JOURNAL OF MODERN DYNAMICS
VOLUME 18, 2022, 209–289

doi: 10.3934/jmd.2022009

MULTIPLE BOREL–CANTELLI LEMMA IN DYNAMICS AND MULTILOG LAW FOR RECURRENCE

DMITRY DOLGOPYAT, BASSAM FAYAD AND SIXU LIU

(Communicated by Dmitry Kleinbock)

ABSTRACT. A classical Borel–Cantelli Lemma gives conditions for deciding whether an infinite number of rare events will happen almost surely. In this article, we propose an extension of Borel–Cantelli Lemma to characterize the multiple occurrence of events on the same time scale. Our results imply multiple Logarithm Laws for recurrence and hitting times, as well as Poisson Limit Laws for systems which are exponentially mixing of all orders. The applications include geodesic flows on compact negatively curved manifolds, geodesic excursions on finite volume hyperbolic manifolds, Diophantine approximations and extreme value theory for dynamical systems.

CONTENTS

Received January 15, 2021; revised October 9, 2021.
2020 *Mathematics Subject Classification:* Primary: 37A50; Secondary: 11K60, 60F20.
Key words and phrases: Borel–Cantelli Lemma, recurrence, hitting times, Poisson limit laws, exponentially mixing systems, geodesic excursions, Diophantine approximations.

IN THE PUBLIC DOMAIN AFTER 2050

©2022 AIMSCIENCES

1. INTRODUCTION

The study of rare events constitutes an important subject in probability theory. On one hand, in many applications there are significant costs associated to certain rare events, so one needs to know how often those events occur. On the other hand, there are many phenomena in science which are driven by rare events including metastability, anomalous diffusion (Levy flights), and traps for motion in random media, to mention just a few examples.

In the independent setting there are three classical regimes. For the first two, consider an array $\{\Omega_n^k\}_{k=1}^n$ of independent events such that $p_n = \mathbb{P}(\Omega_n^k)$ does not depend on k. Let N_n be the number of events from the n-th row of the array which have occurred. The first two regimes are:

(i) *CLT regime:* $np_n \to \infty$. In this case N_n is asymptotically normal.
(ii) *Poisson regime:* $np_n \to \lambda$. In this case N_n is asymptotically Poisson with parameter λ.

For the third, *Borel–Cantelli regime* we consider a sequence (Ω_n) of independent events with different probabilities. The classical Borel–Cantelli Lemma says that infinitely many Ω_ns occur if and only if $\sum_n \mathbb{P}(\Omega_n) = \infty$.

A vast literature is devoted to extending the above classical results to the case where independence is replaced by weak dependence. In particular, there are convenient moment conditions which imply similar results for *weakly* dependent events. One important distinction between the Poisson regime and the other two regimes, is that the Poisson regime requires additional geometric conditions on close-by events to extend the statement to the dependent case. Without such conditions, one can have clusters of rare events where the number of clusters has Poisson distribution while several events may occur inside each cluster. We refer the reader to [5] for a comprehensive discussion of Poisson clustering.

1.1. **The multiple Borel–Cantelli Lemma.**

In the present paper, we consider a regime which is intermediate between the Poisson and Borel–Cantelli. Namely, we consider for the moment a family of events Ω_ρ^n which are nested: $\Omega_{\rho_1}^n \subset \Omega_{\rho_2}^n$ for $\rho_1 < \rho_2$, and that independently of n, $\mathbb{P}(\Omega_\rho^n) = \sigma(\rho)$ for some function σ. (We will see later how these conditions could be slightly weakened to accommodate several applications.) For $\rho > 0$, we define

$$N_\rho^n = \sharp\left\{k : 1 \le k \le n,\ \Omega_\rho^k \text{ occurs}\right\}.$$

The main subject of this paper is to assess, for a sequence (ρ_n) such that $n\sigma(\rho_n)$ goes to 0 as $n \to \infty$, and $r \in \mathbb{N}$, whether the event

$$N_{\rho_n}^n \ge r$$

occurs infinitely many times or not. Even if the events Ω_ρ^n are independent for different n, the variables $N_{\rho_{n_1}}^{n_1}$ and $N_{\rho_{n_2}}^{n_2}$ are strongly dependent if n_1 and n_2 are of the same order. On the other hand, if $n_2 \gg n_1$ then the variables are weakly dependent since, conditioned on $N_{\rho_{n_2}}^{n_2} \ne 0$, it is very likely that all the events

$\Omega^k_{\rho_{n_2}}$ actually occur for $k > n_1$. Using this, one can show under appropriate monotonicity assumptions (see [130]) that $N^n_{\rho_n} \geq r$ infinitely often if and only if

$$\sum_M \mathbb{P}(N^{2^M}_{\rho_{2^M}} = r) = \infty.$$

Under the condition $n\sigma(\rho_n) \to 0$, it follows that in the independent case

$$\mathbb{P}(N^n_{\rho_n} = r) \approx \frac{(n\sigma(\rho_n))^r}{r!}.$$

Therefore, under independence, infinitely many $N^n_{\rho_n} \geq r$ occur if and only if

$$\sum_M 2^{Mr} \left(\sigma\left(\rho_{2^M}\right)\right)^r = \infty. \tag{1.1}$$

This multiple Borel–Cantelli Lemma was extended to the dependent setting in [1]. However, the mixing assumptions made in [1] are quite strong requiring good symbolic dynamics which limits greatly the applicability of that result. In the present paper we present an abstract extension of the multiple Borel–Cantelli Lemma to the dependent variable setting. The dynamical versions that can be extracted from our abstract result use more flexible mixing conditions which open up many interesting applications. Our conditions are similar to the assumptions typically used to prove Poisson limit theorems for dynamical systems.

The precise statements of our abstract results will be given in Sections 2 and 3.

Let us describe in a nutshell the setting for dynamical applications. We will state mixing and regularity conditions on a dynamical system (f, X, μ) and on a family of target sets $\{\Omega_\rho\}, \rho \in \mathbb{R}^*_+$ so that given a (decreasing) sequence (ρ_n), and defining the sets $\Omega^k_{\rho_n} = f^{-k}\Omega_{\rho_n}$, we get the validity of the dichotomy described in (1.1) for the number of hits $N^n_{\rho_n}$. More precisely, when a point y is randomly distributed according to μ, we view $\Omega^k_{\rho_n}$ as an event, and saying that infinitely many $N^n_{\rho_n} \geq r$ occur almost surely just means that for μ-a.e. $y \in X$, there are infinitely many n so that $y \in \Omega^k_{\rho_n}$ for at least r distinct $k \in [1, n]$.

Let us now describe some sample applications to dynamics, geometry, and number theory that will be developed in separate sections after the abstract results are stated and proved.

1.2. **MultiLog Law for recurrence.** Let f be a map preserving a measure μ on a metric space (M, d). Given two points x, y let $d^{(r)}_n(x, y)$ be the r closest distance among $d(x, f^k y)$ for $1 \leq k \leq n$. In particular, $d^{(1)}_n(x, y)$ is the closest distance the orbit of y comes to x up to time n. It is shown in [68] that for systems with superpolynomial decay for Lipschitz observables, for all x and μ-almost all y

$$\lim_{n\to\infty} \frac{|\ln d^{(1)}_n(x, y)|}{\ln n} = \frac{1}{\mathbf{d}},$$

where $\mathbf{d}$ is the local dimension of μ at x provided that it exists.

Under some additional assumptions, one can prove a dynamical Borel–Cantelli Lemma which implies in particular that, if μ is smooth then for all x and almost all y we have

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(1)}(x,y)| - \frac{1}{\mathbf{d}}\ln n}{\ln\ln n} = \frac{1}{\mathbf{d}}.$$

In Section 4 we extend this result to $r>1$, for systems that have multiple exponential mixing properties. For example, if f is an expanding map of the circle, we shall show that for Lebesgue almost all $(x,y)\in\mathbb{T}\times\mathbb{T}$ we have

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,y)| - \ln n}{\ln\ln n} = \frac{1}{r}. \tag{1.2}$$

We sketch briefly the reduction of (1.2) to the multiple Borel–Cantelli lemma leaving the details to Section 4. Given $x\in M$, let $\Omega_{x,\rho} = \{y : d(x,y)\le\rho\}$. We use the notation $\Omega_{x,\rho}^k$ for the event $1_{\Omega_{x,\rho}}\circ f^k$. We let $\sigma(\rho) = \mathrm{Leb}(\Omega_{x,\rho})$.

For $s\ge 0$, we let $\rho_n = n^{-1}(\ln n)^{-s}$, and for $y\in\mathbb{T}$, we denote by $N_{\rho_n}^n(y)$ the number of times $k\in[1,n]$ such that $y\in\Omega_{x,\rho_n}^k$.

Thus (1.2) is equivalent to the following :

(a) If $s>\frac{1}{r}$, then for Lebesgue almost all $(x,y)\in\mathbb{T}\times\mathbb{T}$, we have that for large n, $N_{\rho_n}^n<r$.

(b) If $s\le\frac{1}{r}$, then for Lebesgue almost all $(x,y)\in\mathbb{T}\times\mathbb{T}$, there are infinitely many n such that $N_{\rho_n}^n\ge r$.

With the notation $\mathbf{S}_r = \sum_{j=1}^{\infty} 2^{rj}\sigma(\rho_{2^j})^r$, we see that $\mathbf{S}_r=\infty$ if and only if $s\le\frac{1}{r}$. Hence, (1.2) would follow from an extension of the Multiple Borel–Cantelli Lemma of §1.1 to the case of expanding maps of the circle.

The smoothness assumption on the invariant measure, the Lebesgue typicality assumption on x and the hyperbolicity assumption on f are all essential. Namely, if μ is an invariant Gibbs measure which is **not** conformal, λ is the Lyapunov exponent of μ, then we show in Section 6 that for μ almost all x and y and for all $r\in\mathbb{N}$,

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,y)| - \ln n}{\sqrt{2(\ln n)(\ln\ln\ln n)}} = \frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}$$

for some $\sigma>0$ which will be given in (6.5). We shall also show that there is G_δ–dense set $\mathscr{H}$ such that for all $x\in\mathscr{H}$, Lebesgue almost all y and all $r\ge 1$, we have

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,y)| - \ln n}{\ln\ln n} = 1.$$

Finally if the expanding map is replaced by a rotation T_α then we have (see Theorem 4.7 below) that for almost all (x,y,α)

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,y)| - \ln n}{\ln\ln n} = \begin{cases} 1 & \text{if } r=1,\\ \frac{1}{2} & \text{if } r>1.\end{cases}$$

1.3. **Records of geodesic excursions.** Consider a hyperbolic manifold $\mathcal{Q}$ of dimension $d+1$ which is not compact but has finite volume. Such manifold admits a thick-thin decomposition. Namely $\mathcal{Q}$ is a union of a compact part and several cusps (see §7.1, in particular, formula (7.3) for more details on think-thin decomposition). *A cusp excursion* is a maximal time segment such that the geodesic stays in a cusp for the whole segment. Let

$$H^{(1)}(T) \geq H^{(2)}(T) \geq \dots H^{(r)}(T) \geq \dots$$

be the maximal heights achieved during the excursions which occur before time T placed in the decreasing order. *Sullivan's Logarithm Law* is equivalent to saying that for almost every geodesic

$$\limsup_{T\to\infty} \frac{H^{(1)}(T)}{\ln T} = \frac{1}{d}. \tag{1.3}$$

The proof of (1.3) relies on Sullivan's Borel–Cantelli Lemma and it actually also shows that for almost every geodesic

$$\limsup_{T\to\infty} \frac{H^{(1)}(T) - \frac{1}{d}\ln T}{\ln\ln T} = \frac{1}{d}.$$

Extending the Multiple Borel–Cantelli Lemma of §1.1 to the cusp excursions, we obtain a multiple version of Sullivan's law by showing that for almost every geodesic

$$\limsup_{T\to\infty} \frac{H^{(r)}(T) - \frac{1}{d}\ln T}{\ln\ln T} = \frac{1}{rd}.$$

1.4. **Multiple Khintchine–Groshev Theorem.** Let $\psi : \mathbb{R} \to \mathbb{R}$ be a positive function (in dimension 1 we also assume that ψ is monotone). The classical Khintchine–Groshev Theorem ([76, 103, 149]) says that for almost all $\alpha \in \mathbb{R}^d$ there are infinitely many solutions to

$$|\langle k, \alpha\rangle + m| \leq \psi(\|k\|_\infty) \text{ with } k \in \mathbb{Z}^d, m \in \mathbb{Z} \tag{1.4}$$

if and only if

$$\sum_{r=1}^{\infty} r^{d-1}\psi(r) = \infty. \tag{1.5}$$

In particular the inequality

$$|k|^d |\langle k, \alpha\rangle + m| \leq \frac{1}{\ln|k| (\ln\ln|k|)^s}$$

where $|k| = \sqrt{\sum k_i^2}$, has infinitely many solutions for almost every α if and only if $s \leq 1$. This is one of Khintchine–Groshev $0-1$ laws for Diophantine approximations of linear forms that can all be obtained from suitable dynamical extensions of the classical Borel–Cantelli Lemma to cusp excursions of appropriate diagonal actions on the space of lattices.

One goal of this paper is to extend the Khintchine–Groshev $0-1$ laws to multiple Diophantine approximations. For example, we can replace (1.4) by

$$|k|^d |\langle k, \alpha\rangle + m| \le \frac{1}{\ln N (\ln\ln N)^s}, \ |k| \le N \tag{1.6}$$

and say that α is (r, s) approximable if there are infinitely many Ns for which (1.6) has r positive solutions (that is, solutions with $k_1 > 0$).

In Section 9, we give several versions of the Multiple Borel–Cantelli Lemma for cusp excursions of diagonal actions on the space of lattices, and obtain in one of the applications that almost every $\alpha \in \mathbb{R}^d$ is (r, s) approximable if and only if $s \le \frac{1}{r}$.

1.5. **Plan of the paper.** The layout of the paper is the following. In Section 2 we describe an abstract result on an array of rare events in a probability space which ensures that for a given r, r events in the same row happen for infinitely many (respectively, finitely many) rows. In Section 3 this abstract criterion is applied in the case of rare events that consist of visits to a sublevel set of a Lipschitz function by the orbits of a smooth exponentially mixing dynamical system. The results of Section 3 are then used to obtain MultiLog Laws in various settings. Namely, Section 4 studies hitting and return times for multifold exponentially mixing smooth systems. Section 8 treats similar problems in the configuration space for the geodesic flows on compact negatively curved manifolds. Geodesic excursions are discussed in Section 7, and Diophantine approximations are treated in Section 9. The MultiLog Law for non-conformal measures is discussed in Section 6.

As it was mentioned, the regime we consider is intermediate between the Poisson and Borel–Cantelli. Section 5 contains an application of our results to the Poisson regime. Namely we derive Poisson distribution for hits and mixed Poisson distribution for returns for exponentially mixing systems on smooth manifolds. Section 10 describes the application of our results to the extreme value theory for dynamical systems. Each section ends with some notes where the related literature is discussed.

Some useful auxiliary results are collected in the appendices.

2. MULTIPLE BOREL–CANTELLI LEMMA

2.1. **The result.** The classical Borel–Cantelli Lemma is a standard tool for deciding when an infinite number of rare events occur with probability one. However in case an infinite number of events do occur, the Borel–Cantelli Lemma does not give an information about how well separated in time those occurrences are. In this section we present a criterion which allows to decide when several rare events occur on the same time scale. The criterion is based on various independence conditions between the rare events.

DEFINITION 2.1. Consider a probability space $(\Omega,\mathscr{F},\mathbb{P})$. Given $r\in\mathbb{N}^*$, a sequence (ρ_n) and a family of events $(\Omega^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$, we let $N^n_{\rho_n}$ be the number of times $k\le n$ such that $\Omega^k_{\rho_n}$ occurs, i.e.,

$$N^n_{\rho_n}=\sharp\left\{k:1\le k\le n,\ \Omega^k_{\rho_n}\ \text{ occurs}\right\}.$$

DEFINITION 2.2 (Shrinking targets)**.** When the sequence (ρ_n) is decreasing and

$$\Omega^n_{\rho_1}\subset\Omega^n_{\rho_2}\ \text{if}\ \rho_1\le\rho_2, \tag{2.1}$$

we say that the sequence of targets $(\Omega^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$ is shrinking.

REMARK 2.3. In all our applications, the targets will be shrinking. However, part of our results will not require this condition.

Our goal is to give a criterion that allows to tell exactly when almost surely $N^n_{\rho_n}\ge r$ will hold for infinitely many n. For this, we introduce several conditions quantifying asymptotic independence between the events $\Omega^k_{\rho_n}$. The statement of the conditions requires the existence of:

- an increasing function $\sigma:\mathbb{R}_+\to\mathbb{R}_+$,
- a sequence $\varepsilon_n\to 0$,
- a function $\mathfrak{s}:\mathbb{N}\to\mathbb{N}$ such that $\mathfrak{s}(n)\le(\ln n)^2$,
- a function $\hat{\mathfrak{s}}:\mathbb{N}\to\mathbb{N}$ such that $\varepsilon n\le\hat{\mathfrak{s}}(n)<n(1-q)/(2r)$ for some $0<q<1$, and some $0<\varepsilon<(1-q)/(2r)$,

for which the following holds. For an arbitrary r-tuple $0\le k_1<k_2\cdots<k_r\le n$ we consider the *separation indices*

$$\mathrm{Sep}_n(k_1,\dots,k_r)=\ \mathrm{Card}\left\{j\in\{0,\dots r-1\}:k_{j+1}-k_j\ge\mathfrak{s}(n)\right\},\quad k_0:=0,$$
$$\widehat{\mathrm{Sep}}_n(k_1,\dots,k_r)=\ \mathrm{Card}\left\{j\in\{0,\dots r-1\}:k_{j+1}-k_j\ge\hat{\mathfrak{s}}(n)\right\},\quad k_0:=0.$$

$(M1)_r$ If $0\le k_1<k_2<\dots k_r\le n$ are such that $\mathrm{Sep}_n(k_1,\dots,k_r)=r$ then

$$\sigma(\rho_n)^r(1-\varepsilon_n)\le\mathbb{P}\left(\bigcap_{j=1}^r\Omega^{k_j}_{\rho_n}\right)\le\sigma(\rho_n)^r(1+\varepsilon_n).$$

$(M2)_r$ There exists $K>0$ such that if $0\le k_1<k_2<\dots k_r\le n$ are such that $\mathrm{Sep}_n(k_1,\dots,k_r)=m<r$, then

$$\mathbb{P}\left(\bigcap_{j=1}^r\Omega^{k_j}_{\rho_n}\right)\le\frac{K\sigma(\rho_n)^m}{(\ln n)^{100r}}.$$

$(M3)_r$ If $0\le k_1<k_2<\cdots<k_r<l_1<l_2<\cdots<l_r$, are such that $2^i<k_\alpha\le 2^{i+1}, 2^j<l_\beta\le 2^{j+1}$, for $1\le\alpha,\beta\le r,\ j-i\ge b$ for some constant $b\ge 1$, and such that

$$\widehat{\mathrm{Sep}}_{2^{i+1}}(k_1,\dots,k_r)=r,\quad\widehat{\mathrm{Sep}}_{2^{j+1}}(l_1,\dots,l_r)=r,\quad l_1-k_r\ge\hat{\mathfrak{s}}(2^{j+1}),$$

then

$$\mathbb{P}\left(\left[\bigcap_{\alpha=1}^r\Omega^{k_\alpha}_{\rho_{2^i}}\right]\bigcap\left[\bigcap_{\beta=1}^r\Omega^{l_\beta}_{\rho_{2^j}}\right]\right)\le\sigma(\rho_{2^i})^r\sigma(\rho_{2^j})^r(1+\varepsilon_i).$$

DEFINITION 2.4. For $r \in \mathbb{N}^*$ and a sequence (ρ_n), we say that the events of the family $(\Omega^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$ are $2r$*–almost independent at a fixed scale* if $(M1)_{\bar r}$ and $(M2)_{\bar r}$ are satisfied for every $\bar r \in [1,2r]$. We say that $(\Omega^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$ are $2r$*–almost independent at all scales* if $(M1)_{\bar r}$, $(M2)_{\bar r}$ are satisfied for every $\bar r \in [1,2r]$, and $(M3)_{\bar r}$ is satisfied for every $\bar r \in [1,r]$.

REMARK 2.5. Conditions $(M1)_r$ and $(M3)_r$ are mixing conditions. In the examples considered in this paper they will follow from multiple exponential mixing. Condition $(M2)_r$ is a *non-clustering* condition and it has to be verified on case-by-case basis using the geometry of the targets. We note that $(M2)_r$ holds if, for fixed ρ, the events Ω^k_ρ are *quasi-independent* in the sense that

$$\mathbb{P}\left(\bigcap_{j=1}^r \Omega^{k_j}_\rho\right) \le C(r)\prod_{j=1}^r \mathbb{P}\left(\Omega^{k_j}_\rho\right).$$

The quasi-independence also plays a crucial role in the classical work of Sullivan on dynamical Borel–Cantelli Lemma [152].

THEOREM 2.6. *Given a sequence (ρ_n) and a family of events $(\Omega^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$, define*

$$\mathbf{S}_r = \sum_{j=1}^\infty \left(2^j \sigma(\rho_{2^j})\right)^r.$$

- *(a) If $\mathbf{S}_r < \infty$, and $(\Omega^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$ are shrinking as in Definition 2.2 and are $2r$–almost independent at a fixed scale, then with probability 1, we have that for large n, $N^n_{\rho_n} < r$.*
- *(b) If $\mathbf{S}_r = \infty$, and $(\Omega^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$ are $2r$–almost independent at all scales then with probability 1, there are infinitely many n such that $N^n_{\rho_n} \ge r$.*

REMARK 2.7. An analogous statement has been obtained in [1] under different mixing conditions.

REMARK 2.8. Observe that if the targets are shrinking, we have that

$$\sum_{n=2^j}^{2^{j+1}-1} \sigma^r(\rho_n) n^{r-1} \le \left(2^{j+1}\sigma(\rho_{2^j})\right)^r \le 2^{2r}\sum_{n=2^{j-1}}^{2^j-1} \sigma^r(\rho_n) n^{r-1},$$

in which case the convergence of $\mathbf{S}_r$ is equivalent to the convergence of

$$\sum_{n=1}^\infty \sigma^r(\rho_n) n^{r-1}.$$

2.2. **Estimates on a fixed scale.** For $m \in \mathbb{N}$ let

$$\mathscr{U}_m = \{(k_1,\dots,k_r) \text{ s.t. } 2^m < k_1 < k_2 < \cdots < k_r \le 2^{m+1} \text{ and } \widehat{\mathrm{Sep}}_{2^{m+1}}(k_1,\dots k_r) = r\},$$

$$\mathscr{A}_m := \{\exists\, 0 < k_1 < \cdots < k_r \le 2^{m+1} \text{ s.t. } \Omega^{k_\alpha}_{\rho_{2^m}} \text{happens for any } \alpha \in [1,r]\},$$

$$\mathscr{D}_m := \{\exists\,(k_1,\dots,k_r) \in \mathscr{U}_m \text{ s.t. } \Omega^{k_\alpha}_{\rho_{2^{m+1}}} \text{happens for any } \alpha \in [1,r]\}.$$

The goal of this section is to prove the following estimates from which it will be easy to derive Theorem 2.6.

PROPOSITION 2.9. *Suppose*

$$n\sigma(\rho_n) \to 0 \quad \text{as} \quad n \to \infty. \tag{2.2}$$

If $(\Omega^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$ *are* $2r$*–almost independent at a fixed scale, then there exist constants* $C_r, \bar{C}_r, c_r > 0$ *such that for all sufficiently large* m,

$$\mathbb{P}(\mathscr{A}_m) \le C_r \left(2^{rm}\sigma(\rho_{2^m})^r + m^{-10}\right), \tag{2.3}$$

$$\mathbb{P}(\mathscr{D}_m) \ge c_r 2^{rm}\sigma(\rho_{2^{m+1}})^r - \bar{C}_r m^{-10}. \tag{2.4}$$

If $(\Omega^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$ *are* $2r$*–almost independent at all scales, then there exists a sequence* $\theta_m \to 0$ *such that if* $m' - m \ge b$ *(recall that* b *is a constant from* $(M3)_r$*) then for all sufficiently large* m

$$\mathbb{P}(\mathscr{D}_m \cap \mathscr{D}_{m'}) \le (\mathbb{P}(\mathscr{D}_m) + \bar{C}_r m^{-10})(\mathbb{P}(\mathscr{D}_{m'}) + \bar{C}_r m'^{-10})(1+\theta_m). \tag{2.5}$$

We start with some notations and a lemma. For $n \in \mathbb{N}^*$, for $k_1, \dots, k_r \le n$, define

$$A^{k_1,\dots,k_r}_{\rho_n} := \bigcap_{j=1}^{r} \Omega^{k_j}_{\rho_n}.$$

With these notations

$$\mathscr{A}_m = \bigcup_{0<k_1<k_2<\cdots<k_r\le 2^{m+1}} A^{k_1,\dots,k_r}_{\rho_{2^m}}, \tag{2.6}$$

$$\mathscr{D}_m = \bigcup_{(k_1,\dots,k_r)\in\mathscr{U}_m} A^{k_1,\dots,k_r}_{\rho_{2^{m+1}}}. \tag{2.7}$$

LEMMA 2.10. *Fix* $0 \le a_1 < a_2 \le 2$. *If* $(M1)_r$ *and* $(M2)_r$ *hold then there exist two sequences* $\delta_n \to 0$, $\eta_n \to 0$ *such that*

$$\sum_{a_1 n<k_1<k_2<\cdots<k_r\le a_2 n} \mathbb{P}(A^{k_1,\dots,k_r}_{\rho_n}) = \frac{((a_2-a_1)n\sigma(\rho_n))^r}{r!}(1+\delta_n) + \eta_n (\ln n)^{-10}. \tag{2.8}$$

For $a_2 - a_1 \ge \frac{1}{2}$, *there exists constant* c_r *such that*

$$\sum_{\substack{a_1 n<k_1<k_2<\cdots<k_r\le a_2 n \\ \widehat{\mathrm{Sep}}_n(k_1,\dots k_r)=r}} \mathbb{P}(A^{k_1,\dots,k_r}_{\rho_n}) \ge c_r (n\sigma(\rho_n))^r. \tag{2.9}$$

Proof. For $m \le r$, denote

$$S_m := \sum_{\substack{a_1 n<k_1<k_2<\cdots<k_r\le a_2 n \\ \mathrm{Sep}_n(k_1,\dots k_r)=m}} \mathbb{P}(A^{k_1,\dots,k_r}_{\rho_n}).$$

Note that S_r consists of $\frac{(a_2-a_1)^r n^r}{r!}(1+\delta'_n)$ terms for some sequence $\delta'_n \to 0$ as $n \to \infty$. Hence $(M1)_r$ yields

$$S_r = \frac{((a_2-a_1)n\sigma(\rho_n))^r}{r!}(1+\delta''_n). \tag{2.10}$$

where $\delta''_n \to 0$ as $n \to \infty$.

