k} \ge k,\ k = 1, ..., J).$$

Due to (25) it is enough to show

$$\left(S_{\kappa_l,\alpha_l}^{\left\lfloor \frac{t_1}{\Omega r_l^d}\right\rfloor}, S_{\kappa_l,\alpha_l}^{\left\lfloor \frac{t_2}{\Omega r_l^d}\right\rfloor}, ..., S_{\kappa_l,\alpha_l}^{\left\lfloor \frac{t_J}{\Omega r_l^d}\right\rfloor}\right) \overset{\mu}{\Rightarrow} (P_{t_1}, ..., P_{t_J}) \quad \textit{as } l \to \infty.$$

(ii) Taking a further subsequence if necessary, there are $[0, \infty]$-valued $\tilde{P}_{t_1}, ..., \tilde{P}_{t_J}$ such that

$$\left(S_{\kappa_l,\alpha_l}^{\left\lfloor \frac{t_1}{\Omega r_l^d}\right\rfloor}, S_{\kappa_l,\alpha_l}^{\left\lfloor \frac{t_2}{\Omega r_l^d}\right\rfloor}, ..., S_{\kappa_l,\alpha_l}^{\left\lfloor \frac{t_J}{\Omega r_l^d}\right\rfloor}\right) \overset{\mu}{\Rightarrow} (\tilde{P}_{t_1}, ..., \tilde{P}_{t_J}) \quad \textit{as } l \to \infty.$$

We will show that

(A) $\tilde{P}_{t_k} - \tilde{P}_{t_{k-1}}$ is Poisson distributed with intensity $t_k - t_{k-1}$ for $k = 1, ..., d$,

(B) and $(\tilde{P}_{t_1} - \tilde{P}_{t_0}, \tilde{P}_{t_2} - \tilde{P}_{t_1}, ..., \tilde{P}_{t_J} - \tilde{P}_{t_{J-1}})$ is an independent vector, where $t_0 = 0$.[15]

Clearly $P_0 = \tilde{P}_0 = 0$.

For $j = 1, \ldots, J$ denote $S_{t_j,l} = S_{A_l,\kappa_l}^{\left\lfloor \frac{t_j}{\Omega r_l^d}\right\rfloor}$. Assertions (A) and (B) will follow[16] once we show that, for all $m_1, ..., m_d \ge 1$,

$$\int_X \prod_{j=1}^{J} \binom{S_{t_j,l} - S_{t_{j-1},l}}{m_j} \mathrm{d}\mu = \prod_{j=1}^{J} \frac{(t_j - t_{j-1})^{m_j}}{m_j!} + o(1) \quad \textit{as } l \to \infty.$$

In the rest of the proof fix $J \ge 1$, $0 = t_0 < t_1 < \cdots < t_J$ and $m_1, \ldots, m_J \ge 1$.

(iii) First, for each $j = 1, \ldots, J$, rewrite

$$S_{t_j,l} - S_{t_{j-1},l} = \sum_{i=\left\lfloor \frac{t_{j-1}}{\Omega r_l^d}\right\rfloor+1}^{\left\lfloor \frac{t_j}{\Omega r_l^d}\right\rfloor} 1_{A_l^{(\kappa_l^{(i)})}} \circ T^{\tilde{\alpha}_l^{(i)}}.$$

So

$$\prod_{j=1}^{J} \binom{S_{t_j,l} - S_{t_{j-1},l}}{m_j} = \prod_{j=1}^{d} \sum_{\left\lfloor \frac{t_{j-1}}{\Omega r_l^d}\right\rfloor+1 \le k_{1,j} < \ldots < k_{m_j,j} \le \left\lfloor \frac{t_j}{\Omega r_l^d}\right\rfloor} \prod_{i=1}^{m_j} \xi_{i,j} \tag{26}$$

[15] This is essentially Watanabe's characterisation of Poisson-processes.

[16] Here we apply the method of moments, see eg [5, Theorem 30.2].

where $\xi_{i,j} = 1_{A_l^{(\kappa_l^{(k_{i,j})})}} \circ T^{\tilde{\alpha}_l^{(k_{i,j})}}$. To simplify notation we will also denote $m = (m_1 + \cdots + m_J)$, $\omega = \min_{k=1,\ldots,\mathcal{K}} \omega^{(k)}$, $p_l = 2m\frac{mr+1}{\gamma}\left|\log(\omega r_l^d)\right|$,

$$\Delta_l := \{\mathbf{k} = (k_{i,j})_{\substack{j=1,\ldots,J \\ i=1,\ldots,m_j}} \mid \left\lfloor \frac{t_{j-1}}{\Omega r_l^d} \right\rfloor + 1 \le k_{1,j} < \ldots < k_{m_j,j} \le \left\lfloor \frac{t_j}{\Omega r_l^d} \right\rfloor \quad \textit{for } j = 1, \ldots, J\}$$

and

$$\Delta_l' := \{\mathbf{k} \in \Delta_l \mid \min_{\substack{j=1,\ldots,J,\ i=1,\ldots,m_j \\ j'=1,\ldots,J,\ i'=1,\ldots,m_{j'},\ (j,i)\neq(j',i')}} |\tilde{\alpha}_l^{(k_{i,j})} - \tilde{\alpha}_l^{(k_{i',j'})}| \le p_l\}.$$

We will split the sum in (26) into two terms

$$\prod_{j=1}^{J} \binom{S_{t_j,l} - S_{t_{j-1},l}}{m_j} = M_l + R_l,$$

where

$$M_l = M_{(t_j),l,(m_j)} = \sum_{\mathbf{k}\in\Delta_l\setminus\Delta_l'} \prod_{\substack{j=1,\ldots,J \\ i=1,\ldots,m_j}} \xi_{i,j}, \quad R_l = R_{(t_j),l,(m_j)} = \sum_{\mathbf{k}\in\Delta_l'} \prod_{\substack{j=1,\ldots,J \\ i=1,\ldots,m_j}} \xi_{i,j}.$$

We will show that

$$\int_X M_l \,\mathrm{d}\mu \to \prod_{j=1}^{J} \frac{(t_j - t_{j-1})^{m_j}}{m_j!} \quad \textit{and} \quad \int_X R_l \,\mathrm{d}\mu \to 0 \quad \textit{as } l \to \infty. \tag{27}$$

(iv) Let us first treat M_l. For $l \ge 1$ and $\mathbf{k} \in \Delta_l$, by Lemma 8.3 we have

$$\left| \mu\left(\bigcap_{\substack{j=1,\ldots,J \\ i=1,\ldots,m_j}} T^{-\tilde{\alpha}_l^{(k_{i,j})}} A_l^{(\kappa_l^{(k_{i,j})})} \right) - \prod_{j,i} \mu\left(A_l^{(\kappa_l^{(k_{i,j})})}\right) \right|$$
$$\le K \max_{j,i} \mu\left(A_l^{(\kappa_l^{(k_{i,j})})}\right)^{\frac{d-1}{d}\frac{mr}{mr+1}} e^{-\frac{\gamma \min_{i,j} \alpha_l^{(k_{i,j})}}{mr+1}}.$$

For $\mathbf{k} \in \Delta_l \setminus \Delta_l'$, this yields

$$\mu\left(\bigcap_{\substack{j=1,\ldots,J \\ i=1,\ldots,m_j}} T^{-\tilde{\alpha}_l^{(k_{i,j})}} A_l^{(\kappa_l^{(k_{i,j})})} \right) = r_l^{md} \prod_{\substack{j=1,\ldots,J \\ i=1,\ldots,m_j}} \omega_l^{(\kappa_l^{k_{i,j}})} + o(r_l^{md}),$$

and the o-term does not depend on $\mathbf{k}$. Summing over $\mathbf{k} \in \Delta_l \setminus \Delta_l'$, and using (18), yields

$$\int_X M_l \,\mathrm{d}\mu = \int_X \sum_{\mathbf{k}\in\Delta_l\setminus\Delta_l'} \prod_{\substack{j=1,\ldots,J \\ i=1,\ldots,m_j}} 1_{A_l^{(\kappa_l^{(k_{i,j})})}} \circ T^{\tilde{\alpha}_l^{(k_{i,j})}} \,\mathrm{d}\mu$$
$$= r_l^{md} \sum_{\mathbf{k}\in\Delta_l\setminus\Delta_l'} \prod_{\substack{j=1,\ldots,J \\ i=1,\ldots,m_j}} \omega^{(\kappa_l^{(k_{i,j})})} + o(1)$$
$$\overset{*}{=} r_l^{md} \sum_{\mathbf{k}\in\Delta_l} \prod_{\substack{j=1,\ldots,J \\ i=1,\ldots,m_j}} \omega^{(\kappa_l^{(k_{i,j})})} + o(1)$$

$$= r_l^{md} \prod_{j=1}^{J} \frac{1}{m_j!} \left(\sum_{k=1}^{\mathcal{K}} \left\lfloor \frac{t_j - t_{j-1}}{\Omega r_l^d} \right\rfloor p^{(k)}_{l,\left\lfloor \frac{t_j-t_{j-1}}{\Omega r_l^d} \right\rfloor} \omega^{(k)} \right)^{m_j} + o(1)$$

$$= \Omega^{-m} \prod_{j=1}^{J} \frac{(t_j - t_{j-1})^{m_j}}{m_j!} \left(\sum_{k=1}^{\mathcal{K}} p^{(k)} \omega^{(k)} \right)^{m_j} + o(1)$$

$$= \prod_{j=1}^{J} \frac{(t_j - t_{j-1})^{m_j}}{m_j!} + o(1),$$

for $*$ note that $\#\Delta_l' = O(r_l^{-md+1}|\log(r_l)|)$. This shows the first assertion of (27).

(v) In order to treat R_l, first note that under assumption (I) we have $R_l = 0$ for big enough l. In the following, we focus on assumptions (II) and (III). Note that

$$\int_X R_l \,\mathrm{d}\mu = \int_X 1_{R_l \neq 0} R_l \,\mathrm{d}\mu \leq \mu(\mathrm{supp}(R_l)) \|R_l\|_{L^2}$$

and

$$\mathrm{supp}(R_l) \subset \bigcup_{j=1}^{\left\lfloor \frac{t_J}{\Omega r_l^d} \right\rfloor} T^{-\tilde{\alpha}_l^{(j)}} \left(A_l \cap \{\varphi_{A_l} \leq 3p_l\}\right) =: U_l,$$

since $p_l = O(|\log(r_l)|)$, from (20) resp. (22), it follows that

$$\mu(\mathrm{supp}(R_l)) \leq \mu(U_l) = O(r_l^{-d})\mu(A_l)o(1) = o(1).$$

Therefore, in order to show $\int_X R_l \,\mathrm{d}\mu \to 0$, it is enough to show that $(R_l)_{l\geq 1}$ is bounded in L^2. Notice that

$$R_l^2 \leq \sum_{1 \leq k_1, \ldots, k_{2m} \leq \left\lfloor \frac{t_J}{\Omega r_l^d} \right\rfloor} \prod_{i=1}^{m} 1_{A_l^{(\kappa_l^{(k_i)})}} \circ T^{\tilde{\alpha}_l^{(i)}} = S_{t_J,l}^{2m}.$$

We may write

$$S_{t_J,l}^{2m} = \sum_{k=1}^{2m} \left\{ {2m \atop k} \right\} \binom{S_{t_J,l}}{k} \leq C_m \sum_{k=1}^{2m} \binom{S_{t_J,l}}{k} \leq C_m \sum_{k=1}^{2m} (M_{t_J,l,k} + R_{t_J,l,k}),$$

where $\left\{ {2m \atop k} \right\}$ are the Stirling numbers of the second kind and $C_m = \max_{k=1,\ldots,2m} \left\{ {2m \atop k} \right\}$. Now the previous parts of the proof show that

$$\int_X M_{t_J,l,m} \,\mathrm{d}\mu \quad \textit{is bounded as } l \to \infty, \forall m \geq 1,$$

it remains to show that

$$\int_X R_{t_J,l,m} \,\mathrm{d}\mu \quad \textit{is bounded as } l \to \infty, \forall m \geq 1.$$

(vi) In order to bound $\int_X R_{t_J,l,m} \,\mathrm{d}\mu$ for fixed $m \geq 1$ we first split up[17] Δ_l' into