For $m < r$, S_m consists of $O(n^m(\mathfrak{s}(n))^{r-m})$ terms. Hence $(M2)_r$ gives

$$S_m \le Cn^m(\mathfrak{s}(n))^{r-m}\frac{K\sigma(\rho_n)^m}{(\ln n)^{100r}} = \eta_n(n\sigma(\rho_n))^m(\ln n)^{-10} \tag{2.11}$$

for some sequence $\eta_n \to 0$. Combining (2.10) with (2.11) we obtain (2.8). The proof of (2.9) is similar to that of (2.10), except that the number of terms is not anymore equivalent to $\frac{1}{r!}n^r(1+\delta'_n)$ but just larger than $\frac{1}{r!}(\frac{n}{2}-r\hat{\mathfrak{s}}(n))^r$ which is larger than $\frac{q^r}{2^r r!}n^r$, due to the hypothesis $\hat{\mathfrak{s}}(n) < n(1-q)/(2r)$. □

Proof of Proposition 2.9. First, (2.3) follows directly from (2.6) and (2.8). Next, define

$$I_m = \sum_{(k_1,\dots k_r)\in\mathscr{U}_m}\mathbb{P}(A^{k_1,\dots,k_r}_{\rho_{2^{m+1}}}),\quad J_m = \sum_{\substack{(k_1,\dots,k_r)\in\mathscr{U}_m\\(k'_1,\dots,k'_r)\in\mathscr{U}_m\\\{k_1,\dots,k_r\}\neq\{k'_1,\dots,k'_r\}}}\mathbb{P}\left(A^{k_1,\dots,k_r}_{\rho_{2^{m+1}}}\bigcap A^{k'_1,\dots,k'_r}_{\rho_{2^{m+1}}}\right).$$

From (2.7) and Bonferroni inequalities we get that

$$I_m - J_m \le \mathbb{P}(\mathscr{D}_m) \le I_m. \tag{2.12}$$

Now, (2.9) implies that

$$I_m \ge \bar{c}_r 2^{r(m+1)}\sigma(\rho_{2^{m+1}})^r. \tag{2.13}$$

On the other hand, since

$$A^{k_1,\dots,k_r}_{\rho_{2^{m+1}}}\bigcap A^{k'_1,\dots,k'_r}_{\rho_{2^{m+1}}} = A^{\{k_1,\dots,k_r\}\cup\{k'_1,\dots,k'_r\}}_{\rho_{2^{m+1}}},$$

we get that

$$J_m \le \tilde{C}_r\sum_{l=r+1}^{2r}\sum_{k_1<\dots<k_l}\mathbb{P}(A^{k_1,\dots,k_l}_{\rho_{2^{m+1}}}),$$

and (2.8) then implies that

$$J_m \le \overline{C}_r(2^{(r+1)m}\sigma(\rho_{2^{m+1}})^{r+1} + m^{-10}). \tag{2.14}$$

Combining (2.12), (2.13) and (2.14), and using the assumption (2.2) we obtain (2.4).

Finally, observe that

$$\mathbb{P}(\mathscr{D}_m\cap\mathscr{D}_{m'}) \le \sum_{\substack{(k_1,\dots,k_r)\in\mathscr{U}_m\\(l_1,\dots,l_r)\in\mathscr{U}_{m'}}}\mathbb{P}(A^{k_1,\dots,k_r}_{\rho_{2^{m+1}}}\cap A^{l_1,\dots,l_r}_{\rho_{2^{m'+1}}}).$$

But since $m' > m+1$ implies that $l_1 - k_r \ge \hat{\mathfrak{s}}(2^{m'+1})$, $(M3)_r$ then yields

$$\mathbb{P}(A^{k_1,\dots,k_r}_{\rho_{2^{m+1}}}\cap A^{l_1,\dots,l_r}_{\rho_{2^{m'+1}}}) \le \mathbb{P}(A^{k_1,\dots,k_r}_{\rho_{2^{m+1}}})\mathbb{P}(A^{l_1,\dots,l_r}_{\rho_{2^{m'+1}}})(1+\varepsilon_m),$$

so that using $(M1)_r$ and summing over all $(k_1,\dots,k_r)\in\mathscr{U}_m, (l_1,\dots,l_r)\in\mathscr{U}_{m'}$ we get that

$$\mathbb{P}(\mathscr{D}_m\cap\mathscr{D}_{m'}) \le I_m I_{m'}(1+\varepsilon_m).$$

Now (2.5) follows from (2.12), (2.13) and (2.14). □

2.3. **Convergence case. Proof of Theorem 2.6(a).** Suppose that $\mathbf{S}_r < \infty$. Then by monotonicity of $\sigma(\rho_n)$, we have that $n\sigma(\rho_n) \to 0$. By (2.3) of Proposition 2.9 we have that $\sum_m \mathbb{P}(\mathscr{A}_m) < \infty$. By Borel–Cantelli Lemma, with probability one, $\mathscr{A}_m$ happen only finitely many times. Observe that for $n \in (2^m, 2^{m+1}]$ we have $\{N^n_{\rho_n} \geq r\} \subset \mathscr{A}_m$ because $\Omega^k_{\rho_n} \subset \Omega^k_{\rho_{2^m}}$ for $n \geq 2^m$ due to (2.1). Hence with probability one $\{N^n_{\rho_n} \geq r\}$ happen only finitely many times. □

2.4. **Divergence case. Proof of Theorem 2.6(b).** Suppose that $\mathbf{S}_r = \infty$. We give a proof under the assumption (2.2). The case where (2.2) does not hold requires minimal modifications that will be explained at the end of this section.

CLAIM 2.11. Let $Z_n = \sum_{m=1}^{n} 1_{\mathscr{D}_m}$. There is a subsequence $\{Z_{n_k}\}$ such that a.s.

$$\frac{Z_{n_k}}{\mathbb{E}(Z_{n_k})} \to 1.$$

Since $\mathbb{E}(Z_n) \to \infty$, due to (2.4), the claim implies that, almost surely, $Z_n \to \infty$. That is, with probability one infinitely many of $\mathscr{D}_m$ happen. Noting that $\mathscr{D}_m \subset \{N^{2^{m+1}}_{\rho_{2^{m+1}}} \geq r\}$ completes the proof of Theorem 2.6(b) in the case when (2.2) holds.

Proof of Claim 2.11. We first prove that (2.4) and (2.5) imply that

$$\frac{Z_n}{\mathbb{E}(Z_n)} \to 1 \text{ in } L^2,$$

or equivalently that

$$\frac{\mathrm{Var}(Z_n)}{\mathbb{E}^2(Z_n)} \to 0. \tag{2.15}$$

Note that

$$\mathrm{Var}(Z_n) = \sum_{m=1}^{n} \mathbb{P}(\mathscr{D}_m) - \sum_{m=1}^{n} \mathbb{P}(\mathscr{D}_m)^2 + 2\sum_{i<j} \left[\mathbb{P}(\mathscr{D}_i \cap \mathscr{D}_j) - \mathbb{P}(\mathscr{D}_i)\mathbb{P}(\mathscr{D}_j)\right]. \tag{2.16}$$

By (2.5) for each δ there exists $m(\delta) > b$ such that if $i \geq m(\delta)$, $j - i \geq m(\delta)$ then

$$\begin{aligned} \mathbb{P}(\mathscr{D}_i \cap \mathscr{D}_j) - \mathbb{P}(\mathscr{D}_i)\mathbb{P}(\mathscr{D}_j) \leq \delta \mathbb{P}(\mathscr{D}_i)\mathbb{P}(\mathscr{D}_j) + 2\bar{C}_r i^{-10}\mathbb{P}(\mathscr{D}_j) \\ + 2\bar{C}_r j^{-10}\mathbb{P}(\mathscr{D}_i) + 2\bar{C}_r^2 (ij)^{-10}. \end{aligned} \tag{2.17}$$

Split (2.16) into two parts:

(a) Due to (2.17), the terms where $i \geq m(\delta)$, $j - i \geq m(\delta)$ contribute at most

$$\sum_{i \geq m(\delta), j-i \geq m(\delta)} \left[\delta \mathbb{P}(\mathscr{D}_i)\mathbb{P}(\mathscr{D}_j) + 2\bar{C}_r i^{-10}\mathbb{P}(\mathscr{D}_j) + 2\bar{C}_r j^{-10}\mathbb{P}(\mathscr{D}_i) + 2\bar{C}_r^2 (ij)^{-10}\right]$$
$$\leq \delta(\mathbb{E}(Z_n))^2 + 8\bar{C}_r \mathbb{E}(Z_n) + 8\bar{C}_r^2.$$

(b) The terms where $i \leq m(\delta)$ or $j - i \leq m(\delta)$ contribute at most

$$[2m(\delta) + 1] \sum_{j=1}^{n} \mathbb{P}(\mathscr{D}_j) = [2m(\delta) + 1]\mathbb{E}(Z_n).$$

Since $\mathbb{E}(Z_n) \to \infty$, the case (a) dominates for large n giving

$$\limsup_{n\to\infty} \frac{\mathrm{Var}(Z_n)}{(\mathbb{E}(Z_n))^2} \le \delta.$$

Since δ is arbitrary, (2.15) follows.

Let $n_k = \inf\{n : (\mathbb{E}(Z_n))^2 \ge k^2 \,\mathrm{Var}(Z_n)\}$. Then by Chebyshev inequality

$$\mathbb{P}\left(|Z_{n_k} - \mathbb{E}(Z_{n_k})| > \delta\mathbb{E}(Z_{n_k})\right) \le \frac{1}{\delta^2 k^2}.$$

Thus

$$\sum_{k=1}^{\infty} \mathbb{P}(|Z_{n_k} - E(Z_{n_k})| > \delta\mathbb{E}(Z_{n_k})) \le \sum_{k=1}^{\infty} \frac{1}{\delta^2 k^2} < \infty.$$

Therefore, by Borel–Cantelli Lemma, with probability 1, for large k,

$$|Z_{n_k} - \mathbb{E}(Z_{n_k})| < \delta\mathbb{E}(Z_{n_k}).$$

Hence $\frac{Z_{n_k}}{\mathbb{E}(Z_{n_k})} \to 1$ *a.s.*, as claimed. □

It remains to consider the case where (2.2) fails. After passing to a subsequence, we choose a decreasing sequence ν_n such that $\tilde{\sigma}(\rho_n) := \nu_n\sigma(\rho_n)$ satisfies $\lim_{n\to\infty} n\tilde{\sigma}(\rho_n) = 0$ and $\sum_{j=1}^{\infty} (2^j\tilde{\sigma}(\rho_{2^j}))^r = \infty$.

Next, we define for each $n \in \mathbb{N}$ and for each $k \le n$, a sequence of events $\{\tilde{\Omega}^k_{\rho_n}\}_{k\le n}$ as follows: If $\Omega^k_{\rho_n}$ does not occur then $\tilde{\Omega}^k_{\rho_n}$ does not occur and, conditionally on $\Omega^k_{\rho_n}$ occurring, $\tilde{\Omega}^k_{\rho_n}$ occurs with probability ν_n independently of all other events (all other $\Omega^k_{\rho_n}$ with different k or different n).

The events $(\tilde{\Omega}^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$ thus satisfy $(M1)_r$, $(M2)_r$, and $(M3)_r$ the same way as the events $(\Omega^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$, with this difference that $\sigma(\rho_n)$ is now replaced with $\tilde{\sigma}(\rho_n)$.[1] Since condition (2.2) is satisfied by $\tilde{\sigma}(\rho_n)$, and since

$$\sum_{j=1}^{\infty} (2^j\tilde{\sigma}(\rho_{2^j}))^r = \infty,$$

we get that, with probability one, more than r events among the events $\{\tilde{\Omega}^k_{\rho_n}\}_{k\le n}$ occurs for infinitely many n. By definition, this implies that with probability one, more than r events among the events $\{\Omega^k_{\rho_n}\}_{k\le n}$ occurs for infinitely many n. The proof of Theorem 2.6(b) is thus completed. □

2.5. **Prescribing some details.** In the remaining part of Section 2 we describe some extensions of Theorem 2.6(b).

Namely, we assume that $\Omega^n_\rho = \bigcup_{i=1}^{p} \Omega^{n,i}_\rho$ and that there exists a constant $\hat{\varepsilon} > 0$ such that for each i, $\mathbb{P}(\Omega^{n,i}_\rho) \ge \hat{\varepsilon}\mathbb{P}(\Omega^n_\rho)$. We also assume the following extension of $(M1)_r$: for each $(k_1,\dots,k_r)$ with $\mathrm{Sep}_n(k_1,\dots,k_r) = r$ and each $(i_1,\dots,i_r) \in$

[1]Note that the events $\{\tilde{\Omega}^k_{\rho_n}\}$ will not satisfy (2.1) even if the events $\{\Omega^k_{\rho_n}\}$ satisfy it, but in this part of the proof of Theorem 2.6 (b) condition (2.1) is not needed.

$\{1,\dots,p\}^r$,

$$\widetilde{(M1)}_r \quad \left[\prod_{j=1}^{r} \mathbb{P}(\Omega_{\rho_n}^{k_j,i_j})\right](1-\varepsilon_n) \le \mathbb{P}\left(\bigcap_{j=1}^{r} \Omega_{\rho_n}^{k_j,i_j}\right) \le \left[\prod_{j=1}^{r} \mathbb{P}(\Omega_{\rho_n}^{k_j,i_j})\right](1+\varepsilon_n);$$

and the following extension of $(M3)_r$: for each δ there is $b = b(\delta)$ such that letting $\hat{s}(n) = \delta n$ we have that for each $(k_1,\dots,k_r)$, $(l_1,\dots,l_r)$ with

$$\widehat{\mathrm{Sep}}_{2^{i+1}}(k_1,\dots,k_r) = r, \quad \widehat{\mathrm{Sep}}_{2^{j+1}}(l_1,\dots,l_r) = r, \quad l_1 - k_r \ge \hat{\mathfrak{s}}(2^{j+1}), \quad j - i \ge b$$

and for each $(i_1, i_2, \dots, i_r)$, $(j_1, j_2 \dots j_r) \in \{1,\dots,p\}^r$,

$$\widetilde{(M3)}_r \quad \mathbb{P}\left(\left[\bigcap_{\alpha=1}^{r} \Omega_{\rho_{2^i}}^{k_\alpha,i_\alpha}\right] \bigcap \left[\bigcap_{\beta=1}^{r} \Omega_{\rho_{2^j}}^{l_\beta,j_\beta}\right]\right) \le \left[\prod_{\alpha=1}^{r} \mathbb{P}(\Omega_{\rho_{2^i}}^{k_\alpha,i_\alpha})\right]\left[\prod_{\beta=1}^{r} \mathbb{P}(\Omega_{\rho_{2^j}}^{l_\beta,j_\beta})\right](1+\varepsilon_i).$$

THEOREM 2.12. [2] *If $\mathbf{S}_r = \infty$, and $\widetilde{(M1)}_k$, $(M2)_k$ as well as $\widetilde{(M3)}_k$ for $k = 1,\dots,2r$ are satisfied, then for any $i_1, i_2 \dots i_r$ and for any intervals $I_1, I_2 \dots I_r \subset [0,1]$, with probability 1 there are infinitely many n such that for some $k_1(n), k_2(n) \dots k_r(n)$ with $\frac{k_j(n)}{n} \in I_j$, $\Omega_{\rho_n}^{k_j,i_j}$ occur.*

The proof of Theorem 2.12 is similar to the proof of Theorem 2.6(b). Without the loss of generality we may assume that I_j does not contain 0. Then we fix a large constant l and consider the following modification of $\mathscr{D}_m$

$$\begin{aligned}\tilde{\mathscr{D}}_m := \Big\{ &\exists 2^{lm} < k_1 < \dots < k_r \le 2^{l(m+1)} \text{ such that } \frac{k_\alpha}{2^{l(m+1)}} \in I_\alpha, \\ &\Omega_{\rho_{2^{l(m+1)}}}^{k_\alpha,i_\alpha} \text{ happens and } k_{\alpha+1} - k_\alpha \ge \hat{\mathfrak{s}}(2^{l(m+1)}), 0 \le \alpha \le r-1 \Big\}.\end{aligned}$$

Arguing as in Proposition 2.9 we conclude that $\tilde{\mathscr{D}}_{m_1}$ and $\tilde{\mathscr{D}}_{m_2}$ are asymptotically independent (in the sense of (2.5)) if $m_2 > m_1 + p$ and p is so large that $2^{-p} \notin I_\alpha$ for $\alpha = 1, 2 \dots r$. The rest of the proof is identical to the proof of Theorem 2.6(b).

2.6. **Poisson regime.**

THEOREM 2.13. *Suppose $(M1)_r$ and $(M2)_r$ hold for all r and $\lim_{n\to\infty} n\sigma(\rho_n) = \lambda$. Then $N_{\rho_n}^n$ converges in law as $n \to \infty$ to the Poisson distribution with parameter λ.*

Proof. We compute all (factorial) moments of the limiting distribution. Let $\mathscr{X}$ denote the Poisson random variable with parameter λ. Below $\binom{m}{r}$ denotes the binomial coefficient $\dfrac{m!}{r!(m-r)!}$. Since (see, e.g., [139, formula (3.4) in Section 7.3])

$$\mathbb{E}\left(\binom{N_{\rho_n}^n}{r}\right) = \sum_{k_1<k_2<\dots<k_r\le n} \mathbb{P}(A_{\rho_n}^{k_1,\dots,k_r}),$$

[2] This result is not used in the present paper, so it can be skipped during the first reading. In a followup work, we shall use Theorem 2.12 to obtain some analogues of the Functional Law of Iterated Logarithm for heavy tailed random variables.

Lemma 2.10 implies for each r

$$\lim_{n\to\infty} \mathbb{E}\left(\binom{N^n_{\rho_n}}{r}\right) = \frac{\lambda^r}{r!} = \mathbb{E}\left(\binom{\mathscr{X}}{r}\right). \tag{2.18}$$

Since this holds for all r we also have that for all r, $\lim\limits_{n\to\infty} \mathbb{E}((N^n_{\rho_n})^r) = \mathbb{E}(\mathscr{X}^r)$. Since the Poisson distribution is uniquely determined by its moments the result follows. □

Similarly to Borel–Cantelli Lemma, we also have the following extension of Theorem 2.13 in the setting of §2.5. Denote $N^{n,i}_I$ the number of times event $\Omega^{k,i}_{\rho_n}$ occurs with $k/n \in I$. Write $N^{n,i} := N^{n,i}_{[0,1]}$.

THEOREM 2.14. *Suppose that $\widetilde{(M1)}_r$ and $(M2)_r$ hold for all r and that*

$$\lim_{n\to\infty} n\mathbb{P}(\Omega^{n,i}_{\rho_n}) = \lambda_i.$$

Then $\{N^{n,i}_{\rho_n}\}_{i=1}^p$ converge in law as $n \to \infty$ to the independent Poisson random variables with parameter λ_i.

Moreover, if $I_1, I_2, \dots I_s$ are disjoint intervals, then $\{N^{n,i}_{I_j}\}$, $i = 1\dots p$, $j = 1\dots s$ converge in law as $n \to \infty$ to the independent Poisson random variables with parameter $\lambda_i|I_j|$.

Proof. It suffices to prove the second statement. The proof is similar to the proof of Theorem 2.13. Namely, similarly to (2.18), we show that for each set $r_{ij} \in \mathbb{N}$ we have

$$\lim_{n\to\infty} \mathbb{E}\left(\prod_{i,j} \binom{N^{n,i}_{I_j}}{r_{ij}}\right) = \prod_{i,j} \frac{(\lambda_i|I_j|)^{r_{ij}}}{(r_{ij})!} = \prod_{i,j} \mathbb{E}\left(\binom{\mathscr{X}_{ij}}{r_{ij}}\right),$$

where $\mathscr{X}_{ij}$ are independent Poisson random variables with parameters $\lambda_i|I_j|$. □

2.7. **Notes.** The usual Borel–Cantelli Lemma is a classical subject in probability. There are many extensions to weakly dependent random variables, see, e.g., [157, §12.15], [152, §1]. The connection between Borel–Cantelli Lemma and Poisson Limit Theorem is discussed in [55, 62]. The multiple Borel–Cantelli Lemma for independent events is proven in [130]. [1] obtains multiple Borel–Cantelli Lemma for systems admitting good symbolic dynamics. Extending multiple Borel–Cantelli Lemma to more general systems allows to obtain many new applications, see Sections 4–10 of this paper. Separation conditions similar to our have been used in [44, 146] to obtain the Poisson Law.

3. MULTIPLE BOREL–CANTELLI LEMMA FOR EXPONENTIALLY MIXING DYNAMICAL SYSTEMS

3.1. **Good maps, good targets.** Let f be a transformation of a metric space X preserving a measure μ. Given a family of sets $\Omega_\rho \subset X$, $\rho \in \mathbb{R}^*_+$, we will, in a slight abuse of notations, sometimes call Ω_ρ the event 1_{Ω_ρ} and Ω^k_ρ the event $1_{\Omega_\rho} \circ f^k$. We will take $\sigma(\rho) = \mu(\Omega_\rho)$.

To deal with multiple recurrence and not just multiple hitting of targets, we need to consider slightly more complicated events.

Given a family of events $\bar{\Omega}_\rho$ in $X \times X$, let $\bar{\Omega}^k_\rho \subset X$ be the event

$$\bar{\Omega}^k_\rho = \{x : (x, f^k x) \in \bar{\Omega}_\rho\}.$$

We will take $\bar{\sigma}(\rho) = (\mu \times \mu)(\bar{\Omega}_\rho)$.

From now on we will always assume that if $\rho' \le \rho$, then

$$\Omega_{\rho'} \subset \Omega_\rho, \quad \bar{\Omega}_{\rho'} \subset \bar{\Omega}_\rho.$$

For $\phi : X^k \to \mathbb{R}$, $k \in \mathbb{N}^+$, we denote

$$\mu^k(\phi) = \int_{X^k} \phi(x_1, \cdots, x_k) \mathrm{d}\mu(x_1) \cdots \mathrm{d}\mu(x_k). \tag{3.1}$$

Given a sequence $\{\rho_n\}$, we recall that $N^n_{\rho_n}$ denotes the number of times $k \le n$ such that $\Omega^k_{\rho_n}$ (or $\bar{\Omega}^k_{\rho_n}$) occurs. We want to give conditions on the system (f, X, μ) and on the family $(\Omega^k_{\rho_n})_{(n,k)\in\mathbb{N}^2; 1\le k\le 2n}$ or $(\bar{\Omega}^k_{\rho_n})_{(n,k)\in\mathbb{N}^2; 1\le k\le 2n}$, that imply the validity of the dichotomy of Theorem 2.6 for the number of hits $N^n_{\rho_n}$. For this, we take

$$\mathbf{S}_r = \sum_{j=1}^{\infty} \left(2^j \mathbf{v}_j\right)^r$$

where $\mathbf{v}_j = \sigma(\rho_{2^j})$ if we are considering targets of the type Ω^k_ρ and $\mathbf{v}_j = \bar{\sigma}(\rho_{2^j})$ if we are considering targets of the type $\bar{\Omega}^k_\rho$.

The independence conditions $(M1)_r$, $(M2)_r$, $(M3)_r$ will be satisfied due to mixing conditions on the dynamical system (f, X, μ), and to some regularity and shrinking conditions on the targets that we now state.

DEFINITION 3.1 ($(r+1)$-fold exponentially mixing systems for $r \ge 1$)**.** Let $\mathbb{B}$ be a space of real valued functions defined over X^{r+1}, with a norm $\|\cdot\|_{\mathbb{B}}$. For $r \ge 1$, we say that $(f, X, \mu, \mathbb{B})$ is $(r+1)$-fold exponentially mixing, if there exist constants $C > 0, L > 0$ and $\theta < 1$ such that $\forall\, A$, A_1, $A_2 \in \mathbb{B}$,

(Prod) $\|A_1 A_2\|_{\mathbb{B}} \le C \|A_1\|_{\mathbb{B}} \|A_2\|_{\mathbb{B}}$,

(Gr) $\|A \circ (f^{k_0}, \ldots, f^{k_r})\|_{\mathbb{B}} \le C L^{\sum_{i=0}^r k_i} \|A\|_{\mathbb{B}}$,

(EM)$_r$ If $0 = k_0 \le k_1 \le \ldots \le k_r$ are such that $\forall\, j \in [0, r-1]$, $k_{j+1} - k_j \ge m$, then

$$\left| \int_X A(x, f^{k_1} x, \cdots, f^{k_r} x) \mathrm{d}\mu(x) - \mu^{r+1}(A) \right| \le C\theta^m \|A\|_{\mathbb{B}}.$$

Given a system $(f, X, \mu, \mathbb{B})$, we now define the notion of simple admissible targets.

DEFINITION 3.2 (Simple admissible targets)**.** Let $\{\Omega_\rho\}$, $\rho \in \mathbb{R}^*_+$, be a decreasing collection of sets in X for which there are positive η, τ such that for all sufficiently small $\rho > 0$:

(Appr) There are functions $A^-_\rho, A^+_\rho : X \to \mathbb{R}$ such that $A^\pm_\rho \in \mathbb{B}$ and

(i) $\|A^{\pm}_{\rho}\|_{\infty} \leq 2$ and $\|A^{\pm}_{\rho}\|_{\mathbb{B}} \leq \rho^{-\tau}$;
(ii) $A^{-}_{\rho} \leq 1_{\Omega_{\rho}} \leq A^{+}_{\rho}$;
(iii) $\mu(A^{+}_{\rho}) - \mu(A^{-}_{\rho}) \leq \sigma(\rho)^{1+\eta}$, where $\sigma(\rho) = \mu(\Omega_{\rho})$.

Let (ρ_n) be a decreasing sequence of positive numbers. We say that the sequence (Ω_{ρ_n}) is a simple admissible sequence of targets for $(f, X, \mu, \mathbb{B})$ if there exists $u > 0$ such that

$$\text{(Poly)} \qquad \exists n_0 : \forall n \geq n_0 : \quad \rho_n \geq n^{-u}, \quad \sigma(\rho_n) \geq n^{-u},$$

and

$$\text{(Mov)} \qquad \begin{gathered} \forall R, L\, \exists \bar{C}, n_0\; \forall n \geq n_0 : \forall k \in (0, R \ln n), \\ \mu(\Omega_{\rho_n} \cap f^{-k}\Omega_{\rho_n}) \leq \bar{C}\sigma(\rho_n)(\ln n)^{-L}. \end{gathered}$$

REMARK 3.3. Note that properties (Appr)(ii) and (iii) imply that

$$\mu(A^{+}_{\rho}) - \mu(\Omega_{\rho}) \leq \mu(\Omega_{\rho})^{1+\eta}, \quad \mu(\Omega_{\rho}) - \mu(A^{-}_{\rho}) \leq \mu(\Omega_{\rho})^{1+\eta}.$$

A useful situation where one can verify these properties is the following.

LEMMA 3.4. *Suppose that f is Lipschitz and $\mathbb{B}$ is the space of Lipschitz functions. We have that (Prod) and (Gr) hold with L being the Lipschitz constant of f. Moreover, if there exist constants $\xi, \xi' > 0$ and $\Phi : X \to \mathbb{R}$ a (uniformly) Lipschitz function*[3] *such that for any interval $J \in \mathbb{R}$,*

$$\mu(\{x : \Phi(x) \in J\}) \in [|J|^{\xi}, |J|^{\xi'}]$$

and two (uniformly) Lipschitz functions a_1 and $a_2 : \mathbb{R} \to \mathbb{R}$ such that for some $\alpha, \alpha' > 0$ we have

$$a_2(\rho) - a_1(\rho) \in [\rho^{\alpha}, \rho^{\alpha'}]$$

then (Appr) *holds for the targets*

$$\Omega_{\rho} = \left\{\Phi(x) \in [a_1(\rho), a_2(\rho)]\right\}.$$

The same result holds if $\mathbb{B}$ is the space of C^s functions or the space of compactly supported C^s functions with $s > 0$ arbitrary.

The proof of Lemma 3.4 relies on simple approximation of characteristic functions by Lipschitz functions.

Proof. We will construct A^{+}_{ρ} that satisfies (i), (ii) of (Appr) and
$\overline{\text{(iii)}}$ $\mu(A^{+}_{\rho}) - \mu(\Omega_{\rho}) \leq \sigma(\rho)^{1+\eta}$.

The construction of A^{-}_{ρ} is similar. Note that $\sigma(\rho) = \mu(\Omega_{\rho}) \in [\rho^{\alpha\xi}, \rho^{\alpha'\xi'}]$.

Define a family of smooth function $\psi^{+} : \mathbb{R}^4 \to [0, 2]$ such that for $v > u$ and $\varepsilon > 0$ and $x \in \mathbb{R}$ (we are not interested in the form of ψ^{+} outside this domain) we have

$$\psi^{+}(u, v, \varepsilon, x) = \begin{cases} 1, & \text{for } x \in [u, v] \\ 0, & \text{for } x \notin [u - \varepsilon(v-u), v + \varepsilon(v-u)] \end{cases}$$

[3] A typical situation for using Lemma 3.4 will be with $\Phi(x)$ defined by some distance $d(x_0, x)$.

and for which there exist constants $\eta > 0$ and $C > 0$ such that that for any ν_0 and for $\mathbb{R}^4 \supset \mathscr{R}_{\nu_0} := \{v - u \ge \nu_0, \varepsilon \ge \nu_0\}$, we have that

$$\|\psi^+\|_{C^1(\mathscr{R}_{\nu_0})} \le C\nu_0^{-\eta},$$

where $C^1(\mathscr{R}_{\nu_0})$ refers to the C^1 norm in the region $\mathscr{R}_{\nu_0}$.

Define now $A_\rho^+ : X \to \mathbb{R} : x \mapsto \psi^+(a_1(\rho), a_2(\rho), \rho^b, \Phi(x))$, where $b > 1$ will be chosen later. It is clear that A_ρ^+ is Lipschitz and that $1_{\Omega_\rho} \le A_\rho^+$. On the other hand, $\|A_\rho^+\|_\infty \le 2$ and $\|A_\rho^+\|_{\mathbb{B}} \le C(\Phi)\rho^{-b\alpha\eta}$, and (i) holds for $\tau = b\alpha\eta + 1$. We turn now to $\overline{\text{(iii)}}$. We observe that with $J_1 = [a_1(\rho) - \rho^b(a_2(\rho) - a_1(\rho)), a_1(\rho)]$ and $J_2 = [a_2(\rho), a_2(\rho) + \rho^b(a_2(\rho) - a_1(\rho))]$,

$$\mu(A_\rho^+) - \mu(\Omega_\rho) \le 2\mu(\{\Phi(x) \in J_1 \cup J_2\}) \le 4\rho^{\xi'(b+\alpha')}.$$

Hence, if b is chosen sufficiently large we have $\rho > 0$ sufficiently small that $\mu(A_\rho^+) - \mu(\Omega_\rho) \le \sigma(\rho)^2$.

The fact that the same results hold if $\mathbb{B}$ is the space of C^s functions or the space of compactly supported C^s functions with $s > 0$ arbitrary, is a simple consequence of the approximation of Lipschitz functions by smooth functions. □

To deal with recurrence, the following definition is useful.

DEFINITION 3.5 (Composite admissible targets)**.** Let $\{\overline{\Omega}_\rho\}, \rho \in \mathbb{R}_+^*$ be a decreasing collection of sets in $X \times X$ satisfying the following conditions for some positive constants $\overline{C}, \eta, \tau$ and for all sufficiently small $\rho > 0$,

$\overline{\text{(Appr)}}$ There are functions $\overline{A}_\rho^-, \overline{A}_\rho^+ : X \times X \to \mathbb{R}$ such that $\overline{A}_\rho^\pm \in \mathbb{B}$ and

- (i) $\|\overline{A}_\rho^\pm\|_\infty \le 2$ and $\|\overline{A}_\rho^\pm\|_{\mathbb{B}} \le \rho^{-\tau}$;
- (ii) $\overline{A}_\rho^- \le 1_{\overline{\Omega}_\rho} \le \overline{A}_\rho^+$;
- (iii) For any fixed x,

$$\bar{\sigma}(\rho) - \bar{\sigma}(\rho)^{1+\eta} \le \int \overline{A}_\rho^-(x,y)\mathrm{d}\mu(y) \le \int \overline{A}_\rho^+(x,y)\mathrm{d}\mu(y) \le \bar{\sigma}(\rho) + \bar{\sigma}(\rho)^{1+\eta},$$

- (iv) For any fixed y, $\int \overline{A}_\rho^+(x,y)\mathrm{d}\mu(x) \le \overline{C}\bar{\sigma}(\rho)$.

For a decreasing sequence of positive numbers (ρ_n), the sequence $(\overline{\Omega}_{\rho_n})$ is said to be composite admissible if there exists $u > 0$ such that

$$\overline{\text{(Poly)}} \qquad \exists n_0 \,\forall n \ge n_0 : \quad \rho_n \ge n^{-u}, \quad \bar{\sigma}(\rho_n) \ge n^{-u},$$

and there is a constant $a > 0$ such that for any $k_1 < k_2$

$$\overline{\text{(Sub)}} \qquad \overline{\Omega}_\rho^{k_1} \cap \overline{\Omega}_\rho^{k_2} \subset f^{-k_1}\overline{\Omega}_{a\rho}^{k_2-k_1},$$

and

$$\overline{\text{(Mov)}} \qquad \forall L \exists n_0 : \forall n \ge n_0 \quad \forall k \ne 0, \quad \mu(\overline{\Omega}_{a\rho_n}^k) \le \overline{C}(\ln n)^{-L}.$$

Observe that integrating condition $\overline{\text{(Appr)}}$(iii) with respect to x we obtain for each $n \neq 0$,

$$\bar{C}^{-1}\mu\left(\bar{\Omega}^n_\rho\right) \le \mu\left(\bar{A}^-_\rho(x, f^n x)\right) \le \mu\left(\bar{A}^+_\rho(x, f^n x)\right) \le \bar{C}\mu\left(\bar{\Omega}^n_\rho\right). \tag{3.2}$$

Typical composite targets we deal with are of the type $d(x, y) < \rho$ or $d(x, y) < \gamma(x)\rho$, where $\gamma(x)$ is related to the density of μ at the point x. We state here a general Lemma that guarantees the admissibility of such targets. The statement is a bit technical but if we keep in mind that the function $\Phi(x, y)$ is usually defined by a distance, then the hypothesis of the Lemma become natural. The proof of the Lemma is very simple and follows a similar scheme of the proof of Lemma 3.4 for simple targets.

LEMMA 3.6. *Suppose that f is Lipschitz and $\mathbb{B}$ is the space of Lipschitz functions. Suppose there exist constants $C, \xi, \xi', \xi'' > 0$ and $\Phi : X \times X \to \mathbb{R}$ a (uniformly) Lipschitz function such that*

(h1) $\forall (x, y) \in X \times X, \quad \Phi(x, y) \le C\Phi(y, x).$

(h2) For any interval $J \in \mathbb{R}$,

$$\bar{\sigma}(J) := (\mu \times \mu)\left(\{(x, y) \in X \times X : \Phi(x, y) \in J\}\right) \in [|J|^{\xi}, |J|^{\xi'}].$$

(h3) For any $x \in X$, $\mu\left(\left\{y \in X : \Phi(x, y) \in J\right\}\right) = \bar{\sigma}(J)(1 + \mathcal{O}(|J|^{\xi''}))$.

If two (uniformly) Lipschitz functions a_1 and $a_2 : \mathbb{R} \to \mathbb{R}$ are such that for some $\alpha, \alpha' > 0$

$$a_2(\rho) - a_1(\rho) \in [\rho^{\alpha}, \rho^{\alpha'}]$$

then $\overline{\text{(Appr)}}$ holds for the targets

$$\bar{\Omega}_\rho = \{\Phi(x, y) \in [a_1(\rho), a_2(\rho)]\}$$

The same result holds if $\mathbb{B}$ is the space of C^s functions or the space of compactly supported C^s functions with $s > 0$ arbitrary.

Proof. The proof is very similar to that of Lemma 3.4. We just explain the differences. Note that $\bar{\sigma}(\rho) = (\mu \times \mu)(\bar{\Omega}_\rho) \in [\rho^{\alpha\xi}, \rho^{\alpha'\xi'}]$.

We introduce $\bar{A}^+_\rho : X \times X \to \mathbb{R} : (x, y) \mapsto \psi^+(a_1(\rho), a_2(\rho), \rho^b, \Phi(x, y))$, where ψ^+ is as in the proof of Lemma 3.4. Properties (i) and (ii) hold as in the proof of Lemma 3.4.

We turn now to (iii). We fix $x \in X$, and observe that with $I = [a_1(\rho), a_2(\rho)]$ and $J_1 = [a_1(\rho) - \rho^b(a_2(\rho) - a_1(\rho)), a_1(\rho)]$, $J_2 = [a_2(\rho), a_2(\rho) + \rho^b(a_2(\rho) - a_1(\rho))]$ we have that

$$\begin{aligned}\int \bar{A}^+_\rho(x, y)\mathrm{d}\mu(y) - \mu\left(\left\{y \in X : \Phi(x, y) \in I\right\}\right) &\le 2\mu\left(\left\{y \in X : \Phi(x, y) \in J_1 \cup J_2\right\}\right)\\ &\le \bar{\sigma}(\rho)^2\end{aligned}$$

if b is sufficiently large due to $(h2)$ and $(h3)$. Applying $(h2)$ and $(h3)$, we also see that

$$|\mu\left(\left\{y \in X : \Phi(x, y) \in I\right\}\right) - \bar{\sigma}(\rho)| = \mathcal{O}(\bar{\sigma}(\rho)^{1+\eta})$$

for some $\eta > 0$. This proves $\overline{\text{(Appr)}}$(iii).

Finally, fix $y \in X$ and observe that $(h1)$ implies

$$\int \bar{A}^+_\rho(x,y)\mathrm{d}\mu(x) \le C\int \bar{A}^+_\rho(y,x)\mathrm{d}\mu(x) \le 2C\bar{\sigma}(\rho),$$

which proves $\overline{\text{(Appr)}}$(iv). □

3.2. **Multiple Borel–Cantelli Lemma for admissible targets.** The goal of this section is to establish the following Theorem that gives conditions on the system (f,X,μ) and on the family $(\Omega^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$ (or $(\bar{\Omega}^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$), that imply the validity of the dichotomy of Theorem 2.6 for the number of hits $N^n_{\rho_n}$. Recall that

$$\mathbf{S}_r = \sum_{j=1}^{\infty}\left(2^j\mathbf{v}_j\right)^r$$

where $\mathbf{v}_j = \sigma(\rho_{2^j})$ if we are considering targets of the type Ω^k_ρ and $\mathbf{v}_j = \bar{\sigma}(\rho_{2^j})$ if we are considering targets of the type $\bar{\Omega}^k_\rho$.

THEOREM 3.7. *Assume a system $(f,X,\mu,\mathbb{B})$ is $(2r+1)$-fold exponentially mixing.*[4] *Then*

(a) *If (Ω_{ρ_n}) is a sequence of simple admissible targets as in Definition 3.2, then the events of the family $(\Omega^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$ are $2r$–almost independent at all scales.*
(b) *If $(\bar{\Omega}_{\rho_n})$ is a sequence of composite admissible targets as in Definition 3.5, then the events of the family $(\bar{\Omega}^k_{\rho_n})_{(n,k)\in\mathbb{N}^2;1\le k\le 2n}$ are $2r$–almost independent at all scales.*

Hence, Theorem 2.6 implies

COROLLARY 3.8. *If the system $(f,X,\mu,\mathbb{B})$ is $(2r+1)$-fold exponentially mixing, and if (Ω_{ρ_n}) (or $(\bar{\Omega}_{\rho_n})$) are as in Definition 3.2 (or Definition 3.5), then*

(a) *If $\mathbf{S}_r < \infty$, then with probability 1, we have that for large n $N^n_{\rho_n} < r$.*
(b) *If $\mathbf{S}_r = \infty$, then with probability 1, there are infinitely many n such that $N^n_{\rho_n} \ge r$.*

Theorem 3.7 is a direct consequence of Proposition 3.9 below. We accept a convention that $(\mathrm{EM})_k$ for $k \le 0$ is always satisfied.

PROPOSITION 3.9. *Given a dynamical system $(f,X,\mu,\mathbb{B})$ and a sequence of decreasing sets (Ω_{ρ_n}) such that* (Prod), (Poly), *and* (Appr) *hold, then with the function $\sigma(\cdot) := \mu(\Omega_\cdot)$, and*

(i) *If $(\mathrm{EM})_{r-1}$ holds, then $(M1)_r$ is satisfied with the function $\mathfrak{s}:\mathbb{N}\to\mathbb{N}$, $\mathfrak{s}(n) = R\ln n$, where R is sufficiently large (depending on r, the system and the targets).*
(ii) *If* (Gr), (Mov) *and $(\mathrm{EM})_{r-2}$ hold, then $(M2)_r$ is satisfied.*

[4] Part (a) holds for $2r$-fold exponentially mixing systems, as shown by the first part of Proposition 3.9.

(iii) *If* (Gr) *and* $(\mathrm{EM})_r$ *hold, then for arbitrary* $\varepsilon > 0$, $(M3)_r$ *is satisfied with* $\hat{\mathfrak{s}}(n) = \varepsilon n$.

Similarly, given a dynamical system $(f, X, \mu, \mathbb{B})$ *and a sequence of decreasing sets* $\{\overline{\Omega}_{\rho_n}\}$ *such that* (Prod), $(\overline{\mathrm{Poly}})$ *and* $(\overline{\mathrm{Appr}})$ *hold, then, with the function* $\overline{\sigma}(\cdot) := \mu \times \mu(\overline{\Omega}_{\cdot})$*:*

(i) *If* $(\mathrm{EM})_r$ *holds, then* $(M1)_r$ *is satisfied with the function* $\mathfrak{s} : \mathbb{N} \to \mathbb{N}$, $\mathfrak{s}(n) = R \ln n$*, with R sufficiently large (depending on r, the system and the targets).*

(ii) *If* (Gr), $(\overline{\mathrm{Mov}})$, $(\overline{\mathrm{Sub}})$ *and* $(\mathrm{EM})_{r-1}$ *hold, then* $(M2)_r$ *is satisfied.*

(iii) *If* (Gr) *and* $(\mathrm{EM})_r$ *hold, then for arbitrary* $\varepsilon > 0$ $(M3)_r$ *is satisfied with* $\hat{\mathfrak{s}}(n) = \varepsilon n$.

Proof of Proposition 3.9. We use C to denote a constant that may change from line to line but that will not depend on ρ_n, Ω_{ρ_n}, $\overline{\Omega}_{\rho_n}$, the order of iteration of f, etc.

We note that it is sufficient to check (M1)–(M3) for all sufficiently large n since smaller n can be handled by increasing K and ε_n.

Proof of (i). For Ω_{ρ_n}, we prove $(M1)_r$ in case $k_{i+1} - k_i \geq \sqrt{R} \ln n$, where R is a sufficiently large constant. Indeed, using (Appr) and $(\mathrm{EM})_{r-1}$ we get

$$\mu\left(\prod_{i=1}^{r} 1_{\Omega_{\rho_n}}(f^{k_i} x)\right) \leq \mu\left(\prod_{i=1}^{r} A^+_{\rho_n}(f^{k_i} x)\right) \leq \prod_{i=1}^{r} \mu\left(A^+_{\rho_n}\right) + C\rho_n^{-r\tau} \theta^{\sqrt{R}\ln n}$$
$$\leq \left(\mu(\Omega_{\rho_n}) + C\mu(\Omega_{\rho_n})^{1+\eta}\right)^r + C\rho_n^{-r\tau} \theta^{\sqrt{R}\ln n},$$

which yields the RHS of $(M1)_r$, due to (Poly) if R is sufficiently large. The LHS is proved similarly.

For $\overline{\Omega}_{\rho_n}$, we approximate $1_{\overline{\Omega}_{\rho_n}}$ by $\overline{A}^{\pm}_{\rho_n}$, apply $(\overline{\mathrm{Appr}})$, $(\mathrm{EM})_r$ to the functions

$$B^+_{\rho_n}(x_0, \cdots, x_r) = \overline{A}^+_{\rho_n}(x_0, x_1) \cdots \overline{A}^+_{\rho_n}(x_0, x_r),$$
$$B^-_{\rho_n}(x_0, \cdots, x_r) = \overline{A}^-_{\rho_n}(x_0, x_1) \cdots \overline{A}^-_{\rho_n}(x_0, x_r),$$

and get

$$\mu\left(\bigcap_{j=1}^{r} \overline{\Omega}^{k_j}_{\rho_n}\right) \leq \left(\overline{\sigma}(\rho_n) + C\overline{\sigma}(\rho_n)^{1+\eta}\right)^r + C\rho_n{}^{-r\tau} \theta^{\sqrt{R}\ln n},$$

which yields the RHS of $(M1)_r$ due to $(\overline{\mathrm{Poly}})$ if R is taken sufficiently large. The LHS is proved similarly.

Proof of (ii). For Ω_{ρ_n}, it is enough to consider the case $\mathrm{Sep}(k_1, \ldots, k_r) = r - 1$ otherwise we can estimate all $1_{\Omega_{\rho_n}} \circ f^{k_i}$ with $k_i - k_{i-1} < \mathfrak{s}(n)$, except the first, by 1.

So we assume that $0 < k_j - k_{j-1} < R \ln n$ and $k_i - k_{i-1} \geq R \ln n$ for $i \neq j$. Since $(M1)_r$ was proven under the assumption that $\min_i (k_i - k_{i-1}) > \sqrt{R} \ln n$ we may assume that $k_j - k_{j-1} < \sqrt{R} \ln n$. Note that by (Appr) and Remark 3.3

$$\mu\left(A^+_{\rho_n}\left(A^+_{\rho_n} \circ f^k\right)\right) - \mu\left(1_{\Omega_{\rho_n}}\left(1_{\Omega_{\rho_n}} \circ f^k\right)\right) \leq 4\mu\left(A^+_{\rho_n} - 1_{\Omega_{\rho_n}}\right) \leq 4C\mu(\Omega_{\rho_n})^{1+\eta}.$$

Therefore (Mov) with $L = 1000r$ implies that

$$\mu\left(A^+_{\rho_n}\left(A^+_{\rho_n}\circ f^{k_j-k_{j-1}}\right)\right) \le C\mu(\Omega_{\rho_n})(\ln n)^{-1000r}.$$

Take $B = A^+_{\rho_n}\left(A^+_{\rho_n}\circ f^{k_j-k_{j-1}}\right)$, we get using $(\mathrm{EM})_{r-2}$ and (Poly) that

$$\begin{aligned}\mu\left(\prod_{i=1}^r 1_{\Omega_{\rho_n}}\left(f^{k_i}x\right)\right) &\le \mu\left(\prod_{i=1}^r A^+_{\rho_n}\left(f^{k_i}x\right)\right) = \mu\left(\prod_{i\ne j-1,j} A^+_{\rho_n}\left(f^{k_i}x\right)B(f^{k_{j-1}}x)\right)\\ &\le \mu\left(A^+_{\rho_n}\right)^{r-1}\mu(B) + C{\rho_n}^{-r\tau}L^{\sqrt{R}\ln n}\theta^{R\ln n}\\ &\le C\mu(\Omega_{\rho_n})^{r-1}(\ln n)^{-1000r}\end{aligned}$$

proving $(M2)_r$.

For $\overline{\Omega}_{\rho_n}$, we approximate $1_{\overline{\Omega}_{\rho_n}}$ by $\overline{A}^+_{\rho_n}$. Consider

$$\begin{aligned}&\tilde{B}_r(x_0,\cdots,x_{j-1},x_{j+1},\cdots,x_r)\\ &\;= 1_{\overline{\Omega}_{\rho_n}}(x_0,x_1)\cdots 1_{\overline{\Omega}_{\rho_n}}(x_0,x_{j-1})1_{\overline{\Omega}^{k_j-k_{j-1}}_{a\rho_n}}(x_{j-1})1_{\overline{\Omega}_{\rho_n}}(x_0,x_{j+1})\cdots 1_{\overline{\Omega}_{\rho_n}}(x_0,x_r),\\ &\hat{B}_r(x_0,\cdots,x_{j-1},x_{j+1},\cdots,x_r)\\ &\;= \overline{A}^+_{\rho_n}(x_0,x_1)\cdots\overline{A}^+_{\rho_n}(x_0,x_{j-1})\overline{A}^+_{a\rho_n}(x_{j-1},f^{k_j-k_{j-1}}x_{j-1})\overline{A}^+_{\rho_n}(x_0,x_{j+1})\cdots\overline{A}^+_{\rho_n}(x_0,x_r).\end{aligned}$$

Since $(\overline{\text{Appr}})$, and $(\overline{\text{Sub}})$ hold, we obtain from $(\mathrm{EM})_{r-1}$

$$\begin{aligned}\mu\left(\bigcap_{j=1}^r \overline{\Omega}^{k_j}_{\rho_n}\right) &\le \mu\left(\tilde{B}_r(x,\cdots,f^{k_{j-1}}x,f^{k_{j+1}}x,\cdots,f^{k_r}x)\right)\\ &\le \mu\left(\hat{B}_r(x,\cdots,f^{k_{j-1}}x,f^{k_{j+1}}x,\cdots,f^{k_r}x)\right)\\ &\le \mu^r(\hat{B}_r) + \overline{C}{\rho_n}^{-r\tau}L^{\sqrt{R}\ln n}\theta^{R\ln n}.\end{aligned}$$

Integrating with respect to all variables except x_0 and x_{j-1}, then using $(\overline{\text{Appr}})$(iv) when integrating along x_0 for any fixed value of x_{j-1}, then finally integrating along x_{j-1}, we get

$$\mu^r(\hat{B}_r) \le \left(\bar\sigma(\rho_n) + \bar\sigma(\rho_n)^{1+\eta}\right)^{r-1}\mu\left(\overline{A}^+_{a\rho_n}(x,f^{k_j-k_{j-1}}x)\right),$$

which by (3.2) gives

$$\mu^r(\hat{B}_r) \le \left(\bar\sigma(\rho_n) + \bar\sigma(\rho_n)^{1+\eta}\right)^{r-1}\overline{C}\mu(\overline{\Omega}^{k_j-k_{j-1}}_{a\rho_n})$$

Therefore, $(M2)_r$ follows from $\overline{\text{(Mov)}}$, provided that R is sufficiently large and $L = 1000r$.

Proof of (iii). Fix a large constant b that will be given below. Consider first simple targets Ω_{ρ_n}. Denoting $B(x) = \prod_{\alpha=1}^r A^+_{\rho_{2^i}}(f^{k_\alpha}x)$ for $2^i < k_1 < \cdots < k_r \le 2^{i+1}$, we obtain from (Prod), (Gr), (Appr), (Poly), and $(\mathrm{EM})_r$, that $\|B\|_{\mathbb{B}} \le CL^{r2^{i+1}}$. Thus

$$\mu\left(\left(\prod_{\alpha=1}^{r} 1_{\Omega_{\rho_{2^i}}}(f^{k_\alpha}x)\right)\left(\prod_{\beta=1}^{r} 1_{\Omega_{\rho_{2^j}}}(f^{l_\beta}x)\right)\right) \le \mu\left(\left(\prod_{\alpha=1}^{r} A^+_{\rho_{2^i}}(f^{k_\alpha}x)\right)\left(\prod_{\beta=1}^{r} A^+_{\rho_{2^j}}(f^{l_\beta}x)\right)\right)$$

$$= \mu\left(B(x)\left(\prod_{\beta=1}^{r} A^+_{\rho_{2^j}}(f^{l_\beta}x)\right)\right)$$

$$\le \mu(B)\mu\left(A^+_{\rho_{2^j}}\right)^r + CL^{r2^{i+1}}\rho_{2^i}^{-r\tau}\rho_{2^j}^{-r\tau}\theta^{2^j\varepsilon}.$$

Applying already established $(M1)_r$ to estimate $\mu(B)$, and observing that the second term is smaller than $C(L^{r2^{-b+1}})^{2^j}2^{2r\tau uj}\theta^{2^j\varepsilon}$, which is thus much smaller than the first when b is sufficiently large, we finally get $(M3)_r$.