[17] Here the sets of sequences should be modified, i.e

$$\Delta_l := \{\mathbf{k} = (k_i)_{i=1,\ldots,m} \mid 1 \leq k_1 < \ldots < k_m \leq \left\lfloor \frac{t_J}{\Omega r_l^d} \right\rfloor\}$$

and

$$\Delta_l' := \{\mathbf{k} \in \Delta_l \mid \min_{\substack{i=1,\ldots,m \\ i'=1,\ldots,m,\ i\neq i'}} |\tilde{\alpha}_l^{(k_i)} - \tilde{\alpha}_l^{(k_{i'})}| \leq p_l\}$$

$$\Delta_l^{(j)} := \left\{ \mathbf{k} = (k_1, ..., k_m) \in \Delta_l \;\middle|\; \begin{array}{l} \exists 1 \le i_i < ... < i_{m-j} \le m \textit{ such that} \\ |\tilde{\alpha}_l^{(k_{i+1})} - \tilde{\alpha}_l^{(k_i)}| \le \tilde{p}_l \; \forall i \in \{1, ..., m\} \setminus \{i_1, ..., i_{m-j}\}, \\ \textit{and } |\tilde{\alpha}_l^{(k_{i_r}+1)} - \tilde{\alpha}_l^{(k_{i_r})}| > \tilde{p}_l \; \forall r = 1, ..., m-j \end{array} \right\},$$

for $j = 1, ..., m-1$, so that $\Delta_l' = \bigcup_{j=1}^{m-1} \Delta_l^{(j)}$. Under assumption (II), because of (21), there is a constant $c > 0$ such that for $\rho = 2m\frac{mr+1}{c\gamma}$ and big enough l we have

$$\#\Delta_l^{(j)} \le \binom{m}{j} \binom{\left\lfloor \frac{t}{\Omega r_l^d} \right\rfloor}{m-j} \rho^j \le m^m t^m \rho^m (\Omega^{-m} + 1) r_l^{-d(m-j)}.$$

On the other hand, for $\mathbf{k} \in \Delta_l^{(j)}$, we can again use Lemma 8.3 to estimate

$$\int_X \prod_{s=1}^m \xi_s \, \mathrm{d}\mu \le \int_X \prod_{s \in \{1,...,m\} \setminus \{i_1,...,i_j\}} \xi_s \le \Omega r_l^{d(m-j)} + o(r_l^{d(m-j)}),$$

where the o-term does not depend on $\mathbf{k}$, and

$$\xi_s = 1_{A_l^{\kappa_l^{(k_s)}}} \circ T^{\tilde{\alpha}_l^{(k_s)}}.$$

Summing over $\mathbf{k} \in \Delta_l'$ we obtain

$$\int_X R_{t_j,l,m} \, \mathrm{d}\mu \le 2m^{m+1} t^m \rho^m (\Omega^{-m+1} + \Omega), \quad \forall m \ge 1,$$

and we conclude $\int_X R_l \, \mathrm{d}\mu \to 0$ under assumption (II).

(vii) Finally assume (III). For fixed $j = 1, ..., m-1$, $l \ge 1$ and $1 \le i_1 < ... < i_j \le \left\lfloor \frac{t}{\Omega r_l^d} \right\rfloor$ with $\min_{s=1,...,j-1} |\tilde{\alpha}_l^{(i_s+1)} - \tilde{\alpha}_l^{(i_s)}| \ge p_l$ set

$$A_{l,i_1,...,i_j} = \bigcap_{s=1}^j T^{-\tilde{\alpha}_l^{(i_s)}} A_l^{\kappa_l^{(i_s)}}.$$

By Lemma 8.3

$$\mu(A_{l,i_1,...,i_j}) \le (\omega r_l)^{dj} + o(r_l^{dj}),$$

and the o term doesn't depend on $(i_1, ..., i_j)$. For $x \in A_{l,i_1,...,i_j}$ consider

$$\mathfrak{K}_{l,i_1,...,i_j}(x) = \left\{ \mathbf{k} = (k_1, ..., k_m) \in \Delta_l^{(j)} \;\middle|\; \begin{array}{l} \exists r_1, ..., r_j \; \textit{ such that } k_{r_s} = i_s \; \forall s = 1, ..., j, \\ \textit{and } \prod_{s=1}^m \xi_s(x) = 1 \end{array} \right\}.$$

Then, since $\varphi_{A_l}(x) \ge c \log(\mu(A_l))$, we have

$$\#\mathfrak{K}_{l,i_1,...,i_j}(x) \le \left(2\frac{p_l}{dc \log(r_l)} \right)^{m-j} =: \rho^{m-j},$$

since $p_l = \textit{constant} * |\log(r_l)|$ this quantity doesn't depend on l. At the same time we have

$$\operatorname{supp} \left(\prod_{s=1}^m \xi_s \right) \subset A_{l,i_1,...,i_j} \quad \textit{if } \mathbf{k} \in \mathfrak{K}_{l,i_1,...,i_j}(x) \; \textit{ for some } x.$$

Also, every $\mathbf{k} \in \tilde{\Delta}_l^{(j)}$ is in some $\mathfrak{K}_{l,i_1,\ldots,i_j}(x)$, therefore

$$\begin{aligned}\sum_{\mathbf{k}\in\tilde{\Delta}_l^{(j)}} \int_X \prod_{s=1}^m \xi_s \,\mathrm{d}\mu &\le \sum_{\substack{1\le i_1<\ldots<i_j\le \left\lfloor \frac{t}{\Omega r_l^d}\right\rfloor \\ \min_{s=1,\ldots,j-1} |\tilde{\alpha}_l^{(i_s+1)}-\tilde{\alpha}_l^{(i_s)}|\ge p_l}} \int_{A_{l,i_1,\ldots,i_j}} \#\mathfrak{K}_{l,i_1,\ldots,i_j}(x)\,\mathrm{d}\mu(x) \\ &\le \sum_{\substack{1\le i_1<\ldots<i_j\le \left\lfloor \frac{t}{\Omega r_l^d}\right\rfloor \\ \min_{s=1,\ldots,j-1} |\tilde{\alpha}_l^{(i_s+1)}-\tilde{\alpha}_l^{(i_s)}|\ge \tilde{p}_l}} \int_{A_{l,i_1,\ldots,i_j}} \rho^j \,\mathrm{d}\mu \\ &\le \rho^j \left(\omega r_l^{dj} + o(r_l^{dj})\right)\left(\left\lfloor \frac{t}{\Omega r_l^d}\right\rfloor\right)^j \\ &\le \rho^j t^j \frac{\omega}{\Omega^j} + o(1).\end{aligned}$$

Summing up over j we get

$$\int_X R_l \,\mathrm{d}\mu \le m \max(1,\rho^m)\max(1,t^m)\frac{\omega}{\max(1,\Omega^m)} + o(1).$$

Following the argument in step (v) this shows $\int_X R_l \,\mathrm{d}\mu \to 0$, hence (27), in case of assumption (III). This concludes the proof. □

Remark 8.5. (i) Note that assumptions (I), (II), or (III) directly correspond to the three possible cases of (BR(x^*, y^*))–(SLR(y^*)), (NSR(x^*)) AND (LR(y^*)), or (NSR(x^*)) AND (LR'(x^*)) respectively.

(ii) In all of the examples we give in section 4, the α_l will satisfy a condition stronger than (I). In fact, in this set-up, there is a $\delta_2 > 0$ such that

$$\min_{i\ge 2} |\alpha_l^{(i)}| \ge \mu(A_l)^{-\delta_2},$$

compare also (9)[18]. If this stronger condition is satisfied instead of (I), then we do not need the full strength of exponential mixing in (MEM). Any superpolynomial rate will be enough. Details are given in the proof of Theorem 3.4, but we shall give a heuristic here.

When using mixing of all orders with indicators of the form $1_{B_{r_l}(x^*)}$, the error term will contain a term coming from the C^r norm in the definition of regularly approximable. In this case, using (16), this term will be of order Cr_l^{-drm}. To compensate, say the rate of mixing is ψ, since the gaps α_l are large we can multiply with $\psi(\min_{i\ge 2} |\alpha_l^{(i)}|)$. So we want to show

$$r_l^{-drm}\psi(r_l^{-d\delta_2}) = o(1),$$

for all $m \ge 1$, thus ψ should decay superpolynomially.

9. The PLT for rectangles.

9.1. Quantitative Ergodic Theorem.

In section 9.2, it will be convenient to use a pointwise (almost everywhere) Quantitative Ergodic Theorem instead of the L^2 bound we assume in (EE). Furthermore, using a Borel-Cantelli argument, we will show such bounds simultaneously along a sequence of functions $(f_l)_{l\ge 1}$.

[18] The constant δ_2 given there is not exactly the same. In the notation there, we have to use $\delta_2 \frac{d'}{d}$.

Proposition 9.1. *Assume (EE) is satisfied, and let* $(f_l)_{l\geq 1}$ *be a sequence of functions in* $C^{r'}$. *Then, for every* $\varepsilon > 0, \varepsilon' > 0$, *there are, for* ν*-a.e* $y \in Y$, $L_y \geq 1$ *such that*

$$\left|\sum_{j=1}^{n} f_l \circ R^j(y) - n\int_Y f_l d\nu\right| \leq ||f_l||_{C^{r'}} n^{\delta+\varepsilon} \quad \text{for } \nu\text{-a.e } y \in Y, \forall n \geq l^{\varepsilon'}, l \geq L_y,$$

Proof. The proof of[19] [8, Theorem 3.1] shows that there is a constant $K > 0$ such that

$$\left|\left|\sup_{n\geq N} \frac{1}{n}\left|S_n(f) - n\int_Y f\, \mathrm{d}\nu\right|\right|\right|_{L^2} \leq K||f||_{C^{r'}} N^{-(1-\delta)} \quad N \geq 1, f \in C^{r'}, \tag{28}$$

where $S_n(f) = \sum_{j=1}^n f_l \circ R^j$.

Now let $k, \delta' > 0$ to be chosen later be such that $2(1-\delta)k > 2\delta' + 1$. Using the Chebyshev inequality, from (28) it follows that

$$\begin{aligned} &\nu\left(\sup_{n\geq N^k} \frac{1}{n}\left|S_n(f) - n\int_Y f\, \mathrm{d}\nu\right| > \frac{1}{2}||f||_{C^{r'}} N^{-\delta'}\right) \\ &\leq 4K^2 N^{2\delta' - 2(1-\delta)k} \quad N \geq 1, f \in C^{r'}. \end{aligned} \tag{29}$$

For $l \geq 1, N = N_l = \lceil l^{\frac{2}{2(1-\delta)k - 2\delta' - 1}} \rceil$ denote by $B_{l,N}$ the set

$$B_{l,N} = \left\{y \in Y \mid \sup_{n\geq N^k} \frac{1}{n}\left|S_n(f_l)(y) - n\int_Y f_l\, \mathrm{d}\nu\right| > \frac{1}{2}||f_l||_{C^{r'}} N^{-\delta'}\right\}, \tag{30}$$

then it holds that $\nu(B_{l,N}) \leq N^{2\delta' - 2(1-\delta)k}$ and

$$\begin{aligned} \sum_{l\geq 1, N\geq N_l} \nu(B_{l,N}) &\leq 4K^2 \frac{1}{2\delta' - 2(1-\delta)k + 1} \sum_{l\geq 1} N_l^{2\delta' - 2(1-\delta)k + 1} \\ &\leq \frac{4K^2}{2\delta' - 2(1-\delta)k + 1} \sum_{l\geq 1} l^{-2} < \infty. \end{aligned}$$

Hence, by Borel-Cantelli, for ν-a.e $y \in Y$ there are only finitely many pairs (l, N) with $N < N_l$ such that $y \in B_{l,N}$, therefore, for such y, there is an $L_y \geq 1$ such that $y \notin B_{l,N}$ whenever $l \geq L_y, N \geq N_l$. For such y, l and $n \geq N_l^k$, say $N^k \leq n \leq (N+1)^k$ and $N \geq N_l$, it holds that

$$\begin{aligned} \frac{1}{n}\left|S_n(f_l)(y) - n\int_Y f\, \mathrm{d}\nu\right| &\geq \sup_{n\geq N^k} \frac{1}{n}\left|S_n(f)(y) - n\int_Y f\, \mathrm{d}\nu\right| \\ &\leq \frac{1}{2}||f_l||_{C^{r'}} N^{-\delta'} \leq ||f_l||_{C^{r'}} n^{-\frac{\delta'}{k}}. \end{aligned}$$

Choosing k so big that $\frac{1}{k} < \frac{\varepsilon\varepsilon'}{2}$, the claim follows by setting $\delta' = k(1-\delta-\varepsilon)$ (then $2(1-\delta)k - 2\delta' - 1 = 2k\varepsilon - 1 > \frac{2}{\varepsilon'}$). ☐

---

[19][8, Theorem 3.1] shows that (28) holds under the stronger assumption that $P^n \xrightarrow{L^2} Id$ polynomially fast on L^∞, where P denotes the Perron-Frobenius transfer operator. Note however that Lemma 3.3 of that paper shows that, under this assumption, (EE) holds; the rest of the proof only uses (EE), and not the stronger assumption on P.

9.2. **Uniform estimates.** Let $(B_l)_{l\geq 1}$ be a sequence of rare events[20] in Y, such that $(B_l)_{l\geq 1}$ is regularly approximable in $C^{r'}$. The goal of this section will be to show that, after taking a subsequence if necessary[21]

$$\sup_{l\geq 1}\left|\frac{\nu(B_l)}{sN_l}\sum_{j=1}^{\lceil sN_l\rceil}\varphi_{B_l}^{(j)}(y)-1\right|\to 0 \quad \textit{as } s\to\infty,\ \nu\textit{-a.e } y\in Y, \tag{31}$$

for some $N_l>0$.