Next, we analyze $\overline{\Omega}_{\rho_n}$. Consider

$$B^*(x, x_1, x_2 \dots x_r) = \left(\prod_{\alpha=1}^{r} 1_{\overline{\Omega}^{k_\alpha}_{\rho_{2^i}}}(x)\right)\left(\prod_{\beta=1}^{r} 1_{\overline{\Omega}_{\rho_{2^j}}}(x, x_\beta)\right).$$

By $\overline{\text{(Appr)}}$ and $(\mathrm{EM})_r$ and the already established $(M1)_r$, we get

$$\mu\left(\bigcap_{1\le\alpha,\beta\le r}(\overline{\Omega}^{k_\alpha}_{\rho_{2^i}}\bigcap\overline{\Omega}^{l_\beta}_{\rho_{2^j}})\right) \le \mu\left(B^*(x, f^{l_1}x, \dots, f^{l_r}x)\right)$$

$$\le \mu\left(\prod_{\alpha=1}^{r}\overline{A}^+_{\rho_{2^i}}(x, f^{k_\alpha}x)\right)\left(\bar\sigma(\rho_{2^j}) + \bar\sigma(\rho_{2^j})^{1+\eta}\right)^r$$

$$+ CL^{r2^{i+1}}\rho_{2^i}^{-r\tau}\rho_{2^j}^{-r\tau}\theta^{2^j\varepsilon}.$$

Using $(M1)_r$ again we observe that

$$\mu\left(\prod_{\alpha=1}^{r}\overline{A}^+_{\rho_{2^i}}(x, f^{k_\alpha}x)\right) \le C\left(\bar\sigma(\rho_{2^i}) + \bar\sigma(\rho_{2^i})^{1+\eta}\right)^r,$$

which allows to conclude the proof of $(M3)_r$ in the case of $\overline{\Omega}_{\rho_n}$. □

REMARK 3.10. In fact, analyzing the proof of Theorem 3.7 we see that the composite targets $\overline{\text{(Appr)}}$(iii) could be replaced by a weaker condition: there is a function $\sigma_r(\rho)$ such that $C^{-1}\sigma^r(\rho) < \sigma_r(\rho) < C\sigma^r(\rho)$ and

$$\int \dots \int \left(\prod_{j=1}^{r}\overline{A}^+(x, y_j)\mathrm{d}\mu(y_j)\right)\mathrm{d}\mu(x) = \sigma_r(\rho)(1 + O(\sigma^\eta(\rho))), \tag{3.3}$$

$$\int \dots \int \left(\prod_{j=1}^{r}\overline{A}^-(x, y_j)\mathrm{d}\mu(y_j)\right)\mathrm{d}\mu(x) = \sigma_r(\rho)(1 + O(\sigma^\eta(\rho))). \tag{3.4}$$

We shall call the composite targets satisfying $\overline{\text{(Mov)}}$, $\overline{\text{(Sub)}}$, $\overline{\text{(Poly)}}$, as well as $\overline{\text{(Appr)}}$ with condition (iii) replaced by (3.3)–(3.4) *weakly admissible.*

3.3. **Notes.** There is a vast literature on Borel–Cantelli Lemmas for dynamical systems starting with [137]. Some representative examples dealing with hyperbolic systems are [4, 37, 61, 75, 78, 82, 83, 89, 98, 114] while [31, 32, 98, 106, 107, 108, 120, 153] deal with systems of zero entropy. The later cases are more complicated as counterexamples in [60, 71] show. Survey [7] reviews the results obtained up to 2009 and contains many applications, some of which parallel the results of Sections 4–9 of the present paper.

Examples of systems with multiple exponential mixing include expanding maps, volume preserving Anosov diffeomorphsims [24, 133], time one maps of contact Anosov flows [124], mostly contracting systems [30, 47], partially hyperbolic translations on homogeneous spaces [112], and partially hyperbolic automorphisms of nilmanifolds [74].

The limit theorems for smooth systems which are only assumed to be multiply exponentially mixing (but without any additional assumptions) are considered in [22, 35, 150]. [68] obtains a Logarithm Law for hitting times under an assumption of superpolynomial mixing which is weaker than our exponentially mixing assumption. We note that in our approach the *exponential* rate of mixing is crucial for verifying the condition $(M3)_r$ pertaining to interscale independence. Therefore it is an open problem to ascertain if similar results hold under weaker mixing assumptions.

4. MULTILOG LAWS FOR RECURRENCE AND HITTING TIMES

In this section we apply the results of Section 3 to obtain MultiLog Laws for multiple exponentially mixing diffeomorphisms and flows. We assume that f is a smooth diffeomorphism of a compact $d-$dimensional Riemannian manifold M preserving a smooth[5] measure μ. From now on, we take $\mathbb{B}$ in Definition 3.1 to be the space of Lipschitz observables defined over M^{r+1}.

4.1. **Results.** Let (f, M, μ) be a smooth dynamical system. Let $d_n^{(r)}(x,y)$ be the r-th minimum of

$$d(x, fy), \cdots, d(x, f^n y).$$

The following result was obtained for a large class of weakly hyperbolic systems as a consequence of dynamical Borel–Cantelli Lemmas

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(1)}(x,x)|}{\ln n} = \frac{1}{d}, \tag{4.1}$$

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(1)}(x,y)|}{\ln n} = \frac{1}{d}. \tag{4.2}$$

In particular, the following results are known.

[5] In this paper a *smooth measure* means a measure which has a Lipschitz density with respect to the Riemannian volume.

THEOREM 4.1.

(a) *If a smooth system (f, M, μ) has superpolynomial decay of correlations for Lipschitz observables, that is,*

$$|\mu(A(x)B(f^n x)) - \mu(A)\mu(B)| \leq a(n)\|A\|_{Lip}\|B\|_{Lip}$$

where $\lim_{n\to\infty} n^s a(n) = 0\ \forall s$, then for **all** *$x$ (4.2) holds for a.e. y. If in addition, f has positive entropy, then (4.1) holds for a.e. x.*

(b) *If, in addition, f is partially hyperbolic then for* **all** *x and a.e. y*

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(1)}(x,y)| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{d}. \tag{4.3}$$

In part (a), (4.1) is proven in [143, Theorem 1] and (4.2) is proven in [68, Theorem 4]. Part (b) is proven in [48, Theorem 7].

QUESTION 4.2. If (f, μ) is exponentially mixing, then (4.3) holds for **all** x and a.e. y.

MULTILOG LAW FOR RECURRENCE AND FOR HITTING TIMES. The goal of this section is to obtain an analogue of (4.3) for multiple hits as well as for returns for multiple exponentially mixing systems as in Definition 3.1.

DEFINITION 4.3. Given a smooth system (f, M, μ), define

$$\mathcal{G}_r = \left\{x:\ \text{for a.e. } y,\quad \limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,y)| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{rd}\right\},$$

$$\overline{\mathcal{G}}_r = \left\{x: \limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,x)| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{rd}\right\}.$$

THEOREM 4.4. *Suppose that $(f, M, \mu, \mathbb{B})$ is $(2r+1)$-fold exponentially mixing.*[6] *Then*

(a) $\mu(\mathcal{G}_r) = 1$;

(b) $\mu(\overline{\mathcal{G}}_r) = 1$.

FAILURE OF THE MULTILOG LAWS FOR GENERIC POINTS. Naturally, one can ask if in fact, $\mathcal{G}_r$ equals to M. If $r = 1$ the answer is often positive (see Theorem 4.1(b)). It turns out that for larger r the answer is often negative.

DEFINITION 4.5. Given a function $\zeta : \mathbb{N} \to \mathbb{N}^*$, define

$$\mathcal{H} = \left\{x:\ \text{for a.e. } y,\quad \text{for all } r \geq 1: \limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,y)| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{d}\right\},$$

$$\overline{\mathcal{H}}_\zeta = \left\{x:\ \text{for all } r \geq 1: \limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,x)|}{\zeta(n)} = \infty\right\}.$$

[6] As seen from Proposition 3.9, part (a) holds for $2r$-fold exponentially mixing systems.

THEOREM 4.6. *Suppose that the periodic points of f are dense. Then*

(a) If $\mathscr{G}_1 = M$, then $\mathscr{H}$ contains a G_δ dense set.

(b) For any $\zeta : \mathbb{N} \to \mathbb{N}^*$, $\overline{\mathscr{H}}_\zeta$ contains a G_δ dense set.

Thus for $r \ge 2$ topologically typical points do **not** belong to $\mathscr{G}_r$ or $\overline{\mathscr{G}}_r$.

FAILURE OF THE MULTILOG LAWS FOR NON MIXING SYSTEMS. THE CASE OF TORAL TRANSLATIONS. Theorem 4.6 emphasizes the necessity of a restriction on x in Theorem 4.4. In a similar spirit, we show that the mixing assumptions made in this paper are essential. To this end we consider the case when the dynamical system is $(T_\alpha, \mathbb{T}^d, \lambda)$ where T_α is the translation of vector α and λ is the Haar measure on $\mathbb{T}^d$.

Define

$$\mathscr{E}_r = \left\{ x : \text{ for a.e. } y, \quad \limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,y)| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{2d} \right\},$$

$$\overline{\mathscr{E}}_r = \left\{ x : \limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,x)| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{d} \right\}.$$

THEOREM 4.7. *For λ-a.e. $\alpha \in \mathbb{T}^d$, the system $(T_\alpha, \mathbb{T}^d, \lambda)$, satisfies*

(a) $\lambda(\mathscr{G}_1) = 1$ and $\lambda(\mathscr{E}_r) = 1$ for $r \ge 2$;

(b) $\overline{\mathscr{E}}_r = M$ for all $r \ge 1$.

The proof requires different techniques from the rest of the results of this section. Namely, it is related to the Borel–Cantelli Lemmas in the context of homogeneous dynamics on the space of lattices, and we will therefore give it in Section 9 after we introduce the necessary tools.

Take for example part (b) in the simple case of a circle rotation and $r = 1$. Clearly, $\overline{\mathscr{E}}_1 = M$ is equivalent to the fact that

$$A_n = \{k \in [1,n] : \|k\alpha\| \le n^{-1}(\ln n)^{-s}\}$$

is not empty for infinitely many n, if and only if $s \le 1$. Using the theory of continued fraction this problem could be reduced to a problem on the growth of partial quotients. Recall ([104, Theorem 16]) that for all α we have

$$\max(k \in [1,n] : \|k\alpha\| = \|q_m\alpha\|,$$

where q_m is the largest denominator of the continued fractions which does not exceed n. In addition, we have ([104, Theorems 9 and 13])

$$\frac{1}{a_{m+1}+1} \le q_m\|q_m\alpha\| \le \frac{1}{a_{m+1}}.$$

Next by a theorem of Paul Levy (see [42, Theorem 7.4]) for almost every α we have

$$\lim_{M\to\infty} \frac{\ln q_M}{M} = \frac{\pi^2}{12\ln 2}. \tag{4.4}$$

From the above fact it follows easily that there is a constant K such that for almost all α (in particular for α satisfying (4.4)) the following holds:

(a) If $\alpha \in A_n$ and n is sufficiently large then $\max_{m \le K \ln n} a_m \ge \frac{(\ln n)^s}{K}$ and

(b) If $\max_{m \le \ln n/K} a_m \ge K(\ln n)^s$ and n is sufficiently large then $\alpha \in A_n$.

Using exponential mixing of the Gauss map, one immediately sees from a classical dynamical Borel–Cantelli Lemma that for each K_1, K_2 for Lebesgue almost every α, the events

$$\left\{ \max_{m \le K_1 M} a_m \ge K_2 M^s \right\}$$

happen infinitely often iff $s \le 1$. Thus $s = 1$ is critical value for the infinite occurrence of the inequality $\max_{k \le n} \|k\alpha\| < n(\ln n)^{-s}$.

We note that the growth of the partial quotients for continued fractions is intimately related to the geodesic excursions on the modular surface (see, e.g., [145]). Due to the so called Dani correspondence principle (see §9.2 and the notes of Section 9), extending (b) in the case $r = 1$ to higher dimensions can be done by the use of a dynamical Borel–Cantelli Lemma for cusp excursions of appropriate diagonal actions on the space of lattices instead of the continued fraction algorithm. Similarly the inhomogeneous case of Theorem 4.7(a) for $r = 1$ and 2 can be reduced to a dynamical Borel–Cantelli Lemma for an appropriate diagonal actions on the space of affine lattices. (The case $r \ge 2$ for returns and $r = 3$ for visits could be reduced to the case of smaller r by elementary means). The details will be given in §9.10.

THE CASE OF FLOWS. Here we describe the analogue results of Theorems 4.4 and 4.6 for flows. Let ϕ be a smooth flow on a $(d+1)$-dimensional Riemannian manifold M preserving a smooth measure μ.

Observe that if $\phi^t(y)$ is close to x for some t, then the same is true for $\phi^{\tilde t}(y)$ with $\tilde t$ close to t. Thus we would like to count only one return for the whole connected component lying in the neighborhood of x. Namely, for some fixed $\rho > 0$, for $i \ge 0$, let $[t_i^-, t_i^+]$ denote the consecutive time intervals such that $\phi^t(y) \in B(x,\rho)$ for $t \in [t_i^-, t_i^+]$. Let t_i be the argmin of $d(x, \phi^t(y))$ for $t \in [t_i^-, t_i^+]$. Let $d_n^{(r)}(x, y)$ be the r−th minimum of

$$d(x, \phi^{t_1}(y)), \dots, d(x, \phi^{t_k}(y)), \quad t_k \le n < t_{k+1}. \tag{4.5}$$

Theorem 4.4 and Theorem 4.6 have the following counterpart in the case of flows. Note that the dimension of the manifold in the case of flows is $d+1$. Let us recall the definition of $(r+1)$-fold exponentially mixing flows, which is similar to Definition 3.1.

DEFINITION 4.8 ($(r+1)$-fold exponentially mixing flows for $r \ge 1$)**.** We say that a flow ϕ is $(r+1)$-fold exponentially mixing if there exist constants $C > 0, \theta < 1$ such that for each $(r+1)$ tuple $0 = t_0 \le t_1 \le \ldots \le t_r$ such that $\forall j \in [0, r-1]$, $t_{j+1} - t_j \ge T$, then

$$\forall A \in \mathbb{B}, \left| \int_X A(x, \phi^{t_1} x, \cdots, \phi^{t_r} x) d\mu(x) - \mu^{r+1}(A) \right| \le C\theta^T \|A\|_{\mathbb{B}},$$

where $\mathbb{B}$ is the space of Lipschitz functions on M^{r+1}.

THEOREM 4.9. *Suppose that the smooth system* $(\phi, M, \mu, \mathbb{B})$ *is* $(2r+1)$*-fold exponentially mixing. Then*

(a) $\mu(\mathcal{G}_r) = 1$;
(b) $\mu(\overline{\mathcal{G}}_r) = 1$.

If, in addition, periodic points of ϕ *are dense then*

(c) *If* $\mathcal{G}_1 = M$ *then* $\mathcal{H}$ *contains a* G_δ *dense set;*
(d) *For any* $\zeta : \mathbb{N} \to \mathbb{N}^*$, $\overline{\mathcal{H}}_\zeta$ *contains a* G_δ *dense set.*

4.2. **Slow recurrence and the proof of Theorem 4.4.** Since μ is a smooth measure, there is a smooth function $\gamma(x)$ such that

$$\mu\left(B(x,\rho)\right) = \gamma(x)\rho^d + O\left(\rho^{d+1}\right), \tag{4.6}$$

where the constant in $O\left(\rho^{d+1}\right)$ is uniform in x.

Given $x \in M$, let

$$\Omega_{x,\rho} = \{y : d(x,y) \le \rho\} \tag{4.7}$$

and[7]

$$\overline{\Omega}_\rho = \left\{(x,y) : d(x,y) \le \frac{\rho}{(\gamma(x))^{1/d}}\right\} \tag{4.8}$$

We use the notation $\Omega^k_{x,\rho}$ for the event $1_{\Omega_{x,\rho}} \circ f^k$. We also recall the notation $\overline{\Omega}^k_\rho = \{x : (x, f^k x) \in \overline{\Omega}_\rho\}$. We also keep the notation $\sigma(\rho) = \mu(\Omega_{x,\rho})$, and $\bar\sigma(\rho) = (\mu \times \mu)(\overline{\Omega}_\rho)$.

For $s \ge 0$, we let $\rho_n = n^{-1/d} \ln^{-s} n$, and recall that $N^n_{\rho_n}$ denotes the number of times $k \le n$ such that Ω^k_{x,ρ_n} (or $\overline{\Omega}^k_{\rho_n}$) occurs.

By compactness, there exists a constant $c > 0$ such that

$$\left\{(x,y) : d(x,y) \le c^{-1}\rho\right\} \subset \overline{\Omega}_\rho \subset \left\{(x,y) : d(x,y) \le c\rho\right\}.$$

Thus the statement of Theorem 4.4 becomes equivalent to the following:

(a) If $s > \frac{1}{rd}$, then for μ-a.e. x (and for μ-a.e. y in the case of $\Omega^k_{x,\rho}$), we have that for large n, $N^n_{\rho_n} < r$.
(b) If $s \le \frac{1}{rd}$, then for μ-a.e. x (and for μ-a.e. y in the case of $\Omega^k_{x,\rho}$), there are infinitely many n such that $N^n_{\rho_n} \ge r$.

With the notation $\mathbf{S}_r = \sum_{j=1}^\infty \left(2^j \mathbf{v}_j\right)^r$ where $\mathbf{v}_j = \sigma(\rho_{2^j})$ (in the Ω_{x,ρ_n} case) or $\mathbf{v}_j = \bar\sigma(\rho_{2^j})$ (in the $\overline{\Omega}_{\rho_n}$ case), we see from (4.6) that $\mathbf{S}_r = \infty$ if and only if $s \le \frac{1}{rd}$.

Hence Theorem 4.4 follows from Corollary 3.8, since $(f, M, \mu, \mathbb{B})$ is $(2r+1)$-fold exponentially mixing, provided we establish the following.

PROPOSITION 4.10.

(a) *For* μ*-a.e.* x *the targets* (Ω_{x,ρ_n}) *are simple admissible targets.*

[7] In the definition of the composite target $\overline{\Omega}_\rho$, we include the factor $(\gamma(x))^{-1/d}$ because we want that for every x, $\int 1_{\overline{\Omega}_\rho}(x,y)d\mu(y)$ be essentially the same number to be able to check $\overline{\text{(Appr)}}$(iii) for these targets.

(b) The targets $(\bar{\Omega}_{\rho_n})$ *are composite admissible targets.*

The rest of this section is devoted to the

Proof of Proposition 4.10. Observe first that with the definition of ρ_n and (4.6), we have that (Poly) and $\overline{\text{(Poly)}}$ hold for every x for the target sequences (Ω_{x,ρ_n}) as well as for the sequence $(\bar{\Omega}_{\rho_n})$.

We proceed with the proof of (Appr) and $\overline{\text{(Appr)}}$ and $\overline{\text{(Sub)}}$ properties.

LEMMA 4.11. *For each x, the targets $\Omega_{x,\rho}$ satisfy* (Appr). *The targets $\bar{\Omega}_\rho$ satisfy* $\overline{\text{(Appr)}}$ *and* $\overline{\text{(Sub)}}$.

Proof. For the targets $\Omega_{x,\rho}$, the statement follows from Lemma 3.4 by taking $\Phi(y) = d(x,y)$ (that is a Lipschitz function), $a_1(\rho) = 0$ and $a_2(\rho) = \rho$.

For the targets $\bar{\Omega}_\rho$, we use Lemma 3.6. We take $\Phi(x,y) = d(x,y)\gamma(x)^{1/d}$, $a_1(\rho) = 0$ and $a_2(\rho) = \rho$. We check $(h1)$ since $\gamma(x)/\gamma(y)$ is bounded for $(x,y) \in X \times X$. Property $(h2)$ is obvious. As for $(h3)$ it follows from the definition of $\gamma(x)$ in (4.6).

Finally, for any k_1, k_2, when $x \in \bar{\Omega}_\rho^{k_1} \cap \bar{\Omega}_\rho^{k_2}$, we have

$$d(f^{k_1}x, f^{k_2}x) \le d(x, f^{k_1}x) + d(x, f^{k_2}x) \le \frac{2\rho}{(\gamma(x))^{1/d}} \le \frac{a\rho}{(\gamma(f^{k_1}x))^{1/d}},$$

for some $a > 0$. Hence $\bar{\Omega}_\rho^{k_1} \cap \bar{\Omega}_\rho^{k_2} \subset f^{-k_1}\bar{\Omega}_{a\rho}^{k_2-k_1}$, which is $\overline{\text{(Sub)}}$. Lemma 4.11 is proved. □

Next we prove the (Mov) (for a.e. x) and $\overline{\text{(Mov)}}$ properties. For this we state a lemma on recurrence for the multiple mixing system (f, M, μ) that is of an independent interest. We first introduce two definitions.

DEFINITION 4.12 (Slowly recurrent points)**.** Call x *slowly recurrent* for the system (f, M, μ) if for each $A, K > 0$, there $\exists \rho_0$ such that for all $\rho < \rho_0$ and for all $n \le K|\ln\rho|$ we have

$$\mu(B(x,\rho) \cap f^{-n}B(x,\rho)) \le \mu(B(x,\rho))|\ln\rho|^{-A}.$$

DEFINITION 4.13 (Slowly recurrent system)**.** Call the system (f, M, μ) *slowly recurrent* if for each $A > 0$ $\exists \rho_0$ such that for all $\rho < \rho_0$ and for all $n \in \mathbb{N}^*$ we have

$$\mu\left(\left\{x : d(x, f^n x) < \rho\right\}\right) \le |\ln\rho|^{-A}.$$

LEMMA 4.14. *Suppose that $(f, M, \mu, \mathbb{B})$ is 2-fold exponentially mixing. Then*

(i) (f, M, μ) is slowly recurrent.
(ii) Almost every point is slowly recurrent.

As a consequence, we have that

(a) For μ-a.e. x, the targets Ω_{x,ρ_n} satisfy (Mov).
(b) The targets $\bar{\Omega}_{\rho_n}$ satisfy $\overline{\text{(Mov)}}$.

Proof. Take $B = A^2$. If $k \geq B \ln|\ln\rho|$, take $\hat\rho = |\ln\rho|^{-A}$. By 2-fold exponential mixing, we get

$$\mu(x : d(x, f^k x) \leq \rho) \leq \mu(x : d(x, f^k x) \leq \hat\rho) \\ \leq \mu\left(\overline{A}^+_{\hat\rho}(x, f^k x)\right) \leq C\left(\hat\rho^d + \hat\rho^{d+d\eta} + \hat\rho^{-\tau}\theta^k\right) \leq |\ln\rho|^{-2A}, \tag{4.9}$$

provided ρ is sufficiently small.

Now fix any $1 \leq k \leq B \ln|\ln\rho|$. Denote $\|f\|_1 = \max_{x\in M} \|Df(x)\|$. Assume that x satisfies $d(x, f^k x) \leq \rho$, then for any l we have that

$$d(f^{(l-1)k}(x), f^{lk}x) \leq \|f\|_1^{(l-1)k}\rho$$

If we take $L = [4B\ln|\ln\rho|/k] + 1$ we find that

$$d(x, f^{Lk}x) \leq \sum_{l\leq L-1} \|f\|_1^{lk}\rho \leq \sqrt{\rho},$$

provided ρ is sufficiently small. But $kL \geq B\ln|\ln\sqrt{\rho}|$, hence (4.9) applies and we get

$$\mu(x : d(x, f^k x) \leq \rho) \leq \mu(x : d(x, f^{Lk}x) \leq \sqrt{\rho}) \leq |\ln\rho|^{-A},$$

proving (i).

We proceed now to the proof of (ii). Define for $j, k \in \mathbb{N}^*$

$$H_{j,k}(x) := \mu(B(x, 1/2^j) \cap f^{-k}B(x, 1/2^j)).$$

Note that

$$\begin{aligned}\int H_{j,k}(x)\mathrm{d}\mu(x) &= \iint 1_{[0,1/2^j]}d(x,y)1_{[0,1/2^j]}d(x, f^k y)\mathrm{d}\mu(x)\mathrm{d}\mu(y) \\ &\leq \iint 1_{[0,1/2^j]}d(x,y)1_{[0,1/2^{j-1}]}d(y, f^k y)\mathrm{d}\mu(x)\mathrm{d}\mu(y) \\ &\leq C\mu(B(x,1/2^j))\int 1_{[0,1/2^{j-1}]}d(y, f^k y)\mathrm{d}\mu(y)\end{aligned}$$

where we used that $\mu(B(y, 1/2^j)) \leq C\mu(B(x, 1/2^j))$ for any $x, y \in M$. Part i) then implies that for sufficiently large j it holds that

$$\int H_{j,k}(x)\mathrm{d}\mu(x) \leq \mu(B(x, 1/2^j))\, j^{-A-3}.$$

For such j we get from Markov inequality

$$\mu\left(x : \exists k \in (0, Kj] : H_{j,k}(x) > \mu(B(x, 1/2^j))\, j^{-A}\right) \leq Kj^{-2}.$$

Hence Borel–Cantelli Lemma implies that for almost every x there exists $\bar{j}$ such that $H_{j,k}(x) \leq \mu(B(x, 1/2^j))\, j^{-A}$ for every $j \geq \bar{j}$ and every $k \in (0, Kj]$, which implies (ii).

Finally, (a) and (b) clearly follow from (ii) and (i) respectively. Lemma 4.14 is thus proved. □

With Lemmas 4.11 and 4.14, the proof of Proposition 4.10 is finished. □

Proof of Theorem 4.4. Theorem 4.4 directly follows from Proposition 4.10 and Corollary 3.8. □

4.3. **Generic failure of the MultiLog Law. Proof of Theorem 4.6.**

Proof. To prove part (a), we first prove that periodic points belong to $\mathcal{H}_r$. By assumption, for any $x \in M$ and almost every y,

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(1)}(x,y)| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{d}. \tag{4.10}$$

Since $d_n^{(r)}(x,y) \geq d_n^{(1)}(x,y)$, it follows that for any $x \in M$, any $r \geq 1$, and a. e. y

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,y)| - \frac{1}{d}\ln n}{\ln\ln n} \leq \frac{1}{d}. \tag{4.11}$$

To prove the opposite inequality let

$$\mathcal{H}_{m,l,r} = \left\{ x : \exists\, \mathcal{Y} \text{ open, } \mu(\mathcal{Y}) > 1 - \frac{1}{l} : \forall y \in \mathcal{Y}, \frac{|\ln d_m^{(r)}(x,y)| - \frac{1}{d}\ln m}{\ln\ln m} > \frac{1}{d} - \frac{1}{l} \right\}.$$

We have that

$$\left\{ x : \text{ for a.e. } y, \forall r \geq 1 : \limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,y)| - \frac{1}{d}\ln n}{\ln\ln n} \geq \frac{1}{d} \right\} = \bigcap_{\substack{l\geq 1\\ r\geq 1}} \bigcup_{m\geq 1} \mathcal{H}_{m,l,r}.$$

But $\mathcal{H}_{m,l,r}$ is an open set. Hence we finish if we show that for any fixed r and l, $\bigcup_m \mathcal{H}_{m,l,r}$ contains the dense set of periodic points. Let $\bar{x}$ be a periodic point of period p. Take U to be some small neighbourhood of $\bar{x}$ and denote by Λ the Lipschitz constant of f^p in U.

By (4.10), there exist $n \geq \exp\circ\exp(\Lambda + pr)$ and $\mathcal{Y}$ such that $\mu(\mathcal{Y}) > 1 - \frac{1}{l}$, such that for every $y \in \mathcal{Y}$, there exists $k \in [1,n]$ satisfying

$$d(\bar{x}, f^k y) \leq \left(\frac{1}{n}\right)^{\frac{1}{d}} \left(\frac{1}{\ln n}\right)^{\frac{1}{d}-\frac{1}{2l}}.$$

Then

$$d(\bar{x}, f^{k+pj} y) = d(f^{pj}\bar{x}, f^{k+pj} y) \leq \Lambda^r \left(\frac{1}{n}\right)^{\frac{1}{d}} \left(\frac{1}{\ln n}\right)^{\frac{1}{d}-\frac{1}{2l}}, \quad 0 \leq j \leq r-1.$$

Hence for $y \in \mathcal{Y}$ and $m = n + p(r-1)$, we have that

$$d_m^{(r)}(\bar{x}, y) \leq \Lambda^r \left(\frac{1}{n}\right)^{\frac{1}{d}} \left(\frac{1}{\ln n}\right)^{\frac{1}{d}-\frac{1}{2l}} < \left(\frac{1}{m}\right)^{\frac{1}{d}} \left(\frac{1}{\ln m}\right)^{\frac{1}{d}-\frac{1}{l}},$$

because we took $n \geq \exp\circ\exp(\Lambda + pr)$. Hence $\bar{x} \in \mathcal{H}_{m,l,r}$ and the proof of (a) is finished.

We now turn to the proof of (b). Given any function $\zeta : \mathbb{N} \to \mathbb{N}^*$, define

$$\mathcal{A}_{m,l} = \left\{ x : |\ln d_m^{(l)}(x,x)| > m\zeta(m) \right\}.$$

Observe that $\overline{\mathcal{H}}_\zeta \subset \bigcap_l \bigcup_m \mathcal{A}_{m,l}$. But $\mathcal{A}_{m,l}$ is open and $\bigcup_m \mathcal{A}_{m,l}$ clearly contains the periodic points. Part (b) is thus proved. □

4.4. **The case of flows. Proof of Theorem 4.9.** The proof proceeds in the same way as for diffeomorphisms with minimal modifications that we now explain. First, we need to modify the targets

$$\Omega_{x,\rho} = \{y : \exists\, s \in [0,1], d(x, \phi^s y) \le \rho\},$$

and

$$\bar{\Omega}_\rho = \left\{(x,y) : \exists\, s \in [0,1], d(x,\phi^s y) \le \frac{\rho}{\gamma(x)^{1/d}}\right\}$$

where $\gamma(x) = \lim_{\rho\to 0} \mu(\Omega_{x,\rho})/\rho^d$. Consider the targets

$$\Omega^n_{x,\rho} = \phi^{-n}\Omega_{x,\rho}, \quad \bar{\Omega}^n_\rho = \{x : (x, \phi^n x) \in \bar{\Omega}_\rho\}$$

for $n \in \mathbb{N}^*$ and let $\sigma(\rho) = \mu(\Omega_{\rho,x})$, $\bar{\sigma}(\rho) = (\mu \times \mu)(\bar{\Omega}_\rho)$.

To prove (a) and (b) of Theorem 4.9 we can apply Corollary 3.8 to the smooth system $(\phi, M, \mu, \mathbb{B})$ and to the targets $\Omega^n_{x,\rho}$ and $\bar{\Omega}^n_\rho$. For this, we just need to see that these targets are admissible. This can be checked as in the proof of Proposition 4.10, with very minor differences. Let us check for instance that ($\overline{\text{Sub}}$) holds for $\bar{\Omega}^n_\rho$. Note that when $x \in \bar{\Omega}^{n_1}_\rho \cap \bar{\Omega}^{n_2}_\rho$ for $n_1 < n_2$, we have some s_1, $s_2 \in [0,1]$ such that

$$d(x, \phi^{n_1+s_1} x) \le \frac{\rho}{(\gamma(x))^{1/d}}, \quad d(x, \phi^{n_2+s_2} x) \le \frac{\rho}{(\gamma(x))^{1/d}}.$$

Hence

$$\begin{aligned} d(\phi^{n_1} x, \phi^{n_2+s_2-s_1} x) &\le \max_{s\in[-1,0]} \|\phi^s\|_{C^1}\, d(\phi^{n_1+s_1} x, \phi^{n_2+s_2} x) \\ &\le \max_{s\in[-1,1]} \|\phi^s\|_{C^1} \frac{2\rho}{(\gamma(x))^{1/d}} \le \frac{a\rho}{(\gamma(\phi^{n_1} x))^{1/d}} \end{aligned}$$

for some $a > 0$. It follows that $\bar{\Omega}^{n_1}_\rho \cap \bar{\Omega}^{n_2}_\rho \subset \phi^{-n_1}\bar{\Omega}^{n_2-n_1}_{a\rho}$, which is ($\overline{\text{Sub}}$). [8] As for the proofs of (Mov) and ($\overline{\text{Mov}}$), they are obtained as in the case of maps *via* the notion of slow recurrence. We say that a point x is *slowly recurrent* for the flow if for each $A, K > 0$, there $\exists \rho_0$ such that for all $\rho < \rho_0$ for all $n \le K|\ln\rho|$ we have

$$\mu\left(\Omega_{x,\rho} \cap \Omega^n_{x,\rho}\right) \le \mu(\Omega_{x,\rho})|\ln\rho|^{-A}.$$

Similarly we say that the flow is *slowly recurrent* if for each $A > 0$ $\exists \rho_0$ such that for all $\rho < \rho_0$ for all $n \in \mathbb{N}^*$ we have

$$\mu\left(\bar{\Omega}^n_\rho\right) \le |\ln\rho|^{-A}.$$

The same proof of Lemma 4.14 then shows that if the system $(\phi, M, \mu, \mathbb{B})$ is exponentially mixing, it holds that μ-a.e. point is slowly recurrent for the flow, and that the flow is *slowly recurrent*. Properties (Mov) and ($\overline{\text{Mov}}$) are immediate consequences.

---

[8] When $s_2 - s_1 < 0$, we modify $\bar{\Omega}_\rho$ by $\tilde{\Omega}_\rho = \left\{(x,y) : \exists\, s \in [-1,1], d(x,\phi^s y) \le \frac{\rho}{\gamma(x)^{1/d}}\right\}$ and get $\bar{\Omega}^{n_1}_\rho \cap \bar{\Omega}^{n_2}_\rho \subset \phi^{-n_1}\tilde{\Omega}^{n_2-n_1}_{a\rho}$, which gives $(M2)_r$ by the argument of Proposition 3.9(ii).

The proof of part c) and part d) also proceeds in the same way as for maps. Namely we first see that periodic orbits of the flow belong to $\mathscr{H}_r$ and $\overline{\mathscr{H}}_r$ and then use the genericity argument. □

4.5. **Notes.** Many authors obtain Logarithm Law (4.2) for hitting times as a consequence of dynamical Borel–Cantelli Lemmas. See [37, 48, 67, 89] and references therein. [68] also studies return times. We note that [68] works under much weaker conditions than those imposed in the present paper, however, its results are valid only for $r = 1$ (the first visit).

The works [94, 98, 105, 111] study the recurrence problem when the lim sup in (4.2) is replaced by lim inf. In particular, [111] proves that for several expanding maps

$$\liminf_{n\to\infty} \frac{n\, d_n^{(1)}(x,y)}{\ln\ln n}$$

exists for almost all y.

Theorem 4.7 shows that some systems may satisfy logarithmic laws for $r = 1$ that are the same as in the exponentially mixing case, but fail to do so for $r \geq 2$. Logarithm Laws for unipotent flows were obtained in [10, 11, 72, 98]. It is not known which kind of MultiLog Laws hold for such flows.

5. POISSON LAW FOR NEAR RETURNS

In this section we suppose that μ is a smooth measure and that $(f, M, \mu, \mathbb{B})$ is an r-fold exponentially mixing system for all r. In the previous section we verified properties $(M1)_r$ and $(M2)_r$ for the targets $\Omega_{x,\rho}$ given by (4.7), for almost every x, and for the targets $\overline{\Omega}_\rho$ given by (4.8). Moreover, we have

$$\lim_{\rho\to 0} \rho^{-d}\sigma(\rho) = \gamma(x) \quad \text{and} \quad \lim_{\rho\to 0} \rho^{-d}\bar{\sigma}(\rho) = 1,$$

where $\sigma(\rho) = \mu(\Omega_{\rho,x})$ and $\bar{\sigma}(\rho) = (\mu\times\mu)(\overline{\Omega}_\rho)$. Accordingly Theorem 2.13 gives the following.

THEOREM 5.1.

(a) *For almost all x the following holds. Let y be uniformly distributed with respect to μ. The number of visits of $\{f^k(y)\}_{k\in[1,\tau\rho^{-d}]}$ to $B(x,\rho)$ converges to a Poisson distribution with parameter $\tau\gamma(x)$ as $\rho \to 0$. Moreover letting $n = \tau\rho^{-d}$ we have the sequence*

$$\frac{d_n^{(1)}(x,y)}{\rho}, \frac{d_n^{(2)}(x,y)}{\rho}, \dots, \frac{d_n^{(r)}(x,y)}{\rho}, \dots \tag{5.1}$$

converges to the Poisson process with measure $\gamma(x)\tau d t^{d-1} \mathrm{d}t$.

(b) *Let x be chosen uniformly with respect to μ. Then the number of visits of $\{f^k(x)\}_{k\in[1,\tau\rho^{-d}]}$ to $B(x, \frac{\rho}{\gamma^{1/d}(x)})$ converges to a Poisson distribution with parameter τ as $\rho\to 0$.*

Proof. All the results except for Poisson limit for (5.1) follow from Theorem 2.13. To prove the Poisson limit for (5.1) we need to check that for each choice of $r_1^- < r_1^+ < r_2^- < r_2^+ < \cdots < r_s^- < r_s^+$ the number of times $k \in [1, \tau\rho^{-d}]$ where $d(x, f^k y) \in \left[r_j^-\rho, r_j^+\rho\right]$ are converging to independent Poisson random variables with parameters

$$\gamma(x)\int_{r_j^-}^{r_j^+} \tau d t^{d-1} \mathrm{d}t = \gamma(x)\tau\left[(r_j^+)^d - (r_j^-)^d\right].$$

But this follows from Theorem 2.14. The latter theorem can be applied since by Lemma 3.4 the targets

$$\Omega_\rho^{k,i} = \{y : d(x, f^k y) \in [r_i^-\rho, r_i^+\rho]\}$$

satisfy (Appr) and hence $(\widetilde{M1})_r$ holds by Proposition 3.9. □

There are two natural questions dealing with improving this result. In part (a) we would like to specify more precisely the set of x where the Poisson limit law for hits holds. In part (b) we would like to remove an annoying factor $\gamma^{1/d}(x)$ from the denominator. Regarding the first question we have

CONJECTURE 5.2. *If f is r-fold exponentially mixing for all r, then the conclusion of Theorem 5.1(a) holds for* **all** *non-periodic points.*

Regarding the second question we have the following.

THEOREM 5.3. *Let x be chosen uniformly with respect to μ. Then the number of visits of $\{f^k(x)\}_{k\in[1,\tau\rho^{-d}]}$ to $B(x,\rho)$ converges to a mixture of Poisson distributions. Namely, for each l,*

$$\lim_{\rho\to 0} \mu\big(\mathrm{Card}(n \le \tau\rho^{-d} : d(x, f^n x) \le \rho) = l\big) = \int_M e^{-\gamma(z)\tau} \frac{(\gamma(z)\tau)^l}{l!} \mathrm{d}\mu(z). \tag{5.2}$$

In other words, to obtain the limiting distribution in Theorem 5.3 we first sample $z \in M$ according to the measure μ and then consider a Poisson random variable with parameter $\tau\gamma(z)$.

COROLLARY 5.4. *If f preserves a smooth measure and is r-fold exponentially mixing for Lipschitz observables for all $r \ge 2$, then*

(a) *For almost all x we have that if $\tau_\varepsilon(y)$ is the first time an orbit of y enters $B(x,\varepsilon)$ then for each t*

$$\lim \mu(y : \tau_\varepsilon(y)\varepsilon^d > t) = e^{-\gamma(x)t}$$

(b) *If $T_\varepsilon(x)$ is the first time the orbit of x returns to $B(x,\varepsilon)$ then*

$$\lim \mu(x : T_\varepsilon(x)\varepsilon^d > t) = \int_M e^{-\gamma(z)t} \mathrm{d}\mu(z).$$

Proof. This is a direct consequence of Theorems 5.1(a) and 5.3. For example to get part (b), take $l = 0$ in (5.2). □

Proof of Theorem 5.3. Consider the targets

$$\hat{\Omega}_\rho(x,y) = \{(x,y) \in M \times M : d(x,y) \le \rho\}$$

and let $\hat{\Omega}_\rho^k = \{x : (x, f^k x) \in \hat{\Omega}_\rho\}$. Note that $(M2)_r$ for $\overline{\Omega}_\rho^k$ implies $(M2)_r$ for $\hat{\Omega}_\rho^k$. However, $(M1)_r$ is false for targets $\hat{\Omega}_\rho^k$. We now argue similarly to the proof of Theorem 3.7 to obtain that for separated tuples $k_1, k_2, \dots, k_r$,

$$\mu\left(\bigcap_{j=1}^r \hat{\Omega}_\rho^{k_j}\right) = \rho^{rd} \int_M \gamma^r(z) \mathrm{d}\mu(z)(1 + o(1)). \tag{5.3}$$

Namely, note that

$$\int 1_{\hat{\Omega}_\rho}(x_0,x_1)\dots 1_{\hat{\Omega}_\rho}(x_0,x_r)\mathrm{d}\mu(x_0)\mathrm{d}\mu(x_1)\dots\mathrm{d}\mu(x_r)$$
$$= \int \mu^r(B(x_0,\rho))\mathrm{d}\mu(x_0) = \rho^{rd}(1+O(\rho))\int_M \gamma^r(x_0)\mathrm{d}\mu(x_0).$$

Thus approximating $1_{\hat{\Omega}_\rho}$ by $\hat{A}_\rho^\pm$ satisfying $\overline{(\mathrm{Appr})}$, and applying $\overline{(\mathrm{EM})}_r$ to the functions

$$\hat{B}_\rho^+(x_0,\cdots,x_r) = \hat{A}_\rho^+(x_0,x_1)\cdots\hat{A}_\rho^+(x_0,x_r),$$
$$\hat{B}_\rho^-(x_0,\cdots,x_r) = \hat{A}_\rho^-(x_0,x_1)\cdots\hat{A}_\rho^-(x_0,x_r),$$

we get that if $k_{j+1} - k_j > R|\ln\rho|$ for all $0 \le j \le r-1$, then

$$\mu\left(\bigcap_{j=1}^r \hat{\Omega}_\rho^{k_j}\right) \le \mu\left(\hat{B}_\rho^+(x_0, f^{k_1}x_0,\cdots,f^{k_r}x_0)\right)$$
$$\le \mu\left(\hat{B}_\rho^+(x_0,\cdots,x_r)\right) + C\rho^{-r\sigma}\theta^{R|\ln\rho|}$$
$$\le \left(\rho^d + C\rho^{d(1+\eta)}\right)^r \int_M \gamma^r(z)\mathrm{d}\mu(z) + C\rho^{-r\sigma}\theta^{R|\ln\rho|},$$

and, likewise,

$$\mu\left(\bigcap_{j=1}^r \hat{\Omega}_\rho^{k_j}\right) \ge \left(\rho^d - C\rho^{d(1+\eta)}\right)^r \int_M \gamma^r(z)\mathrm{d}\mu(z) - C\rho^{-r\sigma}\theta^{R|\ln\rho|}.$$

Taking R large we obtain (5.3).

Summing (5.3) over all well separated couples with $k_j \le \tau\rho^{-d}$ and using that the contribution of non-separated couples is negligible due to $(M2)_r$ we obtain

$$\lim_{\rho\to 0}\int_M \binom{N_{\rho,\tau,x}}{r}\mathrm{d}\mu(x) = \int_M \frac{\gamma^r(z)}{r!}\mathrm{d}\mu(z),$$

where

$$N_{\rho,\tau,x} = \mathrm{Card}\big(k \le \tau\rho^{-d} : d(x, f^k x) \le \rho\big).$$

Since the right-hand side coincides with factorial moments of the Poisson mixture from (5.2), the result follows. □

5.1. **Notes.** Early works on Poisson Limit Theorems for dynamical systems include [39, 46, 90, 91, 92, 138]. [33, 84, 88, 136] prove Poisson law for visits to balls centered at a *good* point for nonuniformly hyperbolic dynamical systems and show that the set of good points has a full measure. [48] obtains Poisson Limit Theorem for partially hyperbolic systems. Some of those papers, including [29, 48, 84, 87] show that in various settings the hitting time distributions are Poisson for **all** non-periodic points (cf. our Conjecture 5.2). The rates of convergence under appropriate mixing conditions are discussed in [2, 3, 85]. The Poisson limit theorems for flows are obtained in [128, 132]. Convergence on the level of random measures where one records some extra information about the close encounters, such as for example, the distance of approach is discussed in [48, 63, 64]. A mixed exponential distribution for a return time for dynamical systems similar to Corollary 5.4(a) have been obtained in [41, 141]. See also review papers [80, 144] and the references therein. Corollary 5.4(b) appears to be the first result establishing the mixture of exponential distributions as a limiting distribution for close returns in dynamical systems.

For more discussion of the distribution of the entry times to small measure sets we refer the readers to [40, 97, 144, 160] and references therein. We also refer to Section 10 for related results in the context of extreme value theory.

6. GIBBS MEASURES ON THE CIRCLE: LAW OF ITERATED LOGARITHM FOR RECURRENCE AND HITTING TIMES

6.1. **Gibbs measures.** The goal of this section is to show how absence of the hypothesis of smoothness on the invariant measure μ may also alter the law of multiple recurrence and hitting times.

For simplicity we consider the case where f is an expanding map of the circle $\mathbb{T}$ and μ is a Gibbs measure with Lipschitz potential g. Adding a constant to g if necessary we may and will assume in all the sequel that the topological pressure of g is 0, that is,

$$P(g) = \int g \mathrm{d}\mu + h_\mu(f) = 0. \tag{6.1}$$

This means (see [148] for background on Gibbs measures) that for each $\varepsilon > 0$ there is a constant K_ε such that if $B_n(x,\varepsilon)$ is the Bowen ball

$$B_n(x,\varepsilon) = \{y : d(f^k y, f^k x) \le \varepsilon \text{ for } k = 0,\dots,n-1\},$$

then

$$K_\varepsilon^{-1} \le \frac{\mu(B_n(x,\varepsilon))}{\exp\left[\left(\sum_{k=0}^{n-1} g(f^k x)\right)\right]} \le K_\varepsilon.$$

We denote

$$f_u = \ln|f'|, \tag{6.2}$$

$\lambda = \lambda(\mu)$ the Lyapunov exponent of μ

$$\lambda = \lim_{n\to\infty} \frac{\ln|(f^n)'(x)|}{n} = \int f_u \mathrm{d}\mu > 0,$$

and by $\mathbf{d}$ the dimension of the measure μ

$$\mathbf{d} = \lim_{\delta\to 0} \frac{\ln \mu(B(x,\delta))}{\ln \delta}.$$

We know from [122] that the limit exists for μ-a.e. x and

$$\mathbf{d} = h_\mu(f)/\lambda = -\frac{\int g \mathrm{d}\mu}{\int f_u \mathrm{d}\mu}$$

where the last step relies on (6.1).

We say that μ is f*-conformal* if there is a constant K such that for each x and each $0 < r \le 1$,

$$K^{-1} \le \frac{\mu(B(x,r))}{r^{\mathbf{d}}} \le K. \tag{6.3}$$

Below, for the sake of brevity, we will call f-conformal measures, simply conformal.[9]

It is known (see, e.g., [133]) that μ is conformal if and only if g can be represented in the form

$$g = t f_u - P(t f_u) + \tilde{g} - \tilde{g} \circ f$$

for some Hölder function $\tilde{g}$ and $t \in \mathbb{R}$.

Denote

$$\psi(x) = g(x) + \mathbf{d} f_u(x); \tag{6.4}$$

then we have $\int \psi \mathrm{d}\mu = 0$ under the assumption $P(g) = 0$. Define $\sigma = \sigma(\mu)$ by the relation

$$\sigma^2 = \int \psi^2 \mathrm{d}\mu + 2\sum_{n=1}^{\infty} \int \psi\left(\psi \circ f^n\right) \mathrm{d}\mu. \tag{6.5}$$

The goal of this section is to prove the following:

THEOREM 6.1.

(a) If μ is conformal then Theorems 4.4 and 4.6 remain valid with d replaced by $\mathbf{d}$.

(b) If μ is **not** *conformal then for μ almost every x and $\mu\times\mu$ almost every (x,y), it holds that*

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,x)| - \frac{1}{\mathbf{d}}\ln n}{\sqrt{2(\ln n)(\ln\ln\ln n)}} = \frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}, \tag{6.6}$$

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,y)| - \frac{1}{\mathbf{d}}\ln n}{\sqrt{2(\ln n)(\ln\ln\ln n)}} = \frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}. \tag{6.7}$$

[9]The measures satisfying (6.3) are called *Ahlfors regular* in geometric measure theory. The term *conformality* comes from the fact that the Jacobian J of f with respect to μ and the expansion coefficient λ_u (with respect to a suitable Riemannian metric) are related by $J = \lambda_u^{\mathbf{d}}$, which is similar to the relation which holds for $\mathbf{d}$-dimensional conformal maps with respect to the volume measure.

6.2. **Preliminaries on expanding circle maps and their Gibbs measures.** Here we prepare for the proof of Theorem 6.1 by collecting some facts on expanding maps of the circle and their Gibbs measures.

We first check multiple mixing for such maps.

Recall we take $\mathbb{B} = \mathrm{Lip}$. Let us denote by $\|\cdot\|_{\mathrm{Lip}}$ the Lipschitz norm

$$\|\phi\|_{\mathrm{Lip}} = \int |\phi| \mathrm{d}\mu + \sup_{x,y\in\mathbb{T}} \frac{|\phi(x)-\phi(y)|}{d(x,y)}$$

for $\phi \in \mathbb{B}$.

PROPOSITION 6.2. *For each Gibbs measure μ, the system $(f, \mathbb{T}, \mu, \mathbb{B})$ is r-fold exponentially mixing for any $r \geq 2$.*

This fact is well known but for the reader's convenience we provide the argument in in §A.2.

In the rest of the argument it will be important that if μ is a Gibbs measure then there are positive constants a, b such that for all sufficiently small ρ and for all x,

$$\rho^a \leq \mu(B(x,\rho)) \leq \rho^b. \tag{6.8}$$

We also need the fact that Gibbs measures are *doubling*, in the sense that there is a constant R such that for each x, ρ we have

$$\mu\big(B(x, 4\rho)\big) \leq R\mu(B(x,\rho)). \tag{6.9}$$

We recall the proofs of (6.8) and (6.9) in §B.2.

We also need a lemma on the fluctuations of the local dimension of Gibbs measures for expanding circle maps.

LEMMA 6.3.

(a) $\sigma(\mu) = 0$ if and only if μ is conformal.

(b) If $\sigma > 0$ then for μ almost every x

$$\limsup_{\delta\to 0} \frac{|\ln\mu(B(x,\delta))| - \mathbf{d}|\ln\delta|}{\sqrt{2|\ln\delta|(\ln\ln|\ln\delta|)}} = \frac{\sigma}{\sqrt{\lambda}}, \quad \liminf_{\delta\to 0} \frac{|\ln\mu(B(x,\delta))| - \mathbf{d}|\ln\delta|}{\sqrt{2|\ln\delta|(\ln\ln|\ln\delta|)}} = -\frac{\sigma}{\sqrt{\lambda}}.$$

The proof of this lemma is also given in Appendix B.

6.3. **The targets.** Given $x \in M$, let

$$\Omega_{x,\rho} = \{y : d(x,y) \leq \rho\}, \quad \bar{\Omega}_\rho = \big\{(x,y) : d(x,y) \leq \rho\big\}.$$

We use the notation $\Omega^k_{x,\rho}$ for the event $1_{\Omega_{x,\rho}} \circ f^k$. We also recall the notation $\bar{\Omega}^k_\rho = \{x : (x, f^k x) \in \bar{\Omega}_\rho\}$. In the sequel we will always assume that (ρ_n) is a sequence such that $\rho_n > n^{-u}$ for some u.

We caution the reader that the targets $\bar{\Omega}_\rho$ are not admissible targets in the non-conformal case, so we need to use a roundabout approach, different from Section 4, for proving Theorem 6.1(b).

On the other hand, we will need a modification of the argument of Lemma 4.14 to show that for any Gibbs measure μ and for μ-a.e. $x \in M$, the targets

Ω_{x,ρ_n} are admissible for $(f, M, \mu, \mathbb{B})$. The difference with the case of smooth measures, is that it does not hold anymore that $\mu(B(y, 1/2^j)) \le C\mu(B(x, 1/2^j))$ for any $x, y \in M$, while this was used in the proof of Lemma 4.14.

LEMMA 6.4. *For any Gibbs measure μ, for μ-a.e. $x \in M$, the targets Ω_{x,ρ_n} are admissible for $(f, M, \mu, \mathbb{B})$.*

Proof. Due to (6.8) and (6.9), all the properties of admissible targets except for (Mov) are obtained exactly as in the smooth measure case. To prove (Mov), we modify the argument of Lemma 4.14 to overcome the fact that it does not hold anymore that $\mu(B(y, 1/2^j)) \le C\mu(B(x, 1/2^j))$ for any $x, y \in M$.

In fact we can prove more than (Mov) in this context of expanding circle maps. Namely we can show that for a.e. x and all k

$$\mu(B(x,\rho) \cap f^{-k}B(x,\rho)) \le \mu(B(x,\rho))^{1+\eta}. \tag{6.10}$$

We consider two cases.

(I) $k > \varepsilon|\ln\rho|$ where ε is sufficiently small (see case (II) for precise bound on ε). Take A^+_ρ such that $A^+_\rho = 1$ on $B(x,\rho)$, $\int A^+_\rho d\mu \le 2\mu(B(x,\rho))$ and $\|A^+_\rho\|_{Lip} \le C\rho^{-\tau}$ for some $\tau = \tau(\mu)$. Let $\hat\rho = \rho^\sigma$ where σ is a small constant. Then (A.3) gives

$$\begin{aligned}\mu(B(x,\rho) \cap f^{-k}B(x,\rho)) &\le \int A^+_{\hat\rho}(A^+_\rho \circ f^k)\mathrm{d}\mu \\ &\le 4\mu(B(x,\rho))\mu(B(x,\hat\rho)) + 2C\bar\theta^k\hat\rho^{-\tau}\mu(B(x,\rho)) \\ &\le C\mu(B(x,\rho))\left(\rho^{\sigma b} + \rho^{\varepsilon|\ln\bar\theta|}\rho^{-\tau\sigma}\right)\end{aligned}$$

for some $0 < \bar\theta < 1$. Taking σ small we can make the second term smaller than $\rho^{\varepsilon|\ln\bar\theta|/2}$ which is enough for (Mov) in view of already established (Poly). Note that no restrictions on x are imposed in case (I).