Lemma 9.2. *Let R satisfy (EE) and let $(B_l)_{l\geq 1}$ be a sequence of rare events in Y, such that $(B_l)_{l\geq 1}$ is regularly approximable in $C^{r'}$. Then, after taking a subsequence if necessary, for $\varepsilon>0$, there is a constant $K>0$ only depending on app$(B_l)_{l\geq 1}$, and, for ν-a.e $y\in Y$, $N_y>0$ only depending on y and app$((B_l)_{l\geq 1})$ such that*

$$\begin{aligned}&\left|\sum_{j=1}^{n}1_{B_l}\circ R^j(y)-n\nu(B_l)\right|\\ &\leq K\nu(B_l)^{\frac{r'(d'-1)}{d'(r'+1)}}n^{1-\frac{1-\delta-\varepsilon}{r'+1}},\quad \forall n\geq N_yN_l,\, l\geq L_y\ \nu\textit{-a.e } y\in Y,\end{aligned} \tag{32}$$

where $N_l=\nu(B_l)^{-\frac{r'+1}{1-\delta-\varepsilon}}$.

Proof. Let $\mathcal{B}=\operatorname{app}((B_l)_{l\geq 1})$, then[22] for $k>\nu(B_l)^{-\frac{1}{d'}}$ there are $\underline{h}_{k,l},\overline{h}_{k,l}\in C^{r'}$ with $\underline{h}_{k,l}\leq 1_{B_l}\leq\overline{h}_{k,l}$ and

$$\nu(1_{B_l}\neq h_{k,l})\leq\nu(B_l)^{\frac{d'-1}{d'}}\frac{1}{k},\quad \textit{while}\quad ||h_{k,l}||_{C^{r'}}\leq\mathcal{B}k^{r'},$$

for $h_{k,l}\in\{\overline{h}_{k,l},\underline{h}_{k,l}\}$. In particular

$$\nu(B_l)-\nu(B_l)^{\frac{d'-1}{d'}}\frac{1}{k}\leq\int_Y\underline{h}_{k,l}\,\mathrm{d}\nu\leq\int_Y\overline{h}_{k,l}\,\mathrm{d}\nu\leq\nu(B_l)+\nu(B_l)^{\frac{d'-1}{d'}}\frac{1}{k}.$$

By Proposition 9.1 there are $I_y\geq 1$ and $N_l\geq 1$ such that, for $h_{k,l}\in\{\overline{h}_{k,l},\underline{h}_{k,l}\}$, we have

$$\begin{aligned}&\left|\sum_{j=1}^{n}h_{k,l}\circ R^j(y)-n\int_Y h_{k,l}\,\mathrm{d}\nu\right|\\ &\leq||h_{k,l}||_{C^{r'}}n^{\delta+\varepsilon}\quad\forall n\geq(l+k)^{\varepsilon'}, l+k\geq I_y,\ \nu\textit{-a.e } y\in Y.\end{aligned}$$

Therefore, for such k,l,n,

$$\begin{aligned}\sum_{j=1}^{n}1_{B_l}\circ R^j(y)&\geq\sum_{j=1}^{n}\underline{h}_{k,l}\circ R^j(y)\geq n\int_Y\underline{h}_{k,l}\,\mathrm{d}\nu-\mathcal{B}k^r n^{\delta+\varepsilon}\\ &\geq n\nu(B_l)-n\nu(B_l)^{\frac{d'-1}{d'}}\frac{1}{k}-\mathcal{B}k^{r'}n^{\delta+\varepsilon}.\end{aligned}$$

[20] Later on in §9.3 we will take finitely many such sequences $(B_l^{(1)})_{l\geq 1},...(B_l^{(\mathcal{K})})_{l\geq 1}$, but the same arguments apply.

[21] From this point on we will often take a subsequence of $(B_l)_{l\geq 1}$ to assume that l is sufficiently small compared to $\nu(B_l)^{-1}$. In an effort to keep notation simple, we will do so without explicitly stating, accepting small imprecisions in exchange for simpler notation.

[22] By Definition 8.1.

Likewise

$$\sum_{j=1}^{n} 1_{B_l} \circ R^j(y) \leq n\nu(B_l) + n\nu(B_l)^{\frac{d'-1}{d'}} \frac{1}{k} + \mathcal{B}k^{r'} n^{\delta+\varepsilon}.$$

Hence,

$$\left| \sum_{j=1}^{n} 1_{B_l} \circ R^j(y) - n\nu(B_l) \right| \leq n\nu(B_l)^{\frac{d'-1}{d'}} \frac{1}{k} + \mathcal{B}k^{r'} n^{\delta+\varepsilon}. \tag{33}$$

Let $k_n = \left\lceil n^{\frac{1-\delta-\varepsilon}{r'+1}} \nu(B_l)^{\frac{d'-1}{d'(r'+1)}} \right\rceil$ and

$$N_y = (100 I_y)^{\frac{r'+1}{1-\delta-\varepsilon}}, \quad \textit{and} \quad N_l = \nu(B_l)^{-\frac{r'+1}{(1-\delta-\varepsilon)d'} - \frac{(d'-1)(r'+1)}{d'(1-\delta-\varepsilon)}} = \nu(B_l)^{\frac{r'+1}{1-\delta-\varepsilon}}.$$

We can apply (33) for $n \geq N_y N_l$, k_n and $l \geq 1$ (since $k_n > I_y$ and[23] $k_n^{\varepsilon'} < k_n^{\frac{r'+1}{1-\delta-\varepsilon}} < n$) to obtain

$$\left| \sum_{j=1}^{n} 1_{B_l} \circ R^j(y) - n\nu(B_l) \right| \leq K\nu(B_l)^{\frac{r'(d'-1)}{d'(r'+1)}} n^{1-\frac{1-\delta-\varepsilon}{r'+1}},$$

where the constant $K > 0$ only depends on $\mathcal{B}$. □

Proposition 9.3. *Suppose R satisfies (EE) and let $(B_l)_{l\geq 1}$ be a sequence of rare events in Y, such that $(B_l)_{l\geq 1}$ is regularly approximable in $C^{r'}$. Then for every $\varepsilon > 0$ there is a constant $K > 0$ and, for ν-a.e $y \in Y$, there are $S_y > 0$ such that, for $s > S_y$, we have*

$$\left| \frac{\nu(B_l)}{sM_l} \sum_{j=0}^{\lceil sM_l \rceil - 1} \varphi_{B_l} \circ R_{B_l}^j(y) - 1 \right| \leq K s^{-\frac{1-\delta-\varepsilon}{r'+1}} \tag{34}$$

where

$$M_l = \nu(B_l)^{1-\frac{d'+r'}{d'(1-\delta-\varepsilon)}}.$$

Proof. Let $y \in Y$ be as in the conclusion of Lemma 9.2. By (32), for $l \geq 1$ and $n \geq N_y N_l$ we have

$$\begin{aligned} &n\nu(B_l) - K\nu(B_l)^{\frac{r'(d'-1)}{d'(r'+1)}} n^{1-\frac{1-\delta-\varepsilon}{r'+1}} \\ &\leq \sum_{j=1}^{n} 1_{B_l} \circ R_{B_l}^j(y) \leq n\nu(B_l) + K\nu(B_l)^{\frac{r'(d'-1)}{d'(r'+1)}} n^{1-\frac{1-\delta-\varepsilon}{r'+1}}. \end{aligned}$$

and thus

$$\sum_{j=0}^{\lfloor n\nu(B_l) - K\nu(B_l)^{\frac{r'(d'-1)}{d'(r'+1)}} n^{1-\frac{1-\delta-\varepsilon}{r'+1}} \rfloor - 1} \varphi_{B_l} \circ R_{B_l}^j(y) \leq n, \tag{35}$$

as well as

$$n \leq \sum_{j=0}^{\lceil n\nu(B_l) + K\nu(B_l)^{\frac{r'(d'-1)}{d'(r'+1)}} n^{1-\frac{1-\delta-\varepsilon}{r'+1}} \rceil - 1} \varphi_{B_l} \circ R_{B_l}^j(y).$$

[23] To be completely correct, one would have to consider $k_n + l$ instead of l, but by choosing a subsequence and renumbering we can assume that l is very small compared to $\nu(B_l)^{-1}$.

For $s > S_y$ with $S_y = (100K)^{\frac{r'+1}{1-\delta-\varepsilon}} N_y$, let

$$n = \xi \nu(B_l)^{-\frac{d'+r'}{d'(1-\delta-\varepsilon)}},$$

where $\xi > 0$ is such that $\xi - K\xi^{1-\frac{1-\delta-\varepsilon}{r'+1}} = s$. Clearly

$$\xi \in \left(s, s + 2Ks^{1-\frac{1-\delta-\varepsilon}{r'+1}}\right)$$

in particular[24] $n > N_y N_l$. Hence, using this n in the lower bound of (35) yields

$$\sum_{j=0}^{\lfloor s\nu(B_l)^{1-\frac{d'+r'}{d'(1-\delta-\varepsilon)}} \rfloor -1} \varphi_{B_l} \circ R^j_{B_l}(y) \le n \le (s + 2Ks^{1-\frac{1-\delta-\varepsilon}{r'+1}})\nu(B_l)^{-\frac{d'+r'}{d'(1-\delta-\varepsilon)}},$$

setting $M_l = \nu(B_l)^{1-\frac{d'+r'}{d'(1-\delta-\varepsilon)}}$ we obtain

$$\frac{\nu(B_l)}{sM_l} \sum_{j=0}^{\lfloor sM_l \rfloor -1} \varphi_{B_l} \circ R^j_{B_l}(y) \le 1 + 2Ks^{-\frac{1-\delta-\varepsilon}{r'+1}}.$$

The upper bound follows analogously by setting $\xi + K\xi^{1-\frac{1-\delta-\varepsilon}{r'+1}} = s$, and the claim (making K a bit bigger) is proven. □

9.3. **PLT scaled by returns to $\{B_l\}$.** For the rest of this exposition let $\mathcal{K} \ge 1$, and $A_l^{(1)}, ..., A_l^{(\mathcal{K})}$ resp. $B_l^{(1)}, ..., B_l^{(\mathcal{K})}$ be subsets of X resp. Y regularly approximable in C^r resp. $C^{r'}$. Suppose that there are $r_l \searrow 0$, and positive constants $\omega^{(k)}, \theta^{(k)} > 0$ such that

$$\mu(A_l^{(k)}) = \omega^{(k)} r_l^d + o(r_l^d), \quad \textit{and} \quad \nu(B_l^{(k)}) = \theta^{(k)} r_l^{d'} + o(r_l^{d'}) \quad \forall l \ge 1, k = 1, ..., \mathcal{K},$$

and, for each l, $B_l^{(1)}, ..., B_l^{(\mathcal{K})}$ are disjoint. Denote $B_l = \bigcup_{k=1}^{\mathcal{K}} B_l^{(k)}$, which, by disjointness of $B_l^{(1)}, ..., B_l^{(\mathcal{K})}$, is also regularly approximable in $C^{r'}$, and $\alpha_l = \Phi_{B_l}$. Consider

$$\kappa_l^{(n)}(y) = k \quad \textit{if} \quad R^n_{B_l}(y) \in B_l^{(k)},$$

which is well-defined as the sets $B_l^{(k)}$ are disjoint, and

$$\sigma^{(1)}_{\kappa_l,\alpha_l,y}(x) = \min\left(n \ge 1 \mid T_y^{\tilde{\alpha}_l^{(n)}}(x) \in A_l^{(\kappa_l^{(n)})}\right),$$

$$\sigma^{(n+1)}_{\kappa_l,\alpha_l,y}(x) = \min\left(k \ge \sigma^{(n)}_{\kappa_l,\alpha_l,y}(x) + 1 \mid T_y^{\tilde{\alpha}_l^{(k)}}(x) \in A_l^{(\kappa_l^{(k)})}\right),$$

$$\Sigma_{\kappa_l,\alpha_l,y} = (\sigma^{(1)}_{\kappa_l,\alpha_l,y}, \sigma^{(2)}_{\kappa_l,\alpha_l,y}, ...).$$

Following the steps outlined in §7 we will first show

$$\Omega r_l^{d'} \Sigma_{\kappa_l,\alpha_l,y} \overset{\mu}{\Rightarrow} \Sigma_{Exp} \quad \textit{as } l \to \infty, \ \nu - a.a\ y \in Y \tag{36}$$

for $\Omega = \frac{\sum_{k=1}^{\mathcal{K}} \omega^{(k)}\theta^{(k)}}{\sum_{k=1}^{\mathcal{K}} \theta^{(k)}}$. Then use the relation

$$\sigma_{\bigcup_{k=1}^{\mathcal{K}} A_l^{(k)} \times B_l^{(k)}}(x, y) = \sum_{j=0}^{\sigma_{\kappa_l,\Phi_{B_l}(y)}(x)-1} \varphi_{B_l} \circ R^j_{B_l}(y) \tag{37}$$

[24] Recall that $N_l = \nu(B_l)^{-\frac{1+r'}{d'(1-\delta-\varepsilon)}}$.

to obtain

$$(\mu\times\nu)\left(\bigcup_{k=1}^{\mathcal{K}} A_l^{(k)}\times B_l^{(k)}\right)\Sigma_{\bigcup_{k=1}^{\mathcal{K}} A_l^{(k)}\times B_l^{(k)}} \overset{\mu\times\nu}{\Rightarrow} \Sigma_{Exp} \quad as\ l\to\infty. \tag{38}$$

Denote

$$p_{l,t}^{(k)}(y) := \frac{1}{t}\#\{j=1,...,t \mid \kappa_l^{(j)}(y)=k\} = \frac{1}{t}\sum_{j=1}^{t} 1_{B_l^{(k)}}(R_{B_l}^j(y)).$$

We first show that, for each k, ν-a.e $y\in Y$, and $t_l = O(r_l^{-d})$, we have

$$p_{l,t_l}^{(k)}(y) \to \frac{\theta^{(k)}}{\sum_{j=1}^{\mathcal{K}}\theta^{(j)}} =: p^{(k)} \quad as\ l\to\infty. \tag{39}$$

Lemma 9.4. *Suppose R satisfies (EE) and*

$$d > \frac{r'+d'\delta}{1-\delta}$$

Then κ_l satisfies (39).

Proof. Let $\varepsilon>0$ be small enough that

$$d > \frac{r'+d'(\delta+\varepsilon)}{1-\delta-\varepsilon}.$$

(i) Using Lemma 9.2, and disjointness we obtain a constant $K>0$ and $N_y\ge 1$ such that, for ν-a.e $y\in Y$, we have

$$\left|\sum_{j=1}^{n} 1_{B_l^{(k)}}\circ R^j(y) - n\nu(B_l^{(k)})\right| \le K r_l^{\frac{r'(d'-1)}{r'+1}} n^{1-\frac{1-\delta-\varepsilon}{r'+1}}, \quad and$$

$$\left|\sum_{j=1}^{n} 1_{B_l}\circ R^j(y) - n\nu(B_l)\right| \le K r_l^{\frac{r'(d'-1)}{r'+1}} n^{1-\frac{1-\delta-\varepsilon}{r'+1}}$$

for all $n\ge N_y, l\ge 1$ and $k=1,...,\mathcal{K}$. On the other hand[25], for $s>N_y$, Proposition 9.3 yields

$$\left|\frac{\nu(B_l)}{sM_l}\sum_{j=0}^{\lceil sM_l\rceil-1}\varphi_{B_l}\circ R_{B_l}^j(y) - 1\right| \le K s^{-\frac{1-\delta-\varepsilon}{r'+1}}.$$

where

$$M_l = \nu(B_l)^{1-\frac{d'+r'}{d'(1-\delta-\varepsilon)}}.$$

(ii) Rewrite

$$p_{l,n}^{(k)} = \frac{1}{n}\sum_{i=1}^{\sum_{j=1}^{n}\varphi_{B_l}\circ R_{B_l}^j(y)} 1_{B_l^{(k)}}\circ R^i.$$

Denote $t_l = s_l M_l$, since

$$M_l = \nu(B_l)^{1-\frac{d'+r'}{d'(1-\delta-\varepsilon)}} = O(r_l^{\frac{d'+r'}{1-\delta-\varepsilon}-d'}) = o(r_l^d),$$

we necessarily have $s_l\to\infty$ as $l\to\infty$. Denote $a=\frac{1-\delta-\varepsilon}{r'+1}$ and let l be big enough so that $s_l>N_y$, then we have

[25]Making K and N_y bigger if needed.

$$\begin{aligned}
t_l p_{l,t_l}^{(k)}(y) &= \sum_{j=1}^{\sum_{j=1}^{t_l} \varphi_{B_l} \circ R_{B_l}^j - 1} 1_{B_l^{(k)}} \circ R^j(y) \\
&\leq \sum_{j=1}^{t_l \nu(B_l)^{-1}(1+K s_l^{-a})-1} 1_{B_l^{(k)}} \circ R^j(y) \\
&\leq t_l \nu(B_l)^{-1} \nu(B_l^{(k)})(1+K s^{-a}) + K r_l^{\frac{r'(d'-1)}{r'+1}} \left(t_l \nu(B_l)^{-1}(1+K s^{-a})\right)^{1-a} \\
&\leq t_l \frac{\nu(B_l^{(k)})}{\nu(B_l)}(1+o(s_l^{-a+a^2})).
\end{aligned}$$

The lower bound is similar. So

$$\left| p_{l,n}^{(k)}(y) - \frac{\nu(B_l^{(k)})}{\nu(B_l)} \right| = o(s_l^{-a+a^2}) = o(1). \tag{40}$$

□

9.4. **Adding in the gaps.** Now all that's left to do is to add back in the gaps. As mentioned in §9.3, having shown (36) (this is the content of Proposition 8.4), we will now explain how to conclude

$$(\mu \times \nu)\left(\bigcup_{k=1}^{\mathcal{K}} A_l^{(k)} \times B_l^{(k)}\right) \Sigma_{\bigcup_{k=1}^{\mathcal{K}} A_l^{(k)} \times B_l^{(k)}} \overset{\mu \times \nu}{\Rightarrow} \Sigma_{Exp}, \quad \textit{as } l \to \infty$$

using the relation

$$\sigma_{\bigcup_{k=1}^{\mathcal{K}} A_l^{(k)} \times B_l^{(k)}} = \sum_{j=0}^{\sigma_{\kappa_l, \alpha_l, y}(x)-1} \varphi_{B_l} \circ R_{B_l}^j(y).$$

This is rather straightforward, given Proposition 9.3, and follows from a more general principle in probability theory. As this principle finds use in various places and has, to the author's knowledge, not been formulated in generality, let us state and prove a more general version than we need here.

Lemma 9.5. *Let $(\Omega, \mathbb{P})$ be a probability space, and $E_l : \Omega \to [0, \infty)$ non-negative real random variables, such that there are positive random variables $\mu_l : \Omega \to (0, 1)$ with*

$$\mu_l E_l \overset{\mathbb{P}}{\Rightarrow} E \quad \textit{as } l \to \infty,$$

for some non-negative random variable E with $\mathbb{P}(E = 0) = 0$. Then for any $M_l : \Omega \to [0, \infty)$ with

$$\mu_l M_l \to 0 \quad \textit{as } l \to \infty \quad \textit{pointwise } \mathbb{P}\textit{-a.e}$$

we have

$$\mathbb{P}(E_l \leq M_l) \to 0 \quad \textit{as } l \to \infty.$$

Proof. Let $\varepsilon > 0$, there is a $\delta > 0$ such that $\mathbb{P}(E \leq \delta) < \varepsilon$, and the distribution function of E is continuous at δ. By Jegorow's Theorem, there is a measurable $K \subset \Omega$ with $\mathbb{P}(K^c) < \varepsilon$ and

$$\mu_l M_l \to 0 \quad \textit{as } l \to \infty \quad \textit{uniformly on } K.$$

Choose $\tilde{L} \geq 1$ so big that $M_l \leq \delta\mu_l^{-1}$ on K for $l \geq \tilde{L}$. Now choose $L \geq \tilde{L}$ so big that

$$\mathbb{P}(\mu_l E_l \leq \delta) \leq 2\varepsilon \quad \forall l \geq L,$$

it follows that

$$\mathbb{P}(E_l \leq M_l) \leq \mathbb{P}(\mu_l E_l \leq \delta) + \mathbb{P}(K^c) \leq 3\varepsilon,$$

for $l \geq L$. □

Proposition 9.6. *Let $(\Omega, \mathbb{P})$ be a probability space, and $E_l : \Omega \to \mathbb{N}$ be positive integer valued observables. Assume there are positive real numbers $q_l \searrow 0$, and a $[0,\infty)$-valued random variable E with $\mathbb{P}(E = 0) = 0$ such that*

$$q_l E_l \overset{\mathbb{P}}{\Rightarrow} E \quad \text{as } l \to \infty,$$

Let $\alpha_j^{(l)} : \Omega \to [0,\infty)$ be non-negative random variables, and assume there are $M_l : \Omega \to (0,\infty)$ with $q_l M_l \to 0$ as $l \to \infty$ $\mathbb{P}$-a.e and positive random variables $b_l : \Omega \to (0,\infty)$ such that

$$\sup_{l\geq 1} \left| \frac{1}{sM_l b_l} \sum_{j=1}^{\lceil sM_l \rceil} \alpha_l^{(j)}(\omega) - 1 \right| \to 0 \quad \text{as } s \to \infty, \quad \mathbb{P}\text{-a.e.} \ . \tag{UC}$$

Then

$$\frac{q_l}{b_l} \sum_{j=1}^{E_l} \alpha_l^{(j)} \overset{\mathbb{P}}{\Rightarrow} E \quad \text{as } l \to \infty.$$

Remark 9.7. In our context we use $(\Omega, \mathbb{P}) = (X \times Y, \mu \times \nu)$, $E_l = \sigma_{A_l, \Phi_{B_l}}$, $\alpha_l^{(j)} = \varphi_{B_l} \circ R_{B_l}^j$, $q_l = \mu(A_l)$ and $b_l = \frac{1}{\nu(B_l)}$. The existence of M_l is the content of Proposition 9.3.

Proof. (i) Let F be the distribution function of E and $C = \{t \mid F \textit{ is continuous at } t\}$ the set of its continuities. Let $t \in C$, and $\varepsilon > 0$ such that $\frac{t}{1+\varepsilon}, \frac{t}{1-\varepsilon} \in C$. We will show

$$\mathbb{P}\left(\frac{q_l}{b_l} \sum_{j=1}^{E_l} \alpha_l^{(j)} \leq t \right) \to F(t) \quad \text{as } l \to \infty.$$

(ii) By Jegorow's Theorem, for $l \geq 1$, there is a measurable set $K \subset \Omega$ with $\mathbb{P}(K^c) < \varepsilon$

$$\sup_{l\geq 1} \sup_{\omega \in K} \left| \frac{1}{sM_l b_l} \sum_{j=1}^{\lceil sM_l \rceil} \alpha_l^{(j)}(\omega) - 1 \right| \to 0 \quad \text{as } s \to \infty.$$

(iii) Choose $S > 0$ so big that

$$(1-\varepsilon) b_l \leq \frac{1}{sM_l} \sum_{j=1}^{\lceil sM_l \rceil} \alpha_l^{(j)} \leq (1+\varepsilon) b_l \quad \text{on } K, \ \forall s \geq S, l \geq 1,$$

and restricting to $\{E_l \geq SM_l\}$, we obtain

$$(1-\varepsilon) q_l E_l \leq \frac{q_l}{b_l} \sum_{j=1}^{E_l} \alpha_l^{(j)}(\omega) \leq (1+\varepsilon) q_l E_l \quad \text{on } K \cap \{E_l \geq SM_l\}, \ \forall l \geq 1.$$

We get

$$\mathbb{P}\left(\frac{q_l}{b_l}\sum_{j=1}^{E_l}\alpha_l^{(j)} \le t\right) - \mathbb{P}(E_l \le SM_l) - \mathbb{P}(K^c)$$
$$\le \mathbb{P}\left(K \cap \{E_l \ge SM_l\} \cap \left\{\frac{q_l}{b_l}\sum_{j=1}^{E_l}\alpha_l^{(j)} \le t\right\}\right) \le \mathbb{P}\left(q_l E_l \le \frac{t}{1-\varepsilon}\right),$$

and likewise

$$\mathbb{P}\left(\frac{q_l}{b_l}\sum_{j=1}^{E_l}\alpha_l^{(j)} \le t\right) \ge \mathbb{P}\left(K \cap \{E_l \ge SM_l\} \cap \left\{q_l E_l \le \frac{t}{1+\varepsilon}\right\}\right)$$
$$\ge \mathbb{P}\left(q_l E_l \le \frac{t}{1+\varepsilon}\right) - \mathbb{P}(E_l \le SM_l) - \mathbb{P}(K^c).$$

Taking $\limsup_{l\to\infty}$ resp. $\liminf_{l\to\infty}$ the above two equations, and using Lemma 9.5, we obtain

$$F\left(\frac{t}{1+\varepsilon}\right) - \varepsilon \le \liminf_{l\to\infty} \mathbb{P}\left(\frac{q_l}{b_l}\sum_{j=1}^{E_l}\alpha_l^{(j)} \le t\right)$$
$$\le \limsup_{l\to\infty} \mathbb{P}\left(\frac{q_l}{b_l}\sum_{j=1}^{E_l}\alpha_l^{(j)} \le t\right) \le F\left(\frac{t}{1-\varepsilon}\right) + \varepsilon.$$

Since $C \subset (0,\infty)$ is dense, we can let $\varepsilon \searrow 0$ while $\frac{t}{1+\varepsilon}, \frac{t}{1-\varepsilon} \in C$, this yields

$$\mathbb{P}\left(\frac{q_l}{b_l}\sum_{j=1}^{E_l}\alpha_l^{(j)} \le t\right) \to F(t) \quad \textit{as } l \to \infty.$$

□

Remark 9.8.