(II) $k \le \varepsilon|\ln\rho|$. In this case for a.e. x the intersection $B(x,\rho) \cap f^{-k}B(x,\rho)$ is empty for small ρ due to the Proposition 6.5 below. □

PROPOSITION 6.5. *([16, Lemma 5]) Let $T : X \to X$ be a Lipschitz map with Lipschitz constant $L > 1$ on a compact metric space X. If μ is an ergodic measure with $h_\mu(T) > 0$. Then for almost every x, there exists $\rho_0(x) > 0$ such that for all $\rho \le \rho_0(x)$, and all $0 < k \le \frac{1}{2L}|\ln\rho|$, we have $T^{-k}B\left(x,\rho\right) \cap B\left(x,\rho\right) = \emptyset$.*

The case of composite targets $\bar\Omega_\rho$ is more complicated, except for the conformal case.

In the conformal case, the following Lemma is obtained exactly as in Proposition 4.10 that dealt with the smooth measure case, so we omit its proof.

LEMMA 6.6. *If μ is conformal, then the targets $\bar\Omega_{\rho_n}$ defined by* (4.8) *are weakly admissible in the sense of Remark 3.10.*

6.4. **The conformal case.**

Proof of Theorem 6.1 (a). We take $\rho_n = n^{-1/\mathbf{d}}(\ln n)^{-s}$. By Lemmas 6.4 and 6.6, the targets targets Ω_{x,ρ_n} are admissible for μ-a.e. $x \in M$ and the targets $\bar\Omega_{\rho_n}$

are composite weakly admissible. Consequently, the proof of Theorem 6.1 (a) follows exactly as that of Theorems 4.4 and 4.6 corresponding to the smooth measure case. □

6.5. **The non conformal case. Proof of Theorem 6.1(b).** The proof of Theorem 6.1(b) relies on the lim inf part of Lemma 6.3(b).

6.5.1. *The iterated logarithm law for hitting times: Proof of* (6.7) *from Theorem 6.1(b).* For $\varepsilon > 0$ and $c > 0$ arbitrary let

$$\rho_n = \rho_n(c) = \frac{1}{n^{1/\mathbf{d}}} \exp\left(-c\sqrt{2(\ln n)(\ln\ln\ln n)}\right). \tag{6.11}$$

$$\vartheta^{\pm}_{\varepsilon}(\delta) = \delta^{\mathbf{d}} \exp\left((1\pm\varepsilon)\frac{\sigma}{\sqrt{\lambda}}\sqrt{2|\ln\delta|\,(\ln\ln|\ln\delta|)}\right),$$

$$\tilde{\vartheta}^{\pm}_{\varepsilon,c}(n) = \vartheta^{\pm}_{\varepsilon}(\rho_n(c)).$$

Then

$$\tilde{\vartheta}^{\pm}_{\varepsilon,c}(n) = \frac{1}{n}\exp\left(\left(-c\mathbf{d} + (1\pm\varepsilon)\frac{\sigma}{\sqrt{\mathbf{d}\lambda}} + \eta_n\right)\sqrt{2\ln n(\ln\ln\ln n)}\right)$$

for some $\eta_n \to 0$ as $n \to \infty$.

The lim inf in Lemma 6.3, has the following straightforward consequences, for any $\varepsilon > 0$ and for μ almost every x:

There exists $n(x)$ such that for $n \geq n(x)$, we have

$$\mu\left(\Omega_{x,\rho_n}\right) \leq \tilde{\vartheta}^{+}_{\varepsilon,c}(n). \tag{6.12}$$

For a subsequence $n_l \to \infty$ we have

$$\mu\left(\Omega_{x,\rho_{n_l}}\right) \geq \tilde{\vartheta}^{-}_{\varepsilon,c}(n_l). \tag{6.13}$$

Now it follows that for any $r \geq 1$, $S_r = \sum_{k=1}^{\infty}\left(2^k\mu(\Omega_{x,\rho_{2^k}})\right)^r$ is finite if $c > (1+\varepsilon)\frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}$ and is infinite if $c < (1-\varepsilon)\frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}$. Hence (6.7) follows from Proposition 6.2, Lemma 6.4 and Corollary 3.8. □

6.5.2. *The iterated logarithm law for return times: Proof of the upper bound in* (6.6). Now we turn to the proof of

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,x)| - \frac{1}{\mathbf{d}}\ln n}{\sqrt{2(\ln n)(\ln\ln\ln n)}} \leq \frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}. \tag{6.14}$$

Since $d_n^{(r)}(x,x) \geq d_n^{(1)}(x,x)$, we only need to show (6.14) for $r = 1$.

Denote

$$r_n = \frac{1}{n^{1/\mathbf{d}}}\exp\left\{-(1+2\varepsilon)\frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}\sqrt{2(\ln n)(\ln\ln\ln n)}\right\}.$$

Let $N_k = 2^k$. Similarly to Section 2 it is enough to show that for almost all x, for all sufficiently large k we have that

$$d(x, f^m x) \geq r_{N_k} \text{ for } m = 1, \ldots, N_k.$$

Proposition 6.5 allows us to further restrict the range of m by assuming $m \geq \bar{\varepsilon}\ln N_k$, where $\bar{\varepsilon}$ is sufficiently small.

We say $x \in \mathbb{T}$ is n-good if $\mu(B(x,r_n)) \leq \vartheta^+(r_n)$. Fix k_0 and let

$$\mathscr{A}_k = \{x : x \text{ is } n\text{-good for } n \geq N_k \\ \text{but } d(x, f^m x) \leq r_{N_k} \text{ for some } m = \bar{\varepsilon}\ln N_k, \dots, N_k\}.$$

Let $\mathscr{X}_k = \{x_{j,k}\}_{j=1}^{l_k}$ to be a maximal r_{N_k} separated set of N_k–good points. Thus if x is N_k good then there is j such that $x \in B(x_{j,k}, r_{N_k})$. Therefore if $f^m x \in B(x, r_{N_k})$ then $f^m x \in B(x_{j,k}, 2r_{N_k})$. Fix a large K, for $m \leq K\ln N_k$, (6.10) is telling us that

$$\mu\big(B(x_{j,k}, 2r_{N_k}) \cap f^{-m}B(x_{j,k}, 2r_{N_k})\big) \leq K\mu(B(x_{j,k}, 2r_{N_k}))^{1+\eta}.$$

while for $m > K\ln N_k$ we get by exponential mixing that

$$\mu\big(B(x_{j,k}, 2r_{N_k}) \cap f^{-m}B(x_{j,k}, 2r_{N_k})\big) \leq K\mu(B(x_{j,k}, 2r_{N_k}))^{2}.$$

Summing those estimates, we obtain

$$\sum_{m=\bar{\varepsilon}\ln N_k}^{N_k} \mu\big(B(x_{j,k}, 2r_{N_k}) \cap f^{-m}B(x_{j,k}, 2r_{N_k})\big) \leq K\mu(B(x_{j,k}, 2r_{N_k}))e^{-\kappa\sqrt{k}}$$

for some $\kappa = \kappa(\bar{\varepsilon}) > 0$. Since $B(x_{j,k}, r_{N_k}/2)$ are disjoint for different j, by (6.9) we conclude that

$$\sum_j \mu\big(B(x_{j,k}, 2r_{N_k})\big) \leq R\sum_j \mu\big(B(x_{j,k}, r_{N_k}/2)\big) \leq R.$$

It follows that

$$\mu(\mathscr{A}_k) \leq KRe^{-\kappa\sqrt{k}}.$$

Now the result follows from the classical Borel–Cantelli Lemma.

6.5.3. *Proof of the lower bound in (6.6).* Next we prove that

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,x)| - \frac{1}{\mathbf{d}}\ln n}{\sqrt{2(\ln n)(\ln\ln\ln n)}} \geq \frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}. \tag{6.15}$$

Suppose p to be a fixed point of f. Take the Markov partition $\mathscr{P}_n$ of $\mathbb{T}$ such that if $P_n \in \mathscr{P}_n$, then $f^n(\partial P_n) = p$. Denote $P_n(x) = \{P_n \in \mathscr{P}_n : x \in P_n\}$, and consider two sequences $(k_j(x))$ and $(n_j(x))$, $j \in \mathbb{N}$ such that $k_0(x) = n_0(x) = 0$,

$$n_j(x) = \min\left\{n > k_{j-1}(x)^2 : \mu(P_n(x)) \geq \vartheta_\varepsilon^-(|P_n(x)|)\right\}$$

and

$$k_j(x) = \frac{2}{\mu(P_{n_j}(x))}.$$

Let

$$\mathscr{A}_j = \left\{x :\ \mathrm{Card}\left\{k_{j-1}(x) \leq k \leq k_j(x) : f^k x \in P_{n_j}(x)\right\} \geq r\right\}.$$

Then

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,x)| - \frac{1}{\mathbf{d}}\ln n}{\sqrt{2(\ln n)(\ln\ln\ln n)}} \geq \frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}(1-2\varepsilon)$$

if x belongs to infinitely many $\mathscr{A}_j$s.

Denote by $\mathbb{P}(\cdot|\cdot)$ the conditional probability, and let $\mathscr{F}_j = \mathscr{B}\left(\mathscr{P}_{k_1}, \cdots, \mathscr{P}_{k_j}\right)$ be the σ algebra generated by the itineraries up to the time k_j. We will use the following Lévy's extension of the Borel–Cantelli Lemma.

THEOREM 6.7. ([157, §12.15]) If $\sum_j \mathbb{P}\left(\mathscr{A}_{j+1}|\mathscr{F}_j\right) = \infty$ *a.s.*, then $\mathscr{A}_j$ happen infinitely many times almost surely.

Hence (6.15) follows from the lemma below.

LEMMA 6.8. *There exists $c^* > 0$, such that for almost all x there is $j_0 = j_0(x)$ such that $\mathbb{P}\left(\mathscr{A}_{j+1}|\mathscr{F}_j\right) \geq c^*$ for all $j \geq j_0$.*

Proof. For any $\Omega \subset \mathbb{T}$, $P_k \in \mathscr{P}_k$,

$$\mu\left(f^k(\Omega \cap P_k)\right) = \frac{\mu\left(f^k(\Omega \cap P_k)\right)}{\mu\left(f^k(P_k)\right)} \leq C\frac{\mu(\Omega \cap P_k)}{\mu(P_k)}$$

by bounded distortion property. Note that

$$\mathbb{P}\left(\mathscr{A}_{j+1}|\mathscr{F}_j\right)(x) = \frac{\mu\left(\mathscr{A}_{j+1} \cap P_{k_j(x)}(x)\right)}{\mu\left(P_{k_j(x)}(x)\right)} \geq C^{-1}\mu\left(f^{k_j(x)}\left(\mathscr{A}_{j+1} \cap P_{k_j(x)}(x)\right)\right).$$

By construction

$$f^{k_j(x)}\left(\mathscr{A}_{j+1} \cap P_{k_j(x)}(x)\right)$$

is the set of points $y \in \mathbb{T}$ which visit $P_{n_{j+1}(x)}$ at least r times before time

$$\bar{k}_{j+1}(x) = k_{j+1}(x) - k_j(x).$$

By Lemma 6.4 for almost all x the targets $P_{n_{j+1}}(x)$ satisfy $(M1)_r$ and $(M2)_r$ for all r. Since by construction $\lim_{j\to\infty} \mu(P_{n_j}(x))\bar{k}_j(x) = 2$ we can apply Theorem 2.13 to get

$$\mathbb{P}\left(\mathscr{A}_{j+1}|\mathscr{F}_j\right)(x) \geq C^{-1}\mu\left(f^{n_j}(\mathscr{A}_{j+1} \cap P_{n_j})\right) \geq C_0 \sum_{k=r}^{\infty} e^{-2}\frac{2^k}{k!} := c^*.$$

proving the lemma. □

6.6. **Notes.** The fact that return times for the non-conformal Gibbs measures are dominated by fluctuations of measures of the balls has been explored in various settings [25, 26, 34, 41, 86, 95, 131, 142, 154]. In particular, [79] obtains a result similar to our Lemma 6.3 in the context of symbolic systems. The papers mentioned above deal with either one-dimensional or symbolic systems. In higher dimensions even the leading term of $\ln\mu(B(x,r))$ is rather non-trivial and is analyzed in [15], while fluctuations are determined only for a limited class of systems [123]. Thus extending the results of this section to higher dimension is an interesting open problem.

7. GEODESIC EXCURSIONS

7.1. **Excursions in finite volume hyperbolic manifolds.** Let $\mathscr{Q}$ be a finite volume non-compact $(d+1)$-dimensional manifold of curvature -1. Let $S\mathscr{Q}$ denote the unit tangent bundle to $\mathscr{Q}$. For $(q,v)\in S\mathscr{Q}$, let $\gamma(t)=\gamma(q,v,t)$ be the geodesic such that $\gamma(0)=q$, $\dot\gamma(0)=v$. We call g^t the corresponding geodesic flow, defined by $g^t(\gamma(0),\dot\gamma(0))=(\gamma(t),\dot\gamma(t))$. g^t preserves a smooth measure μ which is called the Liouville measure. This measure is given by restricting the volume form defined by the symplectic structure to the energy surface. Fix a reference point $O\in\mathscr{Q}$ and let $D(q,v,t)=\mathrm{dist}(O,\gamma(t))$. According to Sullivan's Logarithm Law for excursions [152] for μ-a.e. $(q,v)\in S\mathscr{Q}$, it holds that

$$\limsup_{T\to\infty}\frac{D(q,v,T)}{\ln T}=\frac{1}{d}. \tag{7.1}$$

In fact, the Borel–Cantelli Lemma of [152] also shows that

$$\limsup_{T\to\infty}\frac{D(q,v,T)-\frac{1}{d}\ln T}{\ln\ln T}=\frac{1}{d}. \tag{7.2}$$

Here we present a multiple excursions version of (7.2). Recall ([18, Proposition D.3.12]) that $\mathscr{Q}$ admits a decomposition

$$\mathscr{Q}=\mathscr{K}\bigcup\left(\bigcup_{j=1}^{p}\mathscr{C}_j\right) \tag{7.3}$$

where $\mathscr{K}$ is a compact set and $\mathscr{C}_j$ are cusps. Each cusp is isometric to $V_i\times[L_j,\infty)$ endowed with the metric

$$ds^2=\frac{dx^2+dy^2}{y^2},\quad x\in V_j,\ \ y\in[L_j,\infty)$$

where V_j is a compact flat manifold and dx is the Euclidean metric on V_j. Cusps are disjoint, so that a geodesic cannot pass between different cusps without visiting the thick part $\mathscr{K}$ in between. To simplify the notation we assume throughout this section that the time between consecutive visits to the cusps is at least 1. (This can always be achieved by decreasing slightly the cusps and increasing the compact part. Alternatively, if the minimal time between the cusp visits is smaller than some $\kappa<1$ one can repeat the argument given below replacing g^n action by $g^{\kappa n}$). We note that[10] for each $q_0=(x_0,y_0)\in\mathscr{C}_i$ there is a unique geodesic $(\{x=x_0\})$ which remains in the cusp for all positive time. We will call this geodesic *the escaping geodesic* passing through (x_0,y_0). Let $h(q,v,t)=0$ if $\gamma(q,v,t)\in\mathscr{K}$ and $h(q,v,t)=\ln y(t)$ if $\gamma(q,v,t)=(x(t),y(t))\in\mathscr{C}_i$. It is easy to see using the triangle inequality that there exists a constant C such that

$$\left|D(q,v,t)-h(q,v,t)\right|\le C.$$

A geodesic excursion is a maximal interval I such that $\gamma(t)$ belongs to some cusp $\mathscr{C}_i$ for all $t\in I$. Then, $h(I)=\max_{t\in I}h(q,v,t)$ is called *the height of the excursion I*.

[10] We identify hereafter each cusp $\mathscr{C}_j$ with $V_i\times[L_j,\infty)$.

For every triple (q,v,T) we can order the heights of the excursions that correspond to maximal excursion intervals included inside $(0,T)$ starting from the highest one

$$H^{(1)}(q,v,T) \geq H^{(2)}(q,v,T) \geq \cdots \geq H^{(r)}(q,v,T)\ldots$$

Note that (7.2) is implied by

$$\limsup_{T\to\infty} \frac{H^{(1)}(q,v,T) - \frac{1}{d}\ln T}{\ln\ln T} = \frac{1}{d}. \tag{7.4}$$

Here we prove the following multiple excursions version of (7.4).

THEOREM 7.1. *For a.e. (q,v) and all r we have*

$$\limsup_{T\to\infty} \frac{H^{(r)}(q,v,T) - \frac{1}{d}\ln T}{\ln\ln T} = \frac{1}{rd}.$$

We also have the following byproduct of our analysis.

COROLLARY 7.2. *There are constants a_i, $i=1,\ldots,p$ such that for each $\mathfrak{h}$ the following holds. Suppose that (q,v) is uniformly distributed on $S\mathcal{Q}$. Then the number of excursions in the cusp $\mathcal{C}_i$ which finished before time T and reached the height $\frac{\ln T}{d}+\mathfrak{h}$ is asymptotically Poisson with parameter $a_i e^{-d\mathfrak{h}}$.*

In other words, for every $r\geq 1$, we have

$$\lim_{T\to\infty} \mu\left(H_i^{(r)}(q,v,T) < \frac{\ln T}{d}+\mathfrak{h}\right) = \sum_{l=0}^{r-1} \frac{(a_i e^{-d\mathfrak{h}})^l}{l!} \exp\left(-a_i e^{-d\mathfrak{h}}\right) \tag{7.5}$$

where $H_i^{(r)}$ is the r-th highest excursion to the cusp $\mathcal{C}_i$. In particular, taking $r=1$ in (7.5) we obtain

COROLLARY 7.3. *(Gumbel distribution for the maximal excursion) If (q,v) is uniformly distributed on $S\mathcal{Q}$. Let $H_i^{(1)}(q,v,T)$ denote the maximal height reached by $\gamma(q,v,t)$ up to time T inside cusp $\mathcal{C}_i$. Then*

$$\lim_{T\to\infty} \mu\left(H_i^{(1)}(q,v,T) - \frac{\ln T}{d} < \mathfrak{h}\right) = \exp\left(-a_i e^{-d\mathfrak{h}}\right).$$

7.2. **Height and angle.** We start with discussing the conditions for a geodesic which just entered a cusp to reach a certain height before exiting. This information will be important in describing the geometry of the targets we will use to establish the MultiLog for excursions.

Let Π be the plane passing through γ and the escaping geodesic. In this plane the geodesics are half circles centered at the absolute $\{y=0\}$. The half circle (geodesic) given by $(x-x_0)^2+y^2=R^2$ reaches the maximum height of $\ln R + O(1)$. Let n^* be the first integer moment of time after the beginning of the excursion. Then the y coordinate of $\gamma(n^*)$ is uniformly bounded from above and below so the radius of the circle defining the geodesic is given by $R=\dfrac{y(n^*)}{\sin\theta}$ where θ is the angle with the escaping geodesic. It follows that the condition $R\geq R_0$ is equivalent to the condition $\sin\theta \leq \dfrac{y(n^*)}{R_0}$.

DEFINITION 7.4. Given H we consider the set $\mathscr{A}_{i,H}$ which consists of points $(q,v) \in \mathscr{C}_i$ such that

(i) The first positive time $\bar{t}(q,v)$ such that the backward geodesic $\gamma(q,v,-\bar{t})$ exits the cusp satisfies $\bar{t}(q,v) \in [0,1]$;

(ii) The angle v makes with the escaping geodesic at q is less than e^{-H}.

The above discussion implies that for $(q,v) \in \mathscr{C}_i$ satisfying (i) and (ii), the geodesic starting at (q,v) will exit the cusp in backward time less than 1 and will do an excursion in future time up to height $h \ge H + O(1)$, consuming for this a time comparable to h. Conversely, there is a constant C such that any excursion reaching height $H + C$ in $\mathscr{C}_i$ satisfies (i) and (ii).

We also introduce

$$\mathscr{A}_H = \bigcup_i \mathscr{A}_{i,H}. \tag{7.6}$$

It is a basic fact (e.g., see the proof of Theorem 6 in [152]) that

$$\mu(\mathscr{A}_{i,H}) = a_i e^{-dH}(1+o(1)). \tag{7.7}$$

To prove Theorem 7.1 we define for every $k \ge 0$

$$\Omega_\rho^k = g^{-k}\mathscr{A}_{-\ln\rho}. \tag{7.8}$$

By a slight abuse of notation, we still denote the event $1_{\Omega_\rho^k}$ by Ω_ρ^k. We also keep the notation $\sigma(\rho) = \mu(\Omega_\rho)$.

For $s \ge 0, c > 0$, we let $\rho_n = cn^{-1/d}\ln^{-s} n$. Recall that $N_{\rho_n}^n(q,v)$ denotes the number of times $k \in [1,n]$ such that $\Omega_{\rho_n}^k$ occurs (i.e., $(q,v) \in \Omega_{\rho_n}^k$). Theorem 7.1 becomes equivalent to showing that for each c

(a) If $s > \frac{1}{rd}$, then for μ-a.e. (q,v), we have that for large n, $N_{\rho_n}^n < r$.

(b) If $s \le \frac{1}{rd}$, then for μ-a.e. (q,v), there are infinitely many n such that $N_{\rho_n}^{\frac{n}{2}} \ge r$.

We introduce a factor 1/2 in (b) to make sure the last excursion that starts before $n/2$ finishes before n. Here, we are using that the excursion time is comparable to the excursions height $\ln n \ll n/2$. Moreover, if K is a large constant, we can take in part (a) $\rho_n = Kn^{-1/d}(\ln n)^{-s}$. Then $\mathscr{A}_{-\ln\rho_n}$ contains all geodesics reaching the height $n^{1/d}(\ln n)^s$. Similarly, if we take $\rho_n = 1/\left(Kn^{1/d}(\ln n)^s\right)$ in part (b) then all geodesics in $\mathscr{A}_{-\ln\rho_n}$ reach the height $n^{1/d}(\ln n)^s$.

7.3. **MultiLog law for geodesic excursions.** In this section we reduce Theorem 7.1 to a statement about the quasi independence of different excursions (Lemma 7.6 below).

Observe that by (7.7), we have that

$$\mu(\Omega_{\rho_n}) \in \left[C^{-1}n^{-1}(\ln n)^{-sd}, Cn^{-1}(\ln n)^{-sd}\right] \tag{7.9}$$

Hence with the notation $\mathbf{S}_r = \sum_{j=1}^{\infty}\left(2^j\mathbf{v}_j\right)^r$ where $\mathbf{v}_j = \sigma(\rho_{2^j})$ we see from (7.9) that $\mathbf{S}_r = \infty$ if and only if $s \le \frac{1}{rd}$. We want thus to apply Corollary 3.8, but first we need to verify its conditions.

The system $(g^1, S\mathscr{Q}, \mu, \mathbb{B})$ is r-fold exponentially mixing for every $r \ge 2$ in the sense of Definition 3.1. Indeed (Prod) an (Gr) are clear, while (EM)$_r$ follows from [20, Theorem 1.1] (see also [115, Theorem 1.2]) and Remark A.1 and Theorem A.2 of our appendix.

To apply Corollary 3.8, we also need the admissibility of the targets.

PROPOSITION 7.5. *The family of targets $\{\Omega_{\rho_n}\}$ is admissible as in Definition 3.2.*

Before we prove Proposition 7.5, we first complete the

Proof of Theorem 7.1. From the equivalence stated in (a) and (b) above, and since by Proposition 7.5 the targets $\{\Omega_{\rho_n}\}$ are admissible, the lim sup of Theorem 7.1 follows from Corollary 3.8 and the fact that $\mathbf{S}_r = \infty$ if and only if $s \le \frac{1}{rd}$. □

Proof of Proposition 7.5. First, the definition of ρ_n and (7.9) imply (Poly). Next, the first time $\bar{t}(q,v) \ge 0$ such that $\gamma(q, v, -t)$ exits the cusp is Lipschitz in (q, v). Also, the angle $\Psi(q, v)$ that v makes with the escaping geodesic at q is also a Lipschitz function of (q, v). We conclude that (Appr) for the targets $\{\Omega_\rho\}$ follows from Lemma 3.4 with $\Phi(q, v) = \Psi(q, v)$, $a_1(\rho) = 0$ and $a_2(\rho) = \rho$, *modulo* a very simple modification in the proof of Lemma 3.4 to account for the extra condition that $\bar{t}(q, v) \in [0,1]$.

It remains to prove (Mov). We denote $\mathscr{A}_H^n = g^{-n}\mathscr{A}_H$. (Mov) is an immediate consequence of the quasi-independence result on the excursions given in Lemma 7.6 below. Similar quasi-independence results are obtained in [152, 134]. For completeness we will give a proof adapted to our setting in §7.4.

LEMMA 7.6. *There is a constant K such that for each $H > 0$ and each $n_1 < n_2$,*

$$\mu(\mathscr{A}_H^{n_1} \cap \mathscr{A}_H^{n_2}) \le K\mu(\mathscr{A}_H)^2.$$

Up to proving Lemma 7.6, we finished the proof of Proposition 7.5. □

7.4. **Quasi independence of excursions.**

Here we prove Lemma 7.6.

The idea of the proof is the following. Fix numbers $H_1, H_2, t_1 + 1 < t_2$. We want to show that the event that a point has an excursion which ends during the time interval $[t_1, t_1+1)$ and reaches the height of H_1 or higher, and the event that a point has an excursion which starts during the time interval $[t_2, t_2+1]$ and reaches the height of H_2 or higher, are quasi independent. Since μ is invariant by the geodesic flow we may assume that $t_1 < -1$, $t_2 > 0$. Since the points on the same geodesic experience excursions of the same height, we can take a section transversal to the flow direction. Note that if x has an excursion in the future, then all points in its local stable manifold also have an excursion in the future, while if x had an excursion in the past, then all points in its local unstable manifold also had an excursion in the past. Thus the required quasi independence comes from the local product structure of μ.

The formal proof given below is more complicated since we have to address several technical issues including the following:

(a) We know the starting time of both excursions, rather than an ending time of the first excursion as described above.
(b) Points in the same stable (unstable) manifold do not have excursion of exactly the same height;
(c) Excursions for points on the same orbit happen at different times.

Proof of Lemma 7.6. Let $\tilde{\mathscr{A}}_H = I\mathscr{A}_H$ where I denotes the involution $I(q,v) = (q,-v)$. Given $n_1, \bar{n}$ define

$$\mathscr{B}_{H,n_1,\bar{n}} = \{x : g^{n_1}x \in \mathscr{A}_H,\ g^{\bar{n}}x \in \tilde{\mathscr{A}}_H,\ g^n x \notin \mathscr{K} \text{ for } n_1 < n < \bar{n}\}.$$

Thus $\mathscr{B}_{H,n_1,\bar{n}}$ consists of points which enter a cusp at time n_1, reach the height H, and then exit the cusp at time $\bar{n}$. We have that $\mathscr{A}_H^{n_1} = \bigcup\limits_{\bar{n}>n_1} \mathscr{B}_{H,n_1,\bar{n}}$. Note that $\bar{n} - n_1 \geq H$.