(i) We can extend this statement to sequences in the following manner: under the assumptions of the proposition, let $E_l^{(n)} : \Omega \to \mathbb{N}$ be such that

$$q_l(E_l^{(1)}, E_l^{(2)}, \ldots) \overset{\mu}{\Rightarrow} (E^{(1)}, E^{(2)}, \ldots) \quad \textit{as } l \to \infty,$$

for some $E^{(n)} : \Omega \to [0,\infty)$ with $\mathbb{P}(E^{(n)} = 0) = 0$. Then

$$\frac{q_l}{b_l}\left(\sum_{j=1}^{E_l^{(n)}}\alpha_l^{(j)}, \sum_{j=1}^{E_l^{(n)}}\alpha_l^{(j)}, \ldots\right) \overset{\mathbb{P}}{\Rightarrow} (E^{(1)}, E^{(2)}, \ldots) \quad \textit{as } l \to \infty.$$

The proof of this statement is almost the same as for the proposition, therefore we won't repeat it.

(ii) The probability measure $\mathbb{P}$ can be replaced by a sequence $(\mathbb{P}_l)_{l\ge1}$ by also replacing (UC) with the following condition

$\forall\varepsilon > 0$ *there is are measurable sets* $K_l \subset \Omega$ *with* $\limsup_{l\to\infty} \mathbb{P}_l(K_l^c) < \varepsilon$ *such that*

$$\sup_{l\ge1}\sup_{\omega\in K_l}\left|\frac{1}{sM_l b_l}\sum_{j=1}^{\lceil sM_l\rceil}\alpha_l^{(j)}(\omega) - 1\right| \to 0 \quad \textit{as } s \to \infty.$$

Proof of Theorem 3.2. (i) Let $r_l \searrow 0$, and denote by $Q_l = B_{r_l}(x^*, y^*)$ the geodesic ball of radius r_l centred at (x^*, y^*), and let $\varepsilon > 0$. W.l.o.g. r_1 is small enough that the exponential map at (x^*, y^*) is a diffeomorphism from the ball of radius $2r_1$ in $\mathbb{R}^{d+d'}$ onto $B_{2r_1}(x^*, y^*)$. Let $\varepsilon > 0$, it is easy to construct, for some $\mathcal{K} \geq 1$, sets $A_l^{(1)}, ..., A_l^{(\mathcal{K})}$ and $B_l^{(1)}, ..., B_l^{(\mathcal{K})}$ as in §8. Let $\omega^{(k)}, \theta^{(k)} > 0$ be such that[26]

$$\mu(A_l^{(k)}) = \omega^{(k)} r_l^d + o(r_l^d), \quad \textit{and} \quad \nu(B_l^{(k)}) = \theta^{(k)} r_l^{d'} + o(r_l^{d'}) \quad \forall l \geq 1, k = 1, ..., \mathcal{K},$$

and set $\Omega = \frac{\sum_{k=1}^{\mathcal{K}} \omega^{(k)} \theta^{(k)}}{\sum_{k=1}^{\mathcal{K}} \theta^{(k)}}$.

Due to Lemma 8.2 all those sets can be chosen to be regularly approximable, such that[27] $(\mu \times \nu)_{Q_l} \left(Q_l \setminus \bigcup_{k=1}^{\mathcal{K}} A_l^{(k)} \times B_l^{(k)} \right) < \varepsilon$ for all $l \geq 1$. Let $\Lambda > 0$ be such that $(\mu \times \nu)(Q_l) = \Lambda r_l^{d+d'} + o(r_l^{d+d'})$, then

$$\left| \Lambda - \Omega \sum_{k=1}^{\mathcal{K}} \theta^{(k)} \right| < \varepsilon$$

(ii) Denote $B_l = \bigcup_{k=1}^{\mathcal{K}} B_l^{(k)}$, and, for y as in assumption (MEM), consider $\alpha_l(y) = \Phi_{B_l}(y)$ and

$$\kappa_l^{(n)}(y) = k \quad \textit{if} \quad R_{B_l}^n(y) \in B_l^{(k)},$$

by disjointness $\kappa_l(y)$ is well-defined. By Lemma 9.4, $\kappa_l(y)$ satisfies (39), and $p^{(k)} = \frac{\theta^{(k)}}{\sum_{s=1}^{\mathcal{K}} \theta^{(s)}}$. We can use Proposition 8.4 and Remark 8.5(i) to obtain

$$\Omega r_l^d \Sigma_{\kappa_l(y), \alpha_l(y), y} \overset{\mu}{\Rightarrow} \Sigma_{Exp} \quad \textit{as } l \to \infty, \ \nu\textit{-a.e } y.$$

Since the convergence holds for ν-a.e $y \in Y$, it follows that

$$\Omega r_l^d \Sigma_{\kappa_l, \alpha_l} \overset{\mu \times \nu}{\Rightarrow} \Sigma_{Exp} \quad \textit{as } l \to \infty,$$

where $\Sigma_{\kappa_l, \alpha_l}(x, y) = \Sigma_{\kappa_l(y), \alpha_l(y), y}(x)$.

(iii) By Proposition 9.3, α_l satisfies (UC) with

$$b_l = \frac{1}{\nu(B_l)} = \frac{1}{\sum_{k=1}^{\mathcal{K}} \theta^{(k)}} r_l^{-d'} + o(r_l^{-d'})$$

and

$$M_l = \nu(B_l)^{1 - \frac{d' + r'}{d'(1 - \delta - \varepsilon)}}.$$

Note that, for $Q_l' = \bigcup_{k=1}^{\mathcal{K}} A_l^{(k)} \times B_l^{(k)}$, we have

$$\Sigma_{Q_l'} = \left(\sum_{j=0}^{\sigma_{\kappa_l, \alpha_l}^{(1)} - 1} \varphi_{B_l} \circ R_{B_l}^j, \sum_{j=\sigma_{\kappa_l, \alpha_l}^{(1)}}^{\sigma_{\kappa_l, \alpha_l}^{(2)} - 1} \varphi_{B_l} \circ R_{B_l}^j, ... \right).$$

We can apply Proposition 9.6 resp. Remark 9.8(i)

$$\Omega \left(\sum_{k=1}^{\mathcal{K}} \theta^{(k)} \right) r_l^{d+d'} \Sigma_{Q_l'} \overset{\mu \times \nu}{\Rightarrow} \Sigma_{Exp} \quad \textit{as } l \to \infty.$$

[26] Choose x^*, y^* such that the densities of μ, ν are positive at the respective points.
[27] Here we again use the continuity of the density.

By disjointness

$$(\mu \times \nu)(Q_l') = \Omega\left(\sum_{k=1}^{\mathcal{K}} \theta^{(k)}\right) r_l^{d+d'} + o(r_l^{d+d'})$$

hence

$$(\mu \times \nu)(Q_l')\Sigma_{Q_l'} \overset{\mu\times\nu}{\Rightarrow} \Sigma_{Exp} \quad \textit{as } l \to \infty.$$

By the equivalence (12), we have

$$(\mu \times \nu)(Q_l')\Phi_{Q_l'} \overset{\mu\times\nu}{\Rightarrow} \Phi_{Exp} \quad \textit{as } l \to \infty.$$

(iv) By Theorem 2.2 also

$$\mu \times \nu)(Q_l')\Phi_{Q_l'} \overset{\mu\times\nu_{Q_l'}}{\Rightarrow} \Phi_{Exp} \quad \textit{as } l \to \infty.$$

At the same time, taking a subsequence if necessary, there are $[0,\infty]$-valued processes Φ and $\tilde{\Phi}$ such that

$$(\mu \times \nu)(Q_l)\Phi_{Q_l} \overset{\mu\times\nu}{\Rightarrow} \Phi \quad \textit{as } l \to \infty, \quad \textit{and}$$

$$(\mu \times \nu)(Q_l)\Phi_{Q_l} \overset{(\mu\times\nu)_{Q_l}}{\Rightarrow} \tilde{\Phi} \quad \textit{as } l \to \infty.$$

Hence

$$D\left(\Phi_{Exp}, \tilde{\Phi}\right) \le 7\varepsilon,$$

where D is given by Lemma 7.1. Since this is true for every $\varepsilon > 0$ we have $\tilde{\Phi} \overset{d}{=} \Phi_{Exp}$, and by Theorem 2.2 also $\Phi \overset{d}{=} \Phi_{Exp}$. □

10. **The skewing time.** Here we will prove Theorem 3.4, to do this we will verify that the map $T(x,y) = G_{\tau(y)}(x)$ satisfies superpolynomial mixing of all orders, as in Remark 8.5(ii).

Lemma 10.1. *Under the assumptions of Theorem 3.4, suppose that* $\sum_{l\ge1} r_l^{\frac{1}{2}(d'-\delta_2\kappa)} < \infty$. *Then, for each* $t > 0$*, there is a set* $\mathfrak{G}_t$ *with* $\nu(\mathfrak{G}_t) = 1$ *such that, for* $y \in \mathfrak{G}_t$*, there are* $L_{y,t} > 0$ *and sets* $\mathcal{B}_{l,y,t} \subset \{1, ..., \left\lceil \frac{t}{\mu(A_l)} \right\rceil\}$ *with* $\#\mathcal{B}_{l,y,t} = o(\mu(A_l)^{-1})$ *such that*

$$\begin{aligned}&|\tau_{\tilde{\alpha}_l^{(n)}}(y) - \tau_{\tilde{\alpha}_l^{(m)}}(y)|\\ &\ge \zeta\left(\tilde{\alpha}_l^{(n)}(y) - \tilde{\alpha}_l^{(m)}(y)\right) \quad \forall l \ge L_{y,t}, 1 \le n < m \le \left\lceil \frac{t}{\mu(A_l)} \right\rceil, n \notin \mathcal{B}_{l,y,t}\end{aligned}$$

where ζ *is the function from condition (BA) and*[28] $\alpha_l = \Phi_{B_l}$.

Proof. Fix $t > 0$, to keep notation simple we assume $\mu(A_l) = r_l^d + o(r_l^d)$ and $\nu(B_l) = r_l^{d'} + o(r_l^{d'})$, otherwise there is an extra constant in the estimates below.

(i) We call $n \in \left\{1, ..., \left\lceil \frac{t}{\mu(A_l)} \right\rceil\right\}$ a (l,y)-bad return (or simply (l,y)-bad) if there is a $m > n$ such that

$$|\tau_{\tilde{\alpha}_l^{(n)}}(y) - \tau_{\tilde{\alpha}_l^{(m)}}(y)| < \zeta(\tilde{\alpha}_l^{(n)}(y) - \tilde{\alpha}_l^{(m)}(y)),$$

denote $\mathcal{B}_{l,y} = \{n \ge 1 \mid n \textit{ is a } (l,y) - \textit{bad return}\}$. Let $\varepsilon_1 > 0$, we call $y \in Y$ an l-bad point if $\#\mathcal{B}_{l,y} > r_l^{-d+\varepsilon_1}$.

[28] It can be recalled from Definition 5.1 that $\tilde{\alpha}_l^{(n)} = \sum_{j=1}^{n} \alpha_l^{(j)} = \sum_{j=0}^{n-1} \varphi_{B_l} \circ R^j$.

(ii) Using Proposition 9.3 and Jegorov's Theorem, for $\varepsilon_2 > 0$, we can find a measurable $G = G_{\varepsilon_2} \subset Y$ with $\nu(G^c) < \varepsilon_2$ and an $\tilde{L} \geq 1$ depending on G such that

$$\tilde{\alpha}_l^{\left(\left\lfloor \frac{t}{\mu(A_l} \right\rfloor\right)} \leq 2tr_l^{-(d+d')} \quad \forall y \in G, l \geq \tilde{L}. \tag{41}$$

(iii) For $l \geq 1$ denote $G_l = \{\tilde{\alpha}_l^{\left(\left\lfloor \frac{t}{\mu(A_l)} \right\rfloor\right)} \leq 2tr_l^{-(d+d')}\}$, we have

$$\begin{aligned}
\nu(y \in G_l \mid y \textit{ is } l\textit{-bad}) &\leq \frac{\int_{G_l} \#\mathcal{B}_{l,y}\, \mathrm{d}\nu}{r_l^{-d+\varepsilon_1}} \leq r_l^{d-\varepsilon_1} \sum_{n=1}^{\left\lceil \frac{t}{\mu(A_l)} \right\rceil} \nu(y \in G_l \mid n \textit{ is } (l,y)\textit{-bad}) \\
&\leq r_l^{d-\varepsilon_1} \sum_{j=1}^{\lceil 2tr_l^{-(d+d')} \rceil} \nu\left(\exists i \geq 1 \mid |\tau_{j+\tilde{\alpha}_l^{(i)}} - \tau_j| < \zeta\left(\tilde{\alpha}_l^{(i)}\right)\right) \\
&\leq r_l^{d-\varepsilon_1} \sum_{j=1}^{\lceil 2tr_l^{-(d+d')} \rceil} \nu\left(R^{-j}\left(\exists i \geq 1 \mid |\tau_{\tilde{\alpha}_l^{(i)}}| < \zeta\left(\tilde{\alpha}_l^{(i)}\right)\right)\right) \\
&\leq r_l^{d-\varepsilon_1} \sum_{j=1}^{\lceil 2tr_l^{-(d+d')} \rceil} \nu(\exists i \geq r_l^{-\delta_2} \mid |\tau_i| < \zeta(i)) \\
&\leq 2Ktr_l^{d-\varepsilon_1-d-d'+\delta_2\kappa} \leq 2Ktr_l^{\delta_2\kappa-d'-\varepsilon_1},
\end{aligned}$$

for some constant $K > 0$, for small enough ε_1 this is summable. An application of the Borel-Cantelli Lemma yields that for almost every $y \in Y$; for big enough l, either $y \notin G_l$ or

$$|\tau_{\tilde{\alpha}_l^{(n)}}(y) - \tau_{\tilde{\alpha}_l^{(m)}}(y)| \geq \zeta(\tilde{\alpha}_l^{(n)}(y) - \tilde{\alpha}_l^{(m)}(y)) \quad \forall 1 \leq n < m \leq \left\lceil \frac{t}{\mu(A_l)} \right\rceil, n \notin \mathcal{B}_{l,y},.$$

At the same time, by (41), we have $G_l \nearrow Y$. Thus ν-a.e $y \in Y$ is in G_l for big enough l, and the conclusion follows. □

Proof of Theorem 3.4. In order to keep notation simple we will only show the PLT for regularly approximable rectangles, this can be easily extended to geodesic balls, by following the same arguments as in the proof of Theorem 3.2.

(i) For ν-a.e $y \in Y$ and $0 = t_0 < t_1 < ... < t_J$ choose $L_y = L_{y,t_1+...+t_J}$ and $\mathcal{B}_{l,y} = \mathcal{B}_{l,y,t_1+...+t_J}$ as in Lemma 10.1. For such a y and $l \geq L_y$ (in the following we suppress y from the notation) consider

$$S_{t_j,l} = \sum_{i=1}^{\left\lceil \frac{t_j}{\mu(A_l)} \right\rceil} 1_{A_l} \circ T^{\tilde{\alpha}_l^{(i)}} = S'_{t_j,l} + S''_{t_j,l},$$

where

$$S'_{t_j,l} = \sum_{\substack{i=1,\dots,\left\lceil \frac{t_j}{\mu(A_l)} \right\rceil \\ i \notin \mathcal{B}_l}} 1_{A_l} \circ T^{\tilde{\alpha}_l^{(j)}}.$$

As in Proposition 8.4, the first goal is to show

$$\left(S_{t_1,l} - S_{t_0,l}, ..., S_{t_J,l} - S_{t_{J-1},l}\right) \overset{\mu}{\Rightarrow} \left(P_{t_1-t_0}, ..., P_{t_J-t_{J-1}}\right) \quad \textit{as } l \to \infty,$$

where (P_t) is a standard Poisson process. Since $||S''_{t_j,l}||_{L^1} \to 0$ for all $j = 1, ..., J$, it is equivalent to show

$$\left(S'_{t_1,l} - S'_{t_0,l}, ..., S'_{t_J,l} - S'_{t_{J-1},l}\right) \overset{\mu}{\Rightarrow} \left(P_{t_1-t_0}, ..., P_{t_J-t_{J-1}}\right) \quad as\ l \to \infty.$$

For $m_1, ..., m_J \geq 1$ it will be enough to show

$$\int_X \prod_{j=1}^{J} \binom{S'_{t_j,l} - S'_{t_{j-1},l}}{m_j} \mathrm{d}\mu = \prod_{j=1}^{J} \frac{(t_j - t_{j-1})^{m_j}}{m_j!}. \tag{42}$$

We have

$$\prod_{j=1}^{J} \binom{S'_{t_j,l} - S'_{t_{j-1},l}}{m_j} = \sum_{\substack{\left\lceil \frac{t_{j-1}}{\mu(A_l)} \right\rceil + 1 \leq k_{j,1} < ... < k_{j,m_j} \leq \left\lceil \frac{t_j}{\mu(A_l)} \right\rceil \\ k_{j,i} \notin \mathcal{B}_l \ for\ j=1,...,J,\ i=1,...,m_j}} \prod_{i,j} 1_{A_l} \circ T^{\tilde{\alpha}_l^{(k_{j,i})}}. \tag{43}$$

(ii) Due to assumption (MEM) for G, and Lemma 10.1, we have

$$\left| \int_X \prod_{j=1}^{m} f_j \circ T^{\tilde{\alpha}_l^{(n_j)}} \mathrm{d}\mu - \prod_{j=1}^{m} \int_X f_j \, \mathrm{d}\mu \right| \leq C_y \psi(\min_{j \neq j'} |\tilde{\alpha}_l^{(n_j)} - \tilde{\alpha}_l^{(n_{j'})}|) \prod_{j=1}^{m} ||f_j||_{C^r},$$

where $\psi(x) = e^{-\gamma\zeta(x)}$ and ζ is as in assumption (BA), for $f_1, ..., f_m \in C^r$ and $1 \leq n_1 \leq ... \leq n_m \leq \left\lceil \frac{t1+...+t_J}{r_l^d} \right\rceil$ with $n_i \notin \mathcal{B}_l$. Due to (9) and assumption (BA) we have

$$\psi(\min_{j \neq j'} |\tilde{\alpha}_l^{(n_j)} - \tilde{\alpha}_l^{(n_{j'})}|) = O(r_l^{w_l}),$$

for some $w_l > 0$ with $w_l \to \infty$ as $l \to \infty$. Approximating 1_{A_l} by functions in C^r it is straightforward[29] to show that

$$\left| \int_X \prod_{i,j} 1_{A_l} \circ T^{\tilde{\alpha}_l^{(k_{j,i})}} \mathrm{d}\mu - \mu(A_l)^{m_1+...+m_J} \right| = o(\mu(A_l)^{m_1+...+m_J}),$$

for $k_{j,i}$ as in (43). The sum in (43) has

$$\mu(A_l)^{-(m_1+...+m_J)} \prod_{j=1}^{J} \frac{(t_j - t_{j-1})_j^m}{m_j!} + o(\mu(A_l)^{-(m_1+...+m_J)})$$

many terms, so (42) follows. □

11. **Examples.**

Here, we verify conditions (EE) and (BR) for the examples listed in §4.

11.1. **Diophantine rotations.**

Let $\alpha \in ((0,1) \setminus \mathbb{Q})^{d'}$ satisfy a Diophantine condition, i.e. there are $C > 0$ and $n \geq 1$ such that

$$|\langle k, \alpha \rangle - l| > C|k|^{-n} \quad \forall k \in \mathbb{Z}^{d'}, k \neq 0, l \in \mathbb{Z}, \tag{D}$$

and $R = R_\alpha : x \mapsto x + \alpha \ (mod\ 1)$, for $x \in \mathbb{T}^{d'}$, the rotation by α. Almost all α satisfy (D) for some $n > d'$ (this is a consequence of a higher dimensional version of Khinchin's Theorem, see e.g. [4]). If $d' = n = 1$, then we say α is of *bounded type*.

Note that (D) implies that there is a constant $C' > 0$ such that

$$\left|1 - e^{2\pi i \langle k, \alpha \rangle}\right| \geq C'|k|^{-n} \quad \forall k \in \mathbb{Z}^{d'} \setminus \{0\}.$$

[29] The calculation is analogous to Lemma 8.3.

Property $(SLR(y^*))$ follows directly from (D). Due to the self-symmetry of R_α it is enough to consider returns of $x = 0$ to a rectangle $(-r, r)^{d'}$, but

$$R_\alpha^m(0) \in (-r, r)^{d'} \iff |m\alpha - k| < r \quad \textit{for some } k \in \mathbb{Z}^{d'}.$$

Then, by (D), $Cm^{-n} < r$, equivalently $m > (C^{-1}r)^{-\frac{1}{n}}$. Hence, (4) is satisfied with $\psi(r) = (C^{-1}r)^{-\frac{1}{n}}$.

To show effective equidistribution (EE), we solve the homological equation. Let $f \in H^n(\mathbb{T}^{d'})$ with $\int f = 0$. Then

$$f(x) = \sum_{k \in \mathbb{Z}^{d'}} a_k e^{2\pi i \langle k, x \rangle},$$

where $\sum_{k \in \mathbb{Z}^{d'}} |a_k|^2 \sum_{j_1 + \ldots + j_{d'} = n} \prod_{i=1}^{d'} |k_i|^{2j_i} < \infty$ and $a_0 = 0$. To solve $f = g - g \circ R_\alpha$ we write

$$g(x) = \sum_{k \in \mathbb{Z}^{d'}} b_k e^{2\pi i \langle k, x \rangle}.$$

By comparing coefficients, this is satisfied for $b_k = \frac{a_k}{1 - e^{2\pi i \langle k, \alpha \rangle}}$ for $k \neq 0$ and $b_0 = 0$. We have

$$\sum_{k \in \mathbb{Z}^{d'}} |b_k|^2 \leq (C')^2 \sum_{k \in \mathbb{Z}^{d'}} |a_k|^2 |k|^{2n} \leq (C')^2 \sum_{k \in \mathbb{Z}} |a_k|^2 \sum_{j_1 + \ldots + j_{d'} = n} \prod_{i=1}^{d'} |k_i|^{2j_i} < \infty.$$

In particular $||g||_{L^2} \leq C' ||f||_{H^n}$.

Thus, for every[30] $h \in H^n(\mathbb{T}^{d'})$ we have

$$\left|\left| \sum_{j=1}^{J} h \circ R_\alpha^j - J \int_{\mathbb{T}^{d'}} h \, \mathrm{d}\lambda^{d'} \right|\right|_{L^2} \leq C'' ||h||_{H^n},$$

where $\lambda^{d'}$ is the d'-dimensional Lebesgue measure on $\mathbb{T}^{d'}$. Due to Remark 3.1(iii), condition (EE) is satisfied with $r = n$ and $\delta = 0$.

We can apply Theorem 3.2 with

$$d > n.$$

11.2. **Horocycle flows.** Consider the classical horocycle flow h_t on compact homogeneous space $\Gamma \backslash PSL(2, \mathbb{R})$ generated by

$$h_t = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}.$$

For fixed $t > 0$ we will consider the time t map $R = h_t$.

Condition $(SLR(y^*))$ follows from the relation $h_{e^{2s}t} = g_s \circ h_t \circ g_{-s}$, where g_s is the geodesic flow

$$g_s = \begin{bmatrix} e^s & 0 \\ 0 & e^{-s} \end{bmatrix}.$$

Indeed, we will show that there is a constant $c > 0$ such that for small enough $r > 0$, $0 < |s| < cr^{-1}$, and $y, y^* \in \Gamma / PSL(2, \mathbb{R})$ with $d(y, y^*) < \frac{c}{2} r^{\frac{1}{2}}$ we have

$$d(h_s y, y^*) \geq \frac{c}{2} r^{\frac{1}{2}}.$$

[30] If $\int h \neq 0$ consider $f = h - \int h$.

By the triangle inequality, it is enough to show $d(h_s y, y) \geq c r^{\frac{1}{2}}$. By compactness choose $c = \inf_x d(h_1 x, x) > 0$ (by Hedlund's Theorem there are no periodic orbits). Let $t = \frac{\log(|s|)}{2}$, then

$$d(h_s y, y) = d(g_t h_{sgn(s)} g_{-t} y, g_t g_{-t} y) \geq e^{-|t|} d(h_{sgn(s)} g_{-t} y, g_{-t} y) \geq c r^{\frac{1}{2}}.$$

For small enough r, g_t contracts distances at most by a factor of $e^{-|t|}$. Renaming $r = \frac{c}{2} r^{\frac{1}{2}}$ we obtain (4) with $\psi(r) = 2c^3 r^{-2}$.

In order to show effective ergodicity (EE), we combine [20, Corollary 2.8] and [18, Theorem 1.5] to conclude that there is a constant $C > 0$ with

$$\left\| \sum_{j=0}^{n-1} f(R^j(y)) - n \int f \, \mathrm{d}\nu \right\| \leq C ||f||_{W^{15}} N^{\frac{5}{6}+\varepsilon} \quad \forall f \in W^s, y \in \Gamma/PSL(2,\mathbb{R}), N \geq 1, \tag{44}$$

for all $\varepsilon > 0$.

Indeed, for $s > 3$, [18, Theorem 1.5] yields

$$\left| \int_0^T f(h_t(y)) \, \mathrm{d}t - T \int_{\Gamma/PSL(2,\mathbb{R}} \varphi \, \mathrm{d}\nu \right| \leq C(s) ||f||_{W^s} T^{\frac{1}{2}} \log(T) \\ \forall f \in W^s, y \in \Gamma/PSL(2,\mathbb{R}), T > 0, \tag{45}$$

for some constant $C(s) > 0$.