Fix a small δ and let $\tilde{\mathscr{B}}_{H,n_1,\bar{n}} = \bigcup_{x\in\mathscr{B}_{H,n_1,\bar{n}}} \mathscr{W}^u\left(x, \delta e^{-\bar{n}}\right)$, where $\mathscr{W}^u\left(x,\rho\right)$ denotes the local unstable cube containing x of length ρ. Note that if $y \in \tilde{\mathscr{B}}_{H,n_1,\bar{n}}$ then $g^{\bar{n}_1} y \in \mathscr{A}_{H-1}$ for some $\bar{n}_1 \in [n_1 - 1, n_1 + 1]$ and $g^{\bar{n}_2} y \in \tilde{\mathscr{A}}_{H-1}$ for some $\bar{n}_2 \in [\bar{n} - 1, \bar{n} + 1]$. In particular for each n_1 the sets $\{\tilde{\mathscr{B}}_{H,n_1,\bar{n}}\}_{\bar{n}\geq n_1}$ have intersection multiplicity at most 3. Hence

$$\sum_{\bar{n}=n_1}^{\infty} \mu(\tilde{\mathscr{B}}_{H,n_1,\bar{n}}) \leq 3\mu(\mathscr{A}_{H-1}). \tag{7.10}$$

Since $\tilde{\mathscr{B}}_{H,n_1,\bar{n}} \cap \mathscr{A}_H^{n_2} = \emptyset$ if $\bar{n} \geq n_2$, we have for $n_2 > n_1$

$$\mu(\mathscr{A}_H^{n_1} \cap \mathscr{A}_H^{n_2}) \leq \sum_{n_1<\bar{n}<n_2} \mu\left(\tilde{\mathscr{B}}_{H,n_1,\bar{n}} \cap \mathscr{A}_H^{n_2}\right). \tag{7.11}$$

We claim that there exists a constant $C \geq 1$ such that for each $n_1 < \bar{n} < n_2$ we have

$$\mu\left(\mathscr{B}_{H,n_1,\bar{n}} \cap \mathscr{A}_H^{n_2}\right) \leq C\mu(\tilde{\mathscr{B}}_{H,n_1,\bar{n}})\mu(\mathscr{A}_H). \tag{7.12}$$

Now, (7.7), (7.10), (7.11), and (7.12), imply Lemma 7.6.

It remains to establish (7.12). To this end, fix a large $\overline{H}$ and partition a small neighborhood $\mathscr{U}$ of $\tilde{\mathscr{A}}_{\overline{H}}$ into unstable cubes of size δ. (Note that since the points in $\mathscr{U}$ are not too far from the compact part, we can take δ much smaller than the injectivity radius of any point from $\mathscr{U}$. Then unstable cubes of size δ are nice embedded submanifolds with a boundary). For $H \geq \overline{H}$, let

$$\hat{\mathscr{B}}_{H,n_1,\bar{n}} = \bigcup_{x\in B_{H,n_1,\bar{n}}} \mathscr{W}^u(g^{\bar{n}}x, \delta)$$

where $\mathscr{W}^u(y,\delta)$ denotes the element of the above partition containing y. Note that

$$\mathscr{B}_{H,n_1,\bar{n}} \subset g^{-\bar{n}}\hat{\mathscr{B}}_{H,n_1,\bar{n}} \subset \tilde{\mathscr{B}}_{H,n_1,\bar{n}}. \tag{7.13}$$

Thus

$$\mu\left(\mathscr{B}_{H,n_1,\bar{n}} \cap \mathscr{A}_H^{n_2}\right) \leq \mu\left(g^{-\bar{n}}\hat{\mathscr{B}}_{H,n_1,\bar{n}} \cap \mathscr{A}_H^{n_2}\right) = \mu\left(\hat{\mathscr{B}}_{H,n_1,\bar{n}} \cap \mathscr{A}_H^{n^*}\right)$$

where $n^* = n_2 - \bar{n} > 0$. We thus finish if we show that

$$\mu\left(\hat{\mathscr{B}}_{H,n_1,\bar{n}} \cap \mathscr{A}_H^{n^*}\right) \le C\mu(\hat{\mathscr{B}}_{H,n_1,\bar{n}})\mu(\mathscr{A}_H). \tag{7.14}$$

Indeed (7.13) and (7.14) imply (7.12). By construction, $\hat{\mathscr{B}}_{H,n_1,\bar{n}}$ is partitioned into nice unstable cubes of size δ. It suffices to show that for any such cube $\mathscr{W}$ we have

$$\mu(\mathscr{A}_H^{n^*}|\mathscr{W}) \le Ce^{-dH} \tag{7.15}$$

where $\mu(\cdot|\cdot)$ denotes the conditional expectation. Let $Q = \bigcup_{x\in\mathscr{W}} \bigcup_{|t|<\delta} \mathscr{W}^s(g^t x, \delta)$, where $\mathscr{W}^s\left(y,\delta\right)$ denotes the local stable leaf containing y of length δ. Note that if δ is sufficiently small then due to the local product structure, for each point $y \in Q$ there is unique $x \in \mathscr{W}$ and $t \in [-\delta,\delta]$ such that $y \in \mathscr{W}^s(g^t x,\delta)$. In addition if $g^{n^*}x \in \mathscr{A}_H$ then $g^{n^*}y \in \mathscr{A}_{H-1}$. Since the measure of Q is bounded from below uniformly in $\mathscr{W} \subset \mathscr{U}$, it follows that

$$\mu(\mathscr{A}_H^{n^*}|\mathscr{W}) \le \mu(\mathscr{A}_{H-1}^{n^*}|Q) = \frac{\mu(\mathscr{A}_{H-1}^{n^*} \cap Q)}{\mu(Q)} \le \frac{\mu(\mathscr{A}_{H-1}^{n^*})}{\mu(Q)} \le \bar{c}\mu(\mathscr{A}_{H-1}) \le \hat{c}e^{-dH}$$

where $\bar{c} = 1/\mu(Q)$. This establishes (7.15) and, hence (7.14) completing the proof of Lemma 7.6. □

7.5. **Poisson Law for excursions. Proof of Corollary 7.2.** Here we take

$$\rho_n := n^{-1/d}.$$

We fix $\mathfrak{h} \in \mathbb{R}$ and fix a cusp index i. With the sets $\mathscr{A}_{i,H}$ defined as in Definition 7.4, consider the targets

$$\Omega_{i,\rho_n}^k = g^{-k}\mathscr{A}_{i,-\ln\rho_n-\mathfrak{h}}.$$

As in the proof of Theorem 7.1, we have that $\{\Omega_{i,\rho_n}^k\}$ satisfies the assumptions $(M1)_r$ and $(M2)_r$ for all r. Moreover, by (7.7)

$$\lim_{n\to\infty} n\mu(\Omega_{i,\rho_n}) = a_i e^{-d\mathfrak{h}}.$$

Therefore Corollary 7.2 follows from Theorem 2.13. □

7.6. **Notes.** The logarithm law for the highest excursion was proven in [152]. The extensions for infinite volume hyperbolic manifolds are studied in [151]. Corollary 7.3 for surfaces is obtained in [96] where the authors also consider infinite volume surfaces. Papers [13, 56] obtain stable laws for geodesic windings on hyperbolic manifolds. Those papers are relevant since the main contribution to windings comes from long excursions, so the proofs of stable laws and of the Poisson laws for excursions are closely related, see, e.g., [51, 54]. In case the hyperbolic manifold under consideration is the modular surface, the length of the n-th geodesic excursion is approximately equal to the size of the n-th convergence of the continued fraction expansion of the geodesic endpoint [77], therefore the multiple Borel–Cantelli Lemma in that case follows from the results of [1].

Several authors discussed extended Logarithm Law for excursion to other homogeneous spaces. Namely, [114] studies partially hyperbolic flows on homogeneous spaces and presents applications to metric number theory. Logarithm Law for unipotent flows is considered in [10, 11, 72, 98]. In Section 9 we obtain MultiLog Law for certain diagonal flows on the space of lattices.

8. RECURRENCE IN CONFIGURATION SPACE.

8.1. **The results.** In this section we return to the study of compact manifolds, but we treat targets which have more complicated geometry than the targets from Section 4. We will see that a richer geometry of targets leads to stronger results.

Let $\mathcal{Q}$ be a compact manifold of a variable negative curvature and dimension $d+1$. Denote by $S\mathcal{Q}$ the unit tangent bundle over $\mathcal{Q}$, $\pi : S\mathcal{Q} \to \mathcal{Q}$ the canonical projection, g the geodesic flow on $S\mathcal{Q}$ preserving the Liouville measure μ.

Fix a small number $\bar{\rho} > 0$. Given a point $a \in \mathcal{Q}$ and $(q, v) \in S\mathcal{Q}$, let t_j be consecutive times where the function $t \to d(a, \pi(g^t(q, v)))$ has a local minima such that $d_j := d(a, \pi(g^{t_j}(q, v)) \leq \bar{\rho}$. Let $d_n^{(r)}(a, (q, v))$ be the r-th minima among the numbers $\{d_j\}_{t_j \leq n}$.

THEOREM 8.1.

(a) For each $a \in \mathcal{Q}$ and almost every $(q, v) \in S\mathcal{Q}$,

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(a, (q, v))| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{rd}.$$

(b) For almost every $(q, v) \in S\mathcal{Q}$,

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(q, (q, v))| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{rd}.$$

Note that in contrast with Section 4 there are no exceptional points for hitting. We also obtain a Poisson limit theorem. Denote

$$B_\rho(a) = \{q \in \mathcal{Q} : d(a, q) < \rho\},$$
$$\hat{B}_\rho(a) = \{(q, v) \in S\mathcal{Q} : d(a, q) < \rho, v \in S_q\mathcal{Q}\},$$
$$\Omega_{a,\rho} = \bigcup_{t\in[0,\varepsilon]} \phi^t \hat{B}_\rho(a),$$
$$\overline{\Omega}_\rho = \left\{\big((a, u), (q, v)\big) \in S\mathcal{Q} \times S\mathcal{Q} : \exists s \in [0, \varepsilon], d\left(a, \pi(g^s(q, v))\right) < \rho\right\}.$$

The following fact proven in Appendix C will be helpful in our argument.

LEMMA 8.2. *The following limit exists and does not dependent on $a \in \mathcal{Q}$:*

$$\gamma = \lim_{\rho\to 0} \mu\left(\Omega_{a,\rho}\right) / \left(\varepsilon\rho^d\right). \tag{8.1}$$

The following will be a byproduct of our analysis and the proof will be given in §8.3.

COROLLARY 8.3. *For each $a \in \mathcal{Q}$, for every $\tau > 0$, for every $r \geq 1$, we have*

(a) $\lim_{\rho\to 0}\mu\Big((q,v)\in S\mathcal{Q} : d^{(r)}_{\tau\rho^{-d}}(a,(q,v))<\rho\Big)=\sum_{l=0}^{r-1} e^{-\tau\gamma}\frac{(\gamma\tau)^l}{l!}$,

(b) $\lim_{\rho\to 0}\mu\Big((q,v)\in S\mathcal{Q} : d^{(r)}_{\tau\rho^{-d}}(q,(q,v))<\rho\Big)=\sum_{l=0}^{r-1} e^{-\tau\gamma}\frac{(\gamma\tau)^l}{l!}$.

8.2. **MultiLog Law. Proof of Theorem 8.1.** We fix $r\in\mathbb{N}$ and consider the system $(f,S\mathcal{Q},\mu,\mathrm{Lip})$, where $f=g^{\varepsilon}$ for a small $\varepsilon>0$. We note that it follows from [124][Theorem 2.4], [48][Theorem 2], Remark A.1 and Theorem A.2 that $(f,S\mathcal{Q},\mu,\mathrm{Lip})$ is r-fold exponentially mixing for every $r\geq 2$ as in Definition 3.1.

For $k\neq 0$, we keep the notations $\Omega^k_{a,\rho}$ for the event $1_{\Omega_{a,\rho}}\circ f^k$, and $\bar{\Omega}^k_{\rho}$ for the event $\{(q,v):((q,v),f^k(q,v))\in\bar{\Omega}_{\rho}\}$. We also keep the notation $\sigma(\rho)=\mu(\Omega_{a,\rho})$ and $\bar{\sigma}(\rho)=(\mu\times\mu)(\bar{\Omega}_{\rho})$.

For $s\geq 0$, we let $\rho_n=n^{-1/d}\ln^{-s}n$, and recall that $N^n_{\rho_n}$ denotes the number of times $k\leq n$ such that Ω^k_{a,ρ_n} (or $\bar{\Omega}^k_{\rho_n}$) occurs.

The statement of Theorem 8.1 becomes equivalent to the following:

(a) If $s>\frac{1}{rd}$, then for μ-a.e. (q,v), we have that for large n, $N^n_{\rho_n}<r$.
(b) If $s\leq\frac{1}{rd}$, then for μ-a.e. (q,v), there are infinitely many n such that $N^n_{\rho_n}\geq r$.

With the notation $\mathbf{S}_r=\sum_{j=1}^{\infty}\left(2^j\mathbf{v}_j\right)^r$ where $\mathbf{v}_j=\sigma(\rho_{2^j})$ (in the Ω_{x,ρ_n} case) or $\mathbf{v}_j=\bar{\sigma}(\rho_{2^j})$ (in the $\bar{\Omega}_{\rho_n}$ case), we see from (8.1) that $\mathbf{S}_r=\infty$ if and only if $s\leq\frac{1}{rd}$.

Hence Theorem 8.1 follows from Corollary 3.8, provided we establish the following:

PROPOSITION 8.4.

(a) For any $a\in\mathcal{Q}$, the targets $\{\Omega_{a,\rho_n}\}$ are simple admissible targets as in Definition 3.2.

(b) The targets $(\bar{\Omega}_{\rho_n})$ are composite admissible targets as in Definition 3.5.

The rest of this section is devoted to the

Proof of Proposition 8.4. Properties (Prod) and (Gr) are clear. Note that $\Omega_{a,\rho}$ is a sublevel set of a Lipschitz function

$$h(q,v)=\min_{s\in[0,\varepsilon]} d(a,\pi g^s(q,v)),$$

so (Appr) follows as in Lemma 3.4. To prove the first part of Proposition 8.4, it only remains to check (Mov). That is, we need to prove the following Lemma.

LEMMA 8.5. *There exist $\eta>0$ [11] and $t_0>0$ such that for any $a\in Q$ and ρ sufficiently small,*

$$\mu(\Omega_{a,\rho}\cap g^{-t}\Omega_{a,\rho})\leq\mu(\Omega_{a,\rho})^{1+\eta}, \tag{8.2}$$

for all $t>t_0$.

[11] In fact, it can be seen from the proof that η can be taken to be d, that is, we have quasiindependence in Lemma 8.5 $\mu(\Omega_{a,\rho}\cap g^{-t}\Omega_{a,\rho})\leq C\mu(\Omega_{a,\rho})^2$

Recall that $S_q\mathcal{Q}$ is the unit tangent bundle at the point q. Denote $A_\varepsilon(q) = \bigcup_{s\in[0,\varepsilon]} g^s S_q\mathcal{Q}$, which is an embedded submanifold with boundary in $S\mathcal{Q}$ of dimension $d+1$.

LEMMA 8.6. *We let ν be the restriction of μ on $A_\varepsilon(q)$. For each $a \in \mathcal{Q}$*

$$\nu\left(A_\varepsilon(q) \cap g^{-t}\Omega_{a,\rho}\right) \le C\rho^\eta \nu(A_\varepsilon(q)). \tag{8.3}$$

Lemma 8.5 follows from Lemma 8.6 by integration on $q \in B_\rho(a)$.

Introduce $\Sigma(t,q,\varepsilon) := g^t A_\varepsilon(q)$. Note that $\Sigma(t,q,\varepsilon)$ is an embedded submanifold on $S\mathcal{Q}$ of dimension $d+1$.

The proof of the following result is given in the Appendix C.

LEMMA 8.7 (Geometry of expanded spheres in the configuration space)**.** *We have that $\pi : \Sigma(t,q,\varepsilon) \to \mathcal{Q}$ is a local diffeomorphism. Moreover for the inverse map $d\pi^{-1} : S\mathcal{Q} \to S\Sigma(t,q,\varepsilon)$ the norm $||d\pi^{-1}||$ is uniformly bounded.*

Proof of Lemma 8.6. By elementary geometry and the bounded distortion property

$$\nu(A_\varepsilon(q) \cap g^{-t}\Omega_{a,\rho}) \le C\rho^{-1}\nu(\Sigma(t,q,\varepsilon) \cap \hat{B}_{2\rho}(a)). \tag{8.4}$$

By Lemma 8.7, $||d\pi^{-1}||$ is uniformly bounded. Note that $\pi\Sigma(t,q,\varepsilon)$ is an annulus whose boundaries are spheres of radii t and $t+\varepsilon$ respectively. Note those spheres are perpendicular to the geodesics emanating from q. Since the width of annulus is equal to ε and does not depend on t, taking a maximal 1-separated set in the sphere of radius of $t+(\varepsilon/2)$ and considering associated Voronoi cells we see that $\Sigma(t,q,\varepsilon)$ can be cut into several disjoint pieces $\Sigma_j(t)$ satisfying that for each j, $\pi\Sigma_j(t)$ is contained in a ball of radius ε_2 (independent of t and q) and contains a ball of radius $\varepsilon/2$. Decreasing ε_2 if necessary we obtain that the intersection $\pi\Sigma_j(t) \cap B_{2\rho}(a)$ has only one component and since $d\pi^{-1}$ is bounded we get that

$$\nu(\Sigma_j(t) \cap \hat{B}_{2\rho}(a)) \le C(\varepsilon_1)\rho^{d+1}\nu\left(\Sigma_j(t)\right).$$

Summing over j in (8.4) we obtain (8.3), and this finishes the proof of Lemma 8.6. □

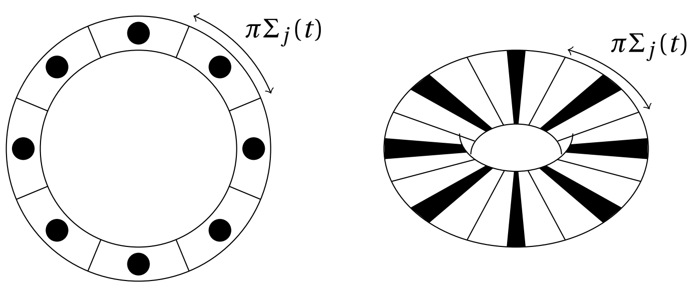

FIGURE 1. Proof of Lemma 8.6

The proof of Proposition 8.4(a) is thus completed.

Now we turn to the proof of Proposition 8.4(b). The task is to verify the conditions $\overline{\text{(Appr)}}$, $\overline{\text{(Mov)}}$ and $\overline{\text{(Sub)}}$ for the targets $\bar{\Omega}_\rho$ defined in Section 8.1. The proof of $\overline{\text{(Appr)}}$ and $\overline{\text{(Sub)}}$ is obtained from Lemma 3.6 exactly as in the proof of Lemma 4.11 that treats the case of the composite targets of Section 4.2. It is left to verify $\overline{\text{(Mov)}}$. Take $x_i \in \mathcal{Q}$, $B_i = B(x_i, \rho)$, $1 \le i \le k$ such that $\mathcal{Q} = \bigcup_{i=1}^{k} B_i$ and $k = O(\rho^{-d})$. By (8.2), for $t > t_0$,

$$
\begin{aligned}
\mu(\bar{\Omega}_\rho^t) &\le \sum_i \left\{(q,v) \in S\mathcal{Q} : \exists s \in [0,\varepsilon], d\left(q, \pi(g^{s+t}(q,v))\right) < \frac{\rho}{\gamma(q)^{1/d}}, q \in B_i\right\} \\
&\le \sum_i \left\{(q,v) \in S\mathcal{Q} : \exists s \in [0,\varepsilon], d\left(x_i, \pi(g^{s+t}(q,v))\right) < c\rho, q \in B_i\right\} \\
&\le \sum_i \mu(\Omega_{x_i, c\rho})^{1+\eta} \le \sum_i C\rho^{d(1+\eta)} \le C\rho^\eta.
\end{aligned}
$$

This completes the proof of Proposition 8.4 and finishes the proof of Theorem 8.1. □

8.3. **Poisson regime. Proof of Corollary 8.3.** Part (a) follows from Theorem 2.13, since conditions $(M1)_r$ and $(M2)_r$ are satisfied for all r, due to the results of §8.2.

The proof of part (b) follows the same argument as the proof of Theorem 5.3 except that now $(M1)_r$ is satisfied since the RHS of (5.3) takes form $\rho^d \lambda$ because λ defined by (8.1) does not depend on a. □

8.4. **Notes.** In [129], Maucourant proved that for all $a \in \mathcal{Q}$ and almost every $(q,v) \in S\mathcal{Q}$

$$\limsup_{t \to +\infty} \frac{\left|\ln d\left(a, \pi\left(g^t(q,v)\right)\right)\right|}{\ln t} = \frac{1}{d}.$$

[118] generalized Maucourant's result to the shrinking target problem for time h map.

The shrinking target problems for sets with complicated geometry is discussed in [69, 70, 72, 98, 99, 100, 135, 140].

Concerning Poisson Limits we note that visits to sets with complicated geometry naturally appear in Extreme Value Theory, see Section 10 for details. [27, 87, 65, 158] provide general conditions for the number of visits to a small neighborhood of arbitrary submanifold to be asymptotically Poisson.

9. MULTIPLE KHINTCHINE–GROSHEV THEOREM

9.1. Statements.

HOMOGENEOUS APPROXIMATIONS. For $x \in \mathbb{R}^d$, we use the notation $|x| = \sqrt{\sum x_i^2}$.

DEFINITION 9.1 ((r,s)-approximable vectors)**.** Given $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}^d$, $s \ge 0$, $c > 0$, let $D_N(\alpha, s, c)$ be the set of $k = (k_1, \dots, k_d) \in \mathbb{Z}^d$ such that

$$|k| \le N \text{ and } \exists m \in \mathbb{Z} : gcd(k_1, \dots, k_d, m) = 1 \text{ and } |k|^d \, |\langle k, \alpha\rangle + m| \le \frac{c}{\ln N (\ln\ln N)^s}.$$

Call α (r,s)*-approximable* if $\forall c>0$, $\mathrm{Card}(D_N(\alpha,s,c)) \geq 2r$ for infinitely many Ns.

THEOREM 9.2. *If $s \leq 1/r$, then the set of (r,s)-approximable vectors $\alpha \in \mathbb{T}^d$ has full measure. If $s > 1/r$, then the set of (r,s)-approximable numbers has zero measure.*

REMARK 9.3. Observe that an equivalent statement of Theorem 9.2 is to replace $2r$ with r in the definition of (r,s)-approximable vectors provided we restrict to $k \in \mathbb{Z}^d$ such that $k_1 > 0$. This will be the version that we will prove in the sequel.

INHOMOGENEOUS APPROXIMATIONS.

DEFINITION 9.4 ((r,s)-approximable pairs)**.** Given $\alpha = (\alpha_1,\ldots,\alpha_d) \in \mathbb{R}^d$ and $z \in \mathbb{R}$, $s \geq 0$ and $c > 0$, let $D_N(\alpha,z,s,c)$ be the set of $k = (k_1,\ldots,k_d) \in \mathbb{Z}^d$ such that

$$|k| \leq N \text{ and } \exists m \in \mathbb{Z} : |k|^d\, |z + \langle k,\alpha\rangle + m| \leq \frac{c}{\ln N (\ln\ln N)^s}. \tag{9.1}$$

Call the pair (α,z) (r,s)*-approximable* if for any $c > 0$, $\mathrm{Card}(D_N(\alpha,z,s,c)) \geq r$ for infinitely many Ns.

THEOREM 9.5. *If $s \leq 1/r$, then the set of (r,s)-approximable pairs $(\alpha,z) \in \mathbb{R}^d \times \mathbb{R}$ has full measure. If $s > 1/r$, then the set of (r,s)-approximable pairs $(\alpha,z) \in \mathbb{R}^d \times \mathbb{R}$ has zero measure.*

REMARK 9.6. Notice that we do not require $gcd(k_1,\ldots,k_d,m) = 1$ in Definition 9.4. This is because, in the inhomogeneous setting, when a vector $k \in \mathbb{Z}^d$ contributes to the Diophantine approximation counting problem there is no reason for the multiples of k to contribute.

EXTENSIONS. One can extend the above results to general Khintchine–Groshev $0-1$ laws for Diophantine approximations of linear forms. For example

DEFINITION 9.7 ((r,s)-simultaneously approximable vectors)**.** Given a vector $\alpha = (\alpha_1,\ldots\alpha_d) \in \mathbb{R}^d$, $s \geq 0$, $c > 0$, let $D_N(\alpha,s,c)$ be the set of $k \in \mathbb{Z}^*$ such that

$$\begin{aligned} &|k| \leq N \text{ and } \exists m \in \mathbb{Z}^d : gcd(k,m_1,\ldots,m_d) = 1 \\ &\qquad \text{and for all } i = 1,\ldots,d, \quad |k|^{\frac{1}{d}}\, |k\alpha_i + m_i| \leq \frac{c}{(\ln N)^{\frac{1}{d}} (\ln\ln N)^{\frac{s}{d}}}. \end{aligned}$$

Call α (r,s)*-simultaneously approximable* if for any $c > 0$, $\mathrm{Card}(D_N(\alpha,s,c)) \geq 2r$ for infinitely many Ns.

THEOREM 9.8. *If $s \leq 1/r$ then the set of (r,s)-simultaneously approximable vectors $\alpha \in \mathbb{T}^d$ has full measure. If $s > 1/r$ then the set of (r,s)-simultaneously approximable numbers has zero measure.*

We omit the proof of Theorem 9.8 since it is obtained by routine modification of the proof of Theorem 9.2.

9.2. **Reduction to a problem on the space of lattices.** Let $\mathscr{M}$ be the space of $(d+1)$-dimensional unimodular lattices. We identify $\mathscr{M}$ with $SL_{d+1}(\mathbb{R})/SL_{d+1}(\mathbb{Z})$. Denote by μ the Haar measure [12] on $\mathscr{M}$. Define

$$\Lambda_\alpha = \begin{pmatrix} \mathrm{Id}_d & 0 \\ \alpha & 1 \end{pmatrix}.$$

For $t \in \mathbb{R}$, we consider $g_t \in SL_{d+1}(\mathbb{R})$

$$g_t = \begin{pmatrix} 2^{-t} & & & \\ & \ddots & & \\ & & 2^{-t} & \\ & & & 2^{dt} \end{pmatrix}. \tag{9.2}$$

For a lattice $\mathscr{L} \in \mathscr{M}_{d+1}$, we say that a vector in $\mathscr{L}$ is primitive if it is not an integer multiple of another vector in $\mathscr{L}$.