A consideration involving twisted integrals as in [20, Corollary 2.8] yields, for $s > 14$,

$$\left| \sum_{n=0}^{N-1} f(h_n(y)) - \int_0^N f(h_t(y)) \, \mathrm{d}t \right| \leq C'(s) ||f||_{W^s} N^{\frac{5}{6}} \log^{\frac{1}{2}}(N) \\ \forall f \in W^s, y \in \Gamma/PSL(2,\mathbb{R}), N \geq 1, \tag{46}$$

for some constant $C'(s) > 0$. Now (45) and (46) together imply (44).

Now, setting $s = 15$ in (46), Theorem 3.2 applies with

$$d > 6(15 + \frac{5}{3}) = 100.$$

11.3. **Skew shifts.** Let $\alpha \in (0,1) \setminus \mathbb{Q}$ satisfy the Diophantine condition (D) for some $n \geq 2$ and $R : \mathbb{T}^2 \to \mathbb{T}^2$ be given by

$$R(x, y) = (x + \alpha, y + x).$$

Since R has a Diophantine rotation as a factor $(SLR(y^*))$ is satisfied by §11.1.

For $k = (k_1, k_2) \in \mathbb{Z}^2$ denote $e_k(x) = e^{2\pi i \langle k, x \rangle}$. Note that

$$\langle e_k, e_{k'} \circ R^j \rangle_{L^2(\mathbb{T}^2)} = \delta^{(k_1,k_2)}_{(k'_1 + j k'_2, k'_2)}.$$

For $f \in H^2(\mathbb{T})$ we can write $f = \sum_{k \in \mathbb{Z}^2} a_k e_k$. If $a_{(k_1,0)} \equiv 0$ (in particular $\int f = 0$) then

$$\sum_{j \geq 1} \left| \langle f, f \circ R^j \rangle_{L^2} \right| = \sum_{j \geq 1} \left| \sum_{k \in \mathbb{Z}^2} a_{(k_1,k_2)} \overline{a_{(k_1 + j k_2, k_2)}} \right| \leq \left(\sum_{k \in \mathbb{Z}^2} |a_{(k_1,k_2)}| \right)^2 \leq C ||f||^2_{H^2},$$

where $C > 0$ does not depend on f. From this we obtain

$$\begin{aligned}\left|\left|\sum_{j=1}^{J} f\circ R^j\right|\right|^2_{L^2(\mathbb{T}^2)} &= \sum_{j=1}^{J}\sum_{j'=1}^{J}\langle f\circ R^{j'}, f\circ R^j\rangle_{L^2(\mathbb{T}^2)}\\ &\le \sum_{j=0}^{J-1}(J-j)\langle f, f\circ R^j\rangle_{L^2(\mathbb{T}^2)} \le CJ||f||^2_{H^2(\mathbb{T}^2)}.\end{aligned}$$

For general $f \in H^n(\mathbb{T}^2)$ with $\int f = 0$, again write $f = \sum_{k\in\mathbb{Z}^2} a_k e_k$ and set

$$f_1 = \sum_{k\in\mathbb{Z}^2, k_2\neq 0} a_k e_k \quad \textit{and} \quad f_2 = \sum_{k\in\mathbb{Z}^2, k_2=0} a_k e_k.$$

Applying the above, and the analysis for Diophantine rotations, we find

$$\left|\left|\sum_{j=1}^{J} f\circ R^j\right|\right|^2_{L^2(\mathbb{T}^2)} \le (CJ + C')||f||^2_{H^n(\mathbb{T}^2)}.$$

Thus, condition (EE) is satisfied with $\delta = \frac{1}{2}$.

So we can apply Theorem 3.2 with

$$d > 2(n+1).$$

11.4. **Example 4.5.** Recall the definition of the Weyl Chamber flow on $\Gamma\backslash SL(d,\mathbb{R})$. Let $d \ge 3$, and Γ be a uniform lattice. Denote by D_+ the subgroup of diagonal elements of $SL(d,\mathbb{R})$ with positive entries. It is easy to see that D_+ is isomorphic to $\mathbb{R}^{d-1}$. D_+ acts on $\Gamma\backslash Sl(d,\mathbb{R})$ by right translation, giving us a $\mathbb{R}^{d-1}$-action. By [6, Theorem 1.1] the action G satisfies (a $\mathbb{R}^{d-1}$ version of) (MEM).

The Diophantine rotation R_α satisfies (EE) and (9) by §11.1. Hence, we can apply Theorem 3.4.

11.5. **Other systems satisfying (EE).** From Example 4.1 it might seem like (EE) is a very special property and only a few systems satisfy this. The opposite is true, in fact, most classical systems have this property.

To convince ourselves of this, let us give some more examples and point out the mechanisms.

Definition 11.1. The system (Y, R, ν) is called *mixing of order* α if, for each $f, g \in C^{r'}$ with $\int_Y f\,\mathrm{d}\nu = \int_Y g\,\mathrm{d}\nu = 0$, we have

$$\left|\int_Y f\circ R^n \cdot g\,\mathrm{d}\nu\right| < ||f||_{C^{r'}}||g||_{C^{r'}}\alpha(n) \quad \forall n \ge 1. \tag{47}$$

We say that (Y, R, ν) is *polynomially mixing* if it is mixing with rate $\alpha(n) = O(n^{-\varepsilon})$ for some $\varepsilon > 0$.

Lemma 11.2. *Polynomial mixing implies (EE). More precisely, if (Y, R, ν) is mixing of order $\alpha(n) = O(n^{-\varepsilon})$, for some $\varepsilon > 0$, then, for all $\varepsilon' > 0$, it satisfies property (EE) with*

$$\delta = \begin{cases} \frac{2-\varepsilon}{2} & \textit{if } \varepsilon < 1\\ \frac{1}{2} + \varepsilon' & \textit{if } \varepsilon = 1\\ \frac{1}{2} & \textit{if } \varepsilon > 1.\end{cases} \tag{48}$$

Proof. For $f \in C^{r'}$ with $\int_Y f \, \mathrm{d}\nu = 0$ we have, for $N \geq 1$,

$$\left|\left|\sum_{n=0}^{N-1} f \circ R^n\right|\right|_{L^2}^2 \leq \sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1}\left|\int_Y f \circ R^{n_1} f \circ R^{n-2} \, \mathrm{d}\nu\right|$$
$$\leq 2N \sum_{n=0}^{N-1}\left|\int_Y f \circ R^n \cdot f \, \mathrm{d}\nu\right| \leq K||f||_{C^{r'}}^2 N^{2\delta}$$

for some $K > 0$. □

Remark 11.3. In fact, the proof above remains valid if (47) holds for all $n \leq N$ except for a subset of $\{1, ..., N\}$ of size $N^{1-\bar{\varepsilon}}$ for some $\bar{\varepsilon} > 0$. We call such systems *polynomially weakly mixing.*[31]

Many classical systems exhibit polynomial (or faster) mixing we list just a few examples referring to [15, Section 8] for a more comprehensive list

- mixing piecewise expanding interval maps [7, Theorem 3.1] as well as expanding interval maps with critical points and singularities [33, Theorem 1.5],
- uniformly hyperbolic systems [31, Theorem 3.9],
- some quadratic maps [41, Theorem 3],
- noncompact translations on finite volume homogeneous spaces of semisimple Lie groups without compact factors [30, §2.4.3],
- time change of horocycle flow [21, Theorem 3].

For parabolic and elliptic systems, one can often use a harmonic analytic argument akin to (but more involved than) §11.1 or 11.3. Other concrete examples include

- nilflows [19, Theorem 1.1],
- almost every interval exchange transformation [2, Theorem 7.1],
- time 1 map of certain smooth surface flows, this follows from a work in progress by the author, where polynomial weak mixing is shown.

12. **Robustness of return times.** Lastly, we mention the proof for the delayed PLT. All of the above proofs can be done using $\Phi_{B_l,\alpha}$ instead of Φ_{B_l}, this shows (with the notation from the proof of Theorem 3.2)

$$(\mu \times \nu)\left(\bigcup_{k=1}^K A_l^{(k)} \times B_l^{(k)}\right)\Phi_{\bigcup_{k=1}^K A_l^{(k)} \times B_l^{(k)},\alpha} \overset{\mu}{\Rightarrow} \Phi_{Exp}.$$

To conclude, we only need a version of the approximation Lemma 7.1 for delayed return times.

Let (M, d_M) be a compact metric space, let $(\tilde{\vartheta}_n)_{n\geq 1}$ be a sequence of Lipschitz functions on M dense in $C(M)$, and denote $\vartheta_n = \frac{\tilde{\vartheta}_n}{||\tilde{\vartheta}_n||_{Lip}}$. The metric

$$D_M(\lambda, \lambda') = \sum_{n\geq 1} 2^{-n}\left|\int_M \vartheta_n \, \mathrm{d}\lambda' - \int_M \vartheta_n \, \mathrm{d}\lambda\right|,$$

for probability measures λ and λ', models distributional convergence[32].

[31] In fact, a slight modification of the proof shows that if R_1 is polynomially mixing and R_2 satisfies (EE) then $R_1 \times R_2$ satisfies (EE).

[32] In the sense that $\lambda_n \Rightarrow \lambda$ if and only if $D_M(\lambda, \lambda_n) \to 0$.

Lemma 12.1. *Let (X,μ,T) be a probability-preserving dynamical system, α be a sequence of natural numbers, $(A_l)_{l\geq 1}$ be a sequence of rare events, and $\Phi = (\phi^{(1)},\phi^{(2)},...)$ be a random process in $[0,\infty)$. Assume that, for each $\delta>0$, there is a sequence of rare events $(A_l^{(\delta)})_{l\geq 1}$ with $A_l^{(\delta)}\subset A_l$ and $\mu_{A_l}(A_l\setminus A_l^{(\delta)})<\delta$ such that*

$$\mu(A_l^{(\delta)})\Phi_{A_l^{(\delta)},\alpha}\overset{\mu}{\Rightarrow}\Phi\quad ;\, as\ l\to\infty,\forall\delta>0.$$

Then

$$\mu(A_l)\Phi_{A_l,\alpha}\overset{\mu}{\Rightarrow}\Phi\quad as\ l\to\infty.$$

Proof. Taking a subsequence if necessary, we may assume that there is a $[0,\infty]$-valued random process Φ' with

$$\mu(A_l)\Phi_{A_l,\alpha}\overset{\mu}{\Rightarrow}\Phi'\quad as\ l\to\infty.$$

For $s,t\in[0,\infty]$ denote $d_{[0,\infty]}(s,t)=|e^{-s}-e^{-t}|$, where by convention $e^{-\infty}=0$, then $([0,\infty],d_{[0,\infty]})$ is a compact metric space. Also, the infinite product $([0,\infty]^{\mathbb{N}},d_{[0,\infty]^{\mathbb{N}}})$ is a compact metric space with $diam([0,\infty]^{\mathbb{N}})=1$, where

$$d_{[0,\infty]^{\mathbb{N}}}((s_j),(t_j))=\sum_{j\geq 1}2^{-j}d_{[0,\infty]}(s_j,t_j).$$

We claim that for every $\varepsilon>0$ there exist $\delta_0>0$ and an $L\geq 1$ such that

$$D_{[0,\infty]^{\mathbb{N}}}\left(law_\mu(\mu(A_l^{(\delta)})\Phi_{A_l^{(\delta)},\alpha}),law_\mu(\mu(A_l)\Phi_{A_l,\alpha})\right)<5\varepsilon\quad\forall l\geq L.\tag{49}$$

Then taking $l\to\infty$ shows $D_{[0,\infty]^{\mathbb{N}}}(\Phi,\Phi')<5\varepsilon$ and the conclusion follows by $\varepsilon\to 0$.

Let $1>\varepsilon>0$. First, note that[33]

$$D_{[0,\infty]^{\mathbb{N}}}\left(law_\mu(\mu(A_l^{(\delta)})\Phi_{A_l^{(\delta)},\alpha}),law_\mu(\mu(A_l)\Phi_{A_l^{(\delta)},\alpha})\right)<\delta,$$

so it is enough to show that there exist $\varepsilon>\delta_0>0$ and an $L\geq 1$ such that

$$D_{[0,\infty]^{\mathbb{N}}}\left(law_\mu(\mu(A_l)\Phi_{A_l^{(\delta)},\alpha}),law_\mu(\mu(A_l)\Phi_{A_l,\alpha})\right)<4\varepsilon\quad\forall l\geq L.\tag{50}$$

Denote $\Phi_{A_l,\alpha}=(\varphi^{(1)}_{A_l,\alpha},\varphi^{(2)}_{A_l,\alpha},...)$ and $\Phi_{A_l^{(\delta)},\alpha}=(\varphi^{(1)}_{A_l^{(\delta)},\alpha},\varphi^{(2)}_{A_l^{(\delta)},\alpha},...)$. Now choose $J\geq 1$ so big that $\sum_{j\geq J}2^{-j}<\varepsilon$, and $T>0$ such that

$$\mathbb{P}\left(\sum_{j=1}^{J}\phi^{(j)}>T\right)<\varepsilon.$$

For $\delta=\min\left(\frac{\varepsilon}{2},\frac{\varepsilon}{2T}\right)$ choose $L\geq 1$ so big that

$$\mu\left(\sum_{j=1}^{J}\mu(A_l^{(\delta)})\varphi^{(j)}_{A_l^{(\delta)},\alpha}>T\right)<2\varepsilon\quad\forall l\geq L.$$

Since $\sum_{j=1}^{J}\varphi^{(j)}_{A_l^{(\delta)},\alpha}>\sum_{j=1}^{J}\varphi^{(j)}_{A_l,\alpha}$ and $\mu(A_l^{(\delta)})>(1-\delta)\mu(A_l)>\frac{1}{2}\mu(A_l)$, in particular

$$\mu\left(\sum_{j=1}^{J}\mu(A_l)\varphi^{(j)}_{A_l,\alpha}>2T\right)<2\varepsilon\quad\forall l\geq L.$$

[33] For $k\geq 1$ and $s,t\in[0,\infty]$ we have $d_{[0,\infty]}(ks,kt)\geq d_{[0,\infty]}(s,t)\leq|s-t|$.

Now, for $j = 1, ..., J$, we have

$$\mu\left(\varphi^{(j)}_{A_l,\alpha} \neq \varphi^{(j)}_{A_l^{(\delta)},\alpha}\right) \leq \mu\left(\varphi^{(J)}_{A_l,\alpha} \neq \varphi^{(J)}_{A_l^{(\delta)},\alpha}\right)$$
$$\leq \mu\left(\bigcup_{i=1}^{\left\lfloor \frac{2T}{\mu(A_l)}\right\rfloor} T^{-\tilde{\alpha}^{(i)}}(A_l \setminus A_l^{(\delta)})\right) + \mu\left(\sum_{j=1}^{J} \mu(A_l)\varphi^{(j)}_{A_l,\alpha} > 2T\right)$$
$$\leq \frac{2T}{\mu(A_l)}\mu(A_l)\delta + 2\varepsilon \leq 3\varepsilon.$$

Thus

$$D_{[0,\infty]^{\mathbb{N}}}\left(law_\mu(\mu(A_l)\Phi_{A_l^{(\delta)},\alpha}), law_\mu(\mu(A_l)\Phi_{A_l,\alpha})\right)$$
$$= \sum_{j\geq 1} 2^{-j}\int_X d_{[0,\infty]}(\mu(A_l)\varphi^{(j)}_{A_l,\alpha}, \mu(A_l)\varphi^{(j)}_{A_l^{(\delta)},\alpha})\,\mathrm{d}\mu \leq 3\varepsilon\sum_{j=1}^{J} 2^{-j} + \varepsilon \leq 4\varepsilon$$

proving (50). □

To conclude, we give a

Proof of Lemma 7.1. Denote $\lambda = \mu \times \nu$ and $M = [0, \infty]$, with $D_{M^{\mathbb{N}}}$ as in the proof of Lemma 12.1 we have

$$D_{M^{\mathbb{N}}}(law_{\lambda_Q}(\lambda(Q)\Phi_Q), law_{\lambda_{Q'}}(\lambda(Q')\Phi_{Q'})$$
$$\leq \lambda_Q(Q')D_{M^{\mathbb{N}}}(law_{\lambda_Q}(\lambda(Q)\Phi_Q), law_{\lambda_Q}(\lambda(Q')\Phi_{Q'})$$
$$+ \lambda_Q(Q \setminus Q')D_{M^{\mathbb{N}}}(law_{\lambda_Q}(\lambda(Q)\Phi_Q), law_{\lambda_{Q'}}(\lambda(Q')\Phi_{Q'})$$
$$\leq D_{M^{\mathbb{N}}}(law_{\lambda_Q}(\lambda(Q)\Phi_Q), law_{\lambda_Q}(\lambda(Q')\Phi_{Q'}) + \lambda_Q(Q \setminus Q').$$

Furthermore

$$D_{M^{\mathbb{N}}}(law_{\lambda_Q}(\lambda(Q)\Phi_Q), law_{\lambda_Q}(\lambda(Q')\Phi_{Q'})$$
$$\leq \sum_{j\geq 0} 2^{-j-1}\int_Q d_M(\lambda(Q)\varphi_Q \circ S_Q^j, \lambda(Q')\varphi_{Q'} \circ S_{Q'}^j)\,\mathrm{d}\lambda_Q$$
$$\leq \sum_{j\geq 0} 2^{-j-1}\int_Q d_M(\lambda(Q)\varphi_Q \circ S_Q^j, \lambda(Q)\varphi_{Q'} \circ S_{Q'}^j)\,\mathrm{d}\lambda_Q + \lambda_Q(Q \setminus Q').$$

For each $j \geq 0$ we have

$$\lambda_Q(\varphi_Q \circ S_Q^j \neq \varphi_{Q'} \circ S_{Q'}^j) \leq \lambda_Q\left(\bigcup_{i=0}^{j} S_Q^i(Q \setminus Q')\right) \leq (j+1)\lambda_Q(Q \setminus Q').$$

The claim follows since $\sum_{j\geq 1} j2^{-j} = 2$. □

Acknowledgments. Special thanks go to my advisor Dmitry Dolgopyat for suggesting this problem and providing helpful comments throughout. Thanks also go to Adam Kanigowski, for supplying his knowledge on many of the examples.

REFERENCES

[1] M. Abadi and B. Saussol, Hitting and returning to rare events for all alpha-mixing processes, *Stochastic Process. Appl.*, **121** (2011), 314-323.

[2] A. Avila, G. Forni and P. Safaee, Quantitative weak mixing for interval exchange transformations, *Geom. Funct. Anal.*, **33** (2023), 1-56.

[3] L. Barreira and B. Saussol, Hausdorff dimension of measures via Poincaré recurrence, *Comm. Math. Phys.*, **219** (2001), 443-463.

[4] V. Beresnevich, F. Ramírez and S. Velani, Metric Diophantine approximation: Aspects of recent work, *Dynamics and Analytic Number Theory*, *London Math. Soc. Lecture Note Ser.*, vol. 437, Cambridge Univ. Press, Cambridge, 2016, 1-95.

[5] P. Billingsley, *Probability and Measure*, John Wiley & Sons, 1995.

[6] M. Björklund, M. Einsiedler and A. Gorodnik, Quantitative multiple mixing, *J. Eur. Math. Soc. (JEMS)*, **22** (2020), 1475-1529.

[7] H. Bruin, B. Saussol, S. Troubetzkoy and S. Vaienti, Return time statistics via inducing, *Ergodic Theory Dynam. Systems*, **23** (2003), 991-1013.

[8] L. A. Bunimovich and Y. Su, Maximal large deviations and slow recurrences in weakly chaotic systems, *Advances in Mathematics*, **432** (2023), Paper No. 109267, 58 pp.

[9] L. A. Bunimovich and Y. Su, Back to boundaries in billiards, *Comm. Math. Phys.*, **405** (2024), Paper No. 140, 74 pp.

[10] J.-R. Chazottes and P. Collet, Poisson approximation for the number of visits to balls in non-uniformly hyperbolic dynamical systems, *Ergodic Theory Dynam. Systems*, **33** (2013), 49-80.

[11] P. Collet and A. Galves, Statistics of close visits to the indifferent fixed point of an interval map, *J. Statist. Phys.*, **72** (1993), 459-478.

[12] M. Denker, M. Gordin and A. Sharova, A Poisson limit theorem for toral automorphisms, *Illinois J. Math.*, **48** (2004), 1-20.

[13] D. Dolgopyat, Limit theorems for partially hyperbolic systems, *Trans. Amer. Math. Soc.*, **356** (2004), 1637-1689.

[14] D. Dolgopyat, C. Dong, A. Kanigowski and P. Nándori, Flexibility of statistical properties for smooth systems satisfying the central limit theorem, *Invent. Math.*, **230** (2022), 31-120.

[15] D. Dolgopyat, C. Dong, A. Kanigowski and P. Nándori, Mixing properties of generalized T, T^{-1} transformations, Israel *J. Math.*, **247** (2022), 21-73.

[16] D. Dolgopyat, B. Fayad and S. Liu, Multiple Borel-Cantelli lemma in dynamics and multilog law for recurrence, *J. Mod. Dyn.*, **18** (2022), 209-289.

[17] D. Dolgopyat, A. Kanigowski and F. Rodriguez Hertz, Exponential mixing implies Bernoulli, *Annals of Mathematics*, **199** (2024), 1225-1292.

[18] L. Flaminio and G. Forni, Invariant distributions and time averages for horocycle flows, *Duke Math. J.*, **119** (2003), 465-526.

[19] L. Flaminio and G. Forni, Equidistribution of nilflows and applications to theta sums, *Ergodic Theory Dynam. Systems*, **26** (2006), 409-433.

[20] L. Flaminio, G. Forni and J. Tanis, Effective equidistribution of twisted horocycle flows and horocycle maps, *Geom. Funct. Anal.*, **26** (2016), 1359-1448.

[21] G. Forni and C. Ulcigrai, Time-changes of horocycle flows, *Journal of Modern Dynamics*, **6** (2012), 251-273.

[22] A. C. M. Freitas, J. M. Freitas and M. Todd, Extreme value laws in dynamical systems for non-smooth observations, *J. Stat. Phys.*, **142** (2011), 108-126.

[23] A. C. M. Freitas, J. M. Freitas and S. Vaienti, Extreme value laws for non stationary processes generated by sequential and random dynamical systems, *Ann. Inst. Henri Poincaré Probab. Stat.*, **53** (2017), 1341-1370.

[24] N. Haydn, Y. Lacroix and S. Vaienti, Hitting and return times in ergodic dynamical systems, *Ann. Probab.*, **33** (2005), 2043-2050.

[25] N. Haydn, M. Nicol, T. Persson and S. Vaienti, A note on Borel-Cantelli lemmas for non-uniformly hyperbolic dynamical systems, *Ergodic Theory Dynam. Systems*, **33** (2013), 475-498.

[26] N. Haydn and S. Vaienti, The compound Poisson distribution and return times in dynamical systems, *Probab. Theory Related Fields*, **144** (2009), 517-542.

[27] N. T. A. Haydn, J. Rousseau and F. Yang, Exponential law for random maps on compact manifolds, *Nonlinearity*, **33** (2020), 6760–6789.

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

[29] M. Hussain, B. Li, D. Simmons and B. Wang, Dynamical Borel-Cantelli lemma for recurrence theory, *Ergodic Theory Dynam. Systems*, **42** (2022), 1994-2008.

[30] D. Y. Kleinbock and G. A. Margulis, Bounded orbits of nonquasiunipotent flows on homogeneous spaces, *Sinaĭ's Moscow Seminar on Dynamical Systems*, Amer. Math. Soc. Transl. Ser. 2, vol. 171, Amer. Math. Soc., Providence, RI, 1996, 141-172.

[31] C. Liverani, Decay of correlations, *Annals of Mathematics*, **142** (1995), 239-301.

[32] V. Lucarini, D. Faranda, A. C. M. Freitas, J. M. Freitas, M. Holland, T. Kuna, M. Nicol, M. Todd and S. Vaienti, *Extremes and Recurrence in Dynamical Systems*, Pure and Applied Mathematics (Hoboken), John Wiley & Sons, Inc., Hoboken, NJ, 2016.

[33] S. Luzzatto and I. Melbourne, Statistical properties and decay of correlations for interval maps with critical points and singularities, Comm. *Math. Phys.*, **320** (2013), 21-35.

[34] M. J. Pacifico and F. Yang, Hitting times distribution and extreme value laws for semi-flows, *Discrete Contin. Dyn. Syst.*, **37** (2017), 5861-5881.

[35] F. Pène and B. Saussol, Poisson law for some non-uniformly hyperbolic dynamical systems with polynomial rate of mixing, *Ergodic Theory Dynam. Systems*, **36** (2016), 2602-2626.

[36] F. Pène and B. Saussol, Spatio-temporal Poisson processes for visits to small sets, *Israel J. Math.*, **240** (2020), 625-665.

[37] B. Pitskel, Poisson limit law for Markov chains, *Ergodic Theory Dynam. Systems*, **11** (1991), 501-513.

[38] J. Rousseau, B. Saussol and P. Varandas, Exponential law for random subshifts of finite type, *Stochastic Process. Appl.*, **124** (2014), 3260-3276.

[39] M. Rychlik, Bounded variation and invariant measures, *Studia Mathematica*, **76** (1983), 69-80.

[40] Y. Su and L. A. Bunimovich, Poisson approximations and convergence rates for hyperbolic dynamical systems, *Comm. Math. Phys.*, **390** (2022), 131-168.

[41] L.-S. Young, Decay of correlations for certain quadratic maps, *Comm. Math. Phys.*, **146** (1992), 123-138.

[42] R. Zweimüller, The general asymptotic return-time process, *Israel J. Math.*, **212** (2016), 1-36.

[43] R. Zweimüller, Hitting times and positions in rare events, *Ann. H. Lebesgue*, **5** (2022), 1361-1415.

Received December 2023; revised August 2024; early access October 2024.