Given a function f on $\mathbb{R}^{d+1}$, its *Siegel transform* $\mathscr{S}(f) : \mathscr{M} \to \mathbb{R}$ is defined by

$$\mathscr{S}(f)(\mathscr{L}) = \sum_{e \in \mathscr{L},\ e \text{ primitive}} f(e). \tag{9.3}$$

For $a > 0$, let ϕ_a be the indicator of the set[13]

$$E_a := \left\{(x, y) \in \mathbb{R}^d \times \mathbb{R} \mid x_1 > 0, |x| \in [1,2], |x|^d |y| \in [0, a]\right\}.$$

Fix $s \geq 0, c > 0$. For $M \in \mathbb{N}^*$, define

$$\nu := \frac{c}{M(\ln M)^s}, \quad \Phi_\nu := \mathscr{S}(\phi_\nu). \tag{9.4}$$

For $t \geq 0$, we define

$$A_t(M) := \{\alpha \in \mathbb{T}^d : \Phi_\nu(g_t \Lambda_\alpha) \geq 1\}.$$

It is readily checked that $\alpha \in A_t(M)$ if and only if there exists $k = (k_1, \dots, k_d)$ with $k_1 \geq 0$, and $2^t < |k| \leq 2^{t+1}$ such that

$$\exists m, \quad gcd(k_1, \dots k_d, m) = 1, \quad |k|^d |\langle k, \alpha \rangle + m| \leq \frac{c}{M(\ln M)^s}. \tag{9.5}$$

If α is such that $\Phi_\nu(g_t \Lambda_\alpha) \leq 1$ for every $t \in \mathbb{N}$, then we get that α is (r, s)-approximable if and only if there exist infinitely many Ms for which there exist $0 < t_1 < t_2 < \dots < t_r \leq M$ satisfying $\alpha \in \bigcap_{j=1}^r A_{t_j}(M)$.

But in general, for α and $t \leq M$ such that $\alpha \in A_t(M)$, there may be multiple solutions k to inequality (9.1) such that $2^t < |k| \leq 2^{t+1}$ for the same t. Since in Theorem 9.2 we are counting all solutions we have to deal with this issue.

The following proposition which will be proved in §9.3 shows that for a.e. α, multiple solutions do not occur.

PROPOSITION 9.9. *For almost every α, we have that for every M sufficiently large, for every $t \in [0, M]$, it holds that $\Phi_\nu(g_t \Lambda_\alpha) \leq 1$.*

Hence, Theorem 9.2 is equivalent to the following.

[12] the Haar measure is the unique left–translation invariant probability measure on $\mathscr{M}$.

[13] We added $x_1 > 0$ in the definition of E_a since we will restrict to vectors $k \in \mathbb{Z}^d$ with $k_1 \geq 0$.

THEOREM 9.10. *If $rs \le 1$, then for almost every $\alpha \in \mathbb{T}^d$, there exists infinitely many M for which there exist $0 < t_1 < t_2 < \ldots < t_r \le M$ satisfying*

$$\alpha \in \bigcap_{j=1}^{r} A_{t_j}(M).$$

If $rs > 1$, then for almost every $\alpha \in \mathbb{T}^d$, there exists at most finitely many M for which there exist $0 < t_1 < t_2 < \ldots < t_r \le M$ satisfying

$$\alpha \in \bigcap_{j=1}^{r} A_{t_j}(M).$$

9.3. **Modifying the initial distribution: homogeneous case.** We transformed our problem into a problem of multiple recurrence of the diagonal action g_t when applied to a piece of horocycle in the direction of $\Lambda_\alpha : \alpha \in \mathbb{T}^d$. This horocycle is exactly the full strong unstable direction of the rapidly mixing partially hyperbolic action g_t. It is, however, more convenient to work with Haar measure on $\mathscr{M}$ instead of Haar measure on Λ_α for $\alpha \in \mathbb{T}^d$. Hence, we define

$$B_t(M) := \{\mathscr{L} \in \mathscr{M} : \Phi_\nu(g_t \mathscr{L}) \ge 1\},$$

where Φ_ν is given by (9.4).

Our goal becomes to prove the following.

PROPOSITION 9.11. *For μ-almost every $\mathscr{L} \in \mathscr{M}$, we have that for every M sufficiently large, for every $t \in [0, M]$, it holds that $\Phi_\nu(g_t \mathscr{L}) \le 1$.*

THEOREM 9.12. *If $rs \le 1$, then for μ-almost every $\mathscr{L} \in \mathscr{M}$, there exists infinitely many M for which there exist $0 < t_1 < t_2 < \ldots < t_r \le M$ satisfying*

$$\mathscr{L} \in \bigcap_{j=1}^{r} B_{t_j}(M).$$

If $rs > 1$, then for μ-almost every $\mathscr{L} \in \mathscr{M}$, there exists at most finitely many M for which there exist $0 < t_1 < t_2 < \ldots < t_r \le M$ satisfying

$$\mathscr{L} \in \bigcap_{j=1}^{r} B_{t_j}(M).$$

Proof that Proposition 9.11 and Theorem 9.12 imply Proposition 9.9 and Theorem 9.10.

Recall that for $M \in \mathbb{N}$ we defined $\nu = \frac{c}{M(\ln M)^s}$. Fix $\eta > 0$ and define $\Phi^{\pm}_\nu$ as in (9.4) but with $(1+\eta)c$ and $(1-\eta)c$ instead of c. Next, define for $\beta \in \mathbb{R}^d$ and $B \in SL_d(\mathbb{R})$

$$\Lambda^-_\beta = \begin{pmatrix} \mathrm{Id}_d & \beta \\ 0 & 1 \end{pmatrix}, \quad D_B = \begin{pmatrix} B & 0 \\ 0 & 1 \end{pmatrix}.$$

Finally let

$$\tilde{\Lambda}_{\alpha,\beta,B} = D_B \Lambda^-_\beta \Lambda_\alpha.$$

Fix $0 < \varepsilon \ll \eta$. If B is distributed according to a smooth density with respect to Haar measure on $SL_d(\mathbb{R})$ in an ε neighborhood of the Identity, β is distributed in some ε neighborhood of 0 in $\mathbb{R}^d$ with a smooth density according to Haar measure of $\mathbb{T}^d$, and α is distributed according to any measure with smooth density with respect to Haar measure on $\mathbb{T}^d$, then the lattice $\tilde{\Lambda}_{\alpha,\beta,B}$ is distributed according to a smooth density in $\mathscr{M}$ with respect to the Haar measure μ. Moreover, because Λ_β^- forms the stable direction of g_t and because D_B forms the centralizer of g_t, we have that if M is sufficiently large, then

$$\Phi_\nu^-(g_t\tilde{\Lambda}_{\alpha,\beta,B}) \geq 1 \implies \Phi_\nu(g_t\Lambda_\alpha) \geq 1 \implies \Phi_\nu^+(g_t\tilde{\Lambda}_{\alpha,\beta,B}) \geq 1.$$

This shows that Proposition 9.9 and Theorem 9.10 follow from Proposition 9.11 and Theorem 9.12 respectively. □

9.4. **Rogers identities.** The following identities (see [126, 156]) play an important role in our argument. Denote

$$\mathbf{c}_1 = \zeta(d+1)^{-1}, \quad \mathbf{c}_2 = \zeta(d+1)^{-2}, \quad \text{where } \zeta(d+1) = \sum_{n=1}^{\infty} n^{-(d+1)}$$

is the Riemann zeta function.

Let f, f_1, f_2 be piecewise smooth functions with compact support on $\mathbb{R}^{d+1}$.

Define the following *Siegel transforms*

$$\mathscr{S}(f)(\mathscr{L}) = \sum_{e\in\mathscr{L},\ \text{primitive}} f(e), \quad \mathscr{S}(f_1,f_2)(\mathscr{L}) = \sum_{e_1\neq\pm e_2\in\mathscr{L},\ \text{primitive}} f_1(e_1)f_2(e_2).$$

LEMMA 9.13. *We have*

(a) $\displaystyle\int_{\mathscr{M}} \mathscr{S}(f)(\mathscr{L})\mathrm{d}\mu(\mathscr{L}) = \mathbf{c}_1 \int_{\mathbb{R}^{d+1}} f(x)\mathrm{d}x,$

(b) $\displaystyle\int_{\mathscr{M}} \mathscr{S}(f_1,f_2)(\mathscr{L})\mathrm{d}\mu(\mathscr{L}) = \mathbf{c}_2 \int_{\mathbb{R}^{d+1}} f_1(x)\mathrm{d}x \int_{\mathbb{R}^{d+1}} f_2(x)\mathrm{d}x.$

9.5. **Multiple solutions on the same scale. Proof of Proposition 9.11.** Recall that $\nu = \frac{c}{M\ln M^s}$.

LEMMA 9.14. *There exists a constant $C>0$, such that for every M, for every $t\in\mathbb{R}$, it holds that*

$$\mu\big(\Phi_\nu(g_t\mathscr{L}) > 1\big) \leq Cc^2M^{-2}(\ln M)^{-2s}.$$

For $K\geq 0$, apply the lemma for $M = 2^K$ and sum over all $t\in[0,M]$, then

$$\mu\big(\exists t\leq 2^K, \Phi_{4\nu}(g_t\mathscr{L}) > 1\big) \leq 16Cc^2 2^{-K}K^{-2s}.$$

The straightforward side of Borel–Cantelli lemma gives that for almost every $\mathscr{L}$, for K sufficiently large, for any $t\leq 2^K, \Phi_{4\nu}(g_t\mathscr{L}) \leq 1$. For the same $\mathscr{L}$, it then holds that for M sufficiently large, for any $t\leq M$, $\Phi_\nu(g_t\mathscr{L}) \leq 1$.

To finish the proof of Proposition 9.11 we give

Proof of Lemma 9.14. Since g_t preserves Haar measure on $\mathscr{M}$ it suffices to prove the lemma for $t=0$. But the condition $k_1\geq 0$ implies that

$$\Phi_\nu^2(\mathscr{L}) - \Phi_\nu(\mathscr{L}) = \sum_{e_1\neq e_2\in L \text{ primitive}} \phi_\nu(e_1)\phi_\nu(e_2) = \sum_{e_1\neq\pm e_2\in L \text{ primitive}} \phi_\nu(e_1)\phi_\nu(e_2).$$

It then follows from Rogers identity of Lemma 9.13(b) that

$$\mu(\Phi_\nu(\mathscr{L})>1) \le \mathbb{E}\left(\Phi_\nu^2(\mathscr{L})-\Phi_\nu(\mathscr{L})\right)$$
$$\le \mathbf{c}_2\left(\int_{\mathbb{R}^{d+1}}\phi_\nu(u)\mathrm{d}u\right)^2 \le Cc^2M^{-2}(\ln M)^{-2s}. \qquad \square$$

9.6. **Proof of Theorem 9.12.** We want to apply Corollary 3.8. For the system (f,X,μ) we take $(g_1,\mathscr{M},\mu)$, where μ is the Haar measure on $\mathscr{M}$. For the targets, we take $\Omega_\rho=\{\mathscr{L}:\Phi_\rho(\mathscr{L})\ge 1\}$ and $\Omega_\rho^t=g_{-t}\Omega_\rho$. Note that by the invariance of the Haar measure by g_t we have that $\mu(\Omega_\rho^t)=\mu(\Omega_\rho)$ for any t.

For $s\in\mathbb{N}$, we define the sequence $\rho_M:=\frac{c}{M(\ln M)^s}$. The conclusions of Theorem 9.12 will then follow from the conclusion of Corollary 3.8 applied to $N_{\rho_M}^M$, where N_ρ^n is the number of times $t\le n$ such that Ω_ρ^t occurs.

Indeed, recalling the definition of

$$\mathbf{S}_r=\sum_{j=1}^\infty\left(2^j\mathbf{v}_j\right)^r,\quad \mathbf{v}_j=\sigma(\rho_{2^j}),\quad \sigma(\rho)=\mu(\Omega_\rho)$$

we see that $\mathbf{S}_r=\infty$ if and only if $rs\le 1$.

Hence, to apply Corollary 3.8 and finish the proof, we only need to check the conditions of Definition 3.1 and Definition 3.2 for the system $(g_1,\mathscr{M},\mu,\mathbb{B})$ and for the family of targets given by Ω_ρ and the sequence ρ_M. The multiple exponential mixing condition $(\mathrm{EM})_r$ follows from [20, Theorem 1.1], Remark A.1 and Theorem A.2. The approximation condition (Appr) can be checked as follows:

CLAIM. There exists $\sigma>0$ such that, for every $\rho>0$ sufficiently small, there exist $A_\rho^-,A_\rho^+\in\mathrm{Lip}(\mathscr{M})$ such that

(i) $\|A_\rho^\pm\|_\infty\le 2$ and $\|A_\rho^\pm\|_{\mathrm{Lip}}\le\rho^{-\sigma}$;
(ii) $A_\rho^-\le 1_{\Omega_\rho}\le A_\rho^+$;
(iii) $\mu(A_\rho^+)-\mu(A_\rho^-)\le\rho^2$

Clearly the claim implies (Appr) since $\mu(\Omega_\rho)=\mathscr{O}(\rho)$.

Proof of the claim. Recall that $\Phi_\rho=\mathscr{S}(\phi_\rho)$, where ϕ_ρ is the indicator of the set $E_\rho=\left\{(x,y)\in\mathbb{R}^d\times\mathbb{R}\mid x_1>0,|x|\in[1,2],|x|^d|y|\in[0,\rho]\right\}$. We will construct A_ρ^+ that satisfies (i), (ii) and

$\overline{\text{(iii)}}$ $\mu(A_\rho^+)-\mu(1_{\Omega_\rho})\le\rho^2$.

The construction of A_ρ^- is similar.

Pick $f^+\in\mathrm{Lip}(\mathbb{R}^{d+1},[0,2])$ such that for some $\sigma>0$

- $\|f^+\|_{\mathrm{Lip}}\le\rho^{-\sigma}$,
- For $z\in E_\rho$, $f^+(z)=1$,
- For $z\notin E_{\rho+\rho^{10}}$, $f^+(z)=0$.

As the consequence $\mathscr{S}(f^+)\in\mathrm{Lip}(\mathscr{M})$ and $\Phi_\rho\le\mathscr{S}(f^+)$, and using Rogers identity of Lemma 9.13(a) (applied to the Siegel transform of the characteristic function of the set $E_{\rho+\rho^{10}}-E_\rho$) we get for ρ sufficiently small an open set $\mathscr{E}_\rho\subset\mathscr{M}$ such that $\mu(\mathscr{E}_\rho)\le\rho^3$

$(P1)$ For $\mathscr{L} \notin \mathscr{E}_\rho$, if $\mathscr{S}(f^+) > 0$, then $\Phi_\rho \ge 1$.
$(P2)$ If $\mathscr{M}_\rho := \{\mathscr{L} : \mathscr{S}(f^+) < 2\}$, then $\|\mathscr{S}(f^+)\|_{\mathrm{Lip}(\mathscr{M}_\rho)} \le \rho^{-\sigma-1}$.

Let now $u : \mathbb{R} \to [0,1]$ be an increasing C^∞ function such that $u(x) = 0$ for $x \le 0$ and $u(x) = 1$ for $x \ge 1$.

Finally, introduce $A^+_\rho : \mathscr{M} \to \mathbb{R}$ such that for $\mathscr{L} \in \mathscr{M}$

$$A^+_\rho(\mathscr{L}) = u\big(\mathscr{S}(f^+)(\mathscr{L})\big).$$

We now check that A^+_ρ satisfies the requirements of the claim.

Since $u \in C^\infty(\mathscr{M}, [0,1])$ we get that $A^+_\rho \in \mathrm{Lip}(\mathscr{M})$ and $\|A^+_\rho\|_\infty \le 2$. To prove the Lipschitz bound, observe that for $\mathscr{L} \notin \mathscr{M}_\rho$ we have that $A^+_\rho(\mathscr{L}) = 1$, while for $\mathscr{L} \in M_\rho$ we have $(P2)$. Hence $\|A^+_\rho\|_{\mathrm{Lip}} \le \rho^{-2\sigma}$. This proves (i) of the claim. To see (ii), just observe that

$$\Phi_\rho(\mathscr{L}) \ge 1 \implies \mathscr{S}(f^+)(\mathscr{L}) \ge 1 \implies A^+_\rho(\mathscr{L}) = 1.$$

We turn to $\overline{\text{(iii)}}$. If $\mathscr{L} \notin \mathscr{E}_\rho$, then by $(P1)$

$$A^+_\rho(\mathscr{L}) > 0 \implies \mathscr{S}(f^+)(\mathscr{L}) > 0 \implies \Phi_\rho(\mathscr{L}) \ge 1 \implies A^+_\rho(\mathscr{L}) = 1.$$

Since $\mu(\mathscr{E}_\rho) \le \rho^3$ and $\|A^+_\rho\|_\infty \le 2$, we get

$$\mu(A^+_\rho) - \mu(1_{\Omega_\rho}) \le \rho^2,$$

and $\overline{\text{(iii)}}$ is proved. □

Next we show how Rogers identity of Lemma 9.13(b) implies (Mov). Define

$$E^\tau_\nu = \left\{(x,y) \in \mathbb{R}^d \times \mathbb{R} \mid x_1 > 0, 2^{-\tau}|x| \in [1,2], |x|^d |y| \in [0,\nu]\right\}$$

and let ϕ^τ_ν be the indicator function of E^τ_ν. Then

$$\mu(\Omega_\rho \cap g_{-\tau}\Omega_\rho) \le \mathbb{E}(\Phi_\rho \Phi_\rho \circ g_\tau) = \int_{\mathscr{M}} \sum_{e_2 \ne \pm e_1 \in \mathscr{L} \text{ primitive}} \phi_\rho(e_1)\phi^\tau_\rho(e_2)\mathrm{d}\mu(\mathscr{L}),$$

where the contribution of $e_2 = -e_1$ vanishes because the contribution of any pair (e_1, e_2) where not both $e_{1,1}$ and $e_{2,1}$ are positive is zero. Applying Lemma 9.13(b) we get

$$\mu(\Omega_\rho \cap g_{-\tau}\Omega_\rho) \le C\mu(\Omega_\rho)^2$$

which is stronger than the required (Mov).

Finally, (Poly) holds for the sequence $\rho_M = \frac{c}{M(\ln M)^s}$ due to Lemma 9.13(a). □

9.7. **The argument in the inhomogeneous case.** The proof of Theorem 9.5 is very similar to that of Theorem 9.2, and below we only outline the main differences.

Let $\widetilde{\mathscr{M}}$ be the space of $(d+1)$-dimensional unimodular affine lattices. We identify $\widetilde{\mathscr{M}}$ with $SL_{d+1}(\mathbb{R}) \ltimes \mathbb{R}^{d+1} / SL_{d+1}(\mathbb{Z}) \ltimes \mathbb{Z}^{d+1}$, where the multiplication in $SL_{d+1}(\mathbb{R}) \ltimes \mathbb{R}^{d+1}$ is defined as $(A,a)(B,b) = (AB, a + Ab)$. We denote by $\tilde{\mu}$ the Haar measure on $\widetilde{\mathscr{M}}$.

For $\alpha \in \mathbb{R}^d$ and $z \in \mathbb{R}$, we define

$$\Lambda_{\alpha,z} = (\Lambda_\alpha, (0, \dots, 0, z)). \tag{9.6}$$

For $a > 0$, let $\tilde{\phi}_a$ be the indicator of the set[14]

$$\tilde{E}_a := \left\{(x, y) \in \mathbb{R}^d \times \mathbb{R} \mid |x| \in [1,2], |x|^d |y| \in [0, a]\right\}.$$

Fix $s \geq 0, c > 0$. For $M \in \mathbb{N}^*$, define

$$\nu := \frac{c}{M(\ln M)^s}, \quad \tilde{\Phi}_\nu := \tilde{\mathscr{S}}(\phi_\nu),$$

where the Siegel transforms in this affine setting are defined as follows for f, f_1, f_2 piecewise smooth functions with compact support on $\mathbb{R}^{d+1}$:

$$\tilde{S}(f)(\tilde{\mathscr{L}}) = \sum_{e \in \tilde{\mathscr{L}}} f(e), \quad \tilde{S}(f_1, f_2)(\tilde{\mathscr{L}}) = \sum_{e_1 \neq e_2 \in \tilde{\mathscr{L}}} f_1(e_1) f_2(e_2). \tag{9.7}$$

Note that, unlike our definition of the Siegel transform in the case of regular lattices, we do not require in the affine setting that the vectors e in the summation be primitive. On one hand, the notion of primitive vectors is not defined for affine lattices since the origin is not fixed. On the other hand, unlike the homogeneous setting, the multiples of one solution of the inhomogeneous inequality, does not satisfy a similar inequality, (see Remark 9.6). For $t \geq 0$, we then define

$$\tilde{A}_t(M) := \{(\alpha, z) \in \mathbb{R}^d \times \mathbb{R} : \tilde{\Phi}_\nu(g_t \Lambda_{\alpha,z}) \geq 1\}.$$

It is readily checked that $(\alpha, z) \in \tilde{A}_t(M)$ if and only if there exists $k = (k_1, \ldots, k_d)$ such that $2^t < |k| \leq 2^{t+1}$ and that

$$\exists m, \quad |k|^d |z + \langle k, \alpha \rangle + m| \leq \frac{c}{M(\ln M)^s}.$$

If α is such that $\tilde{\Phi}_\nu(g_t \Lambda_{\alpha,z}) \leq 1$ for every $t \in \mathbb{N}$, then we get that (α, z) is (r, s)-approximable if and only if there exist infinitely many Ms for which there exist $0 < t_1 < t_2 < \ldots < t_r \leq M$ satisfying $(\alpha, z) \in \bigcap_{j=1}^{r} \tilde{A}_{t_j}(M)$.

But in general, for α and $t \leq M$ such that $(\alpha, z) \in \tilde{A}_t(M)$, there may be multiple solutions k such that $2^t < |k| \leq 2^{t+1}$ for the same t. As in the case of Theorem 9.2 we have to deal with this issue.

The following proposition shows that almost surely on (α, z), multiple solutions do not occur. Its proof is based on Rogers identity for the second moment of the Siegel transforms in the affine lattices setting that we will recall in Section 9.9.

PROPOSITION 9.15. *For almost every $(\alpha, z) \in \mathbb{R}^d \times \mathbb{R}$, we have that for every M sufficiently large, for every $t \in [0, M]$, it holds that $\tilde{\Phi}_\nu(g_t \Lambda_{\alpha,z}) \leq 1$.*

Hence, Theorem 9.5 is equivalent to the following.

[14] Note that we do not ask in this affine setting that $x_1 > 0$ in the definition of $\tilde{E}_a$ since the symmetric contributions of $-k$ for every $k \in \mathbb{Z}^d$ that contributes to the Diophantine approximation counting problem in the homogeneous case of Theorem 9.2 do not appear in the inhomogeneous Diophantine approximation problem of Theorem 9.5.

THEOREM 9.16. *If $rs \le 1$, then for almost every $(\alpha, z) \in \mathbb{T}^d \times \mathbb{T}$, there exists infinitely many M for which there exist $0 < t_1 < t_2 < \ldots < t_r \le M$ satisfying*

$$\alpha \in \bigcap_{j=1}^{r} \tilde{A}_{t_j}(M).$$

If $rs > 1$, then for almost every $(\alpha, z) \in \mathbb{T}^d \times \mathbb{T}$, there exists at most finitely many M for which there exist $0 < t_1 < t_2 < \ldots < t_r \le M$ satisfying

$$\alpha \in \bigcap_{j=1}^{r} \tilde{A}_{t_j}(M).$$

9.8. **Modifying the initial distribution: inhomogeneous case.** Since the horocycle directions of $\Lambda_{\alpha,z}$, $(\alpha, z) \in \mathbb{T}^d \times \mathbb{T}$ account for all the strong unstable direction of the diagonal flow g_t acting on $\widetilde{\mathcal{M}}$, we can transform the requirement of Proposition 9.15 and Theorem 9.16 into a problem of multiple recurrence of the diagonal action g_t when applied to a random lattice in $\widetilde{\mathcal{M}}$.

We define

$$\tilde{B}_t(M) := \{\tilde{\mathcal{L}} \in \widetilde{\mathcal{M}} : \tilde{\Phi}_\nu(g_t\tilde{\mathcal{L}}) \ge 1\}.$$

Our goal becomes to prove the following.

PROPOSITION 9.17. *For $\tilde{\mu}$-almost every $\tilde{\mathcal{L}} \in \widetilde{\mathcal{M}}$, we have that for every M sufficiently large, for every $t \in [0, M]$, it holds that $\tilde{\Phi}_\nu(g_t\tilde{\mathcal{L}}) \le 1$.*

THEOREM 9.18. *If $rs \le 1$, then for $\tilde{\mu}$-almost every $\tilde{\mathcal{L}} \in \widetilde{\mathcal{M}}$, there exists infinitely many M for which there exist $0 < t_1 < t_2 < \ldots < t_r \le M$ satisfying*

$$\tilde{\mathcal{L}} \in \bigcap_{j=1}^{r} \tilde{B}_{t_j}(M).$$

If $rs > 1$, then for $\tilde{\mu}$-almost every $\tilde{\mathcal{L}} \in \widetilde{\mathcal{M}}$, there exists at most finitely many M for which there exist $0 < t_1 < t_2 < \ldots < t_r \le M$ satisfying

$$\tilde{\mathcal{L}} \in \bigcap_{j=1}^{r} \tilde{B}_{t_j}(M).$$

9.9. **Proofs of Proposition 9.17 and Theorem 9.18.** Again, the proofs of Proposition 9.17 and Theorem 9.18 are very similar to the proofs of their counterpart in the homogeneous case, Proposition 9.11 and Theorem 9.12.

Similarly to the homogeneous case, we want to apply Corollary 3.8. For the system (f, X, μ) we take $(g_1, \widetilde{\mathcal{M}}, \tilde{\mu})$, where $\tilde{\mu}$ is the Haar measure on $\widetilde{\mathcal{M}}$. For the targets, we take $\Omega_\rho = \{\tilde{\mathcal{L}} : \tilde{\Phi}_\rho(\tilde{\mathcal{L}}) \ge 1\}$. Observe that from the invariance of the Haar measure by g_t we have that $\tilde{\mu}(\Omega_\rho^t) = \tilde{\mu}(\Omega_\rho)$ for any t.

The only difference in the proof of Proposition 9.17 and Theorem 9.18 compared to that of Proposition 9.11 and Theorem 9.12, is in the application of Rogers identities to prove Proposition 9.17 as well as in the proof of (Mov) that is part of the proof of Theorem 9.18.

Before explaining the differences and concluding the proofs we recall Rogers identities for affine lattices. Recall the definition of Siegel transforms in the affine setting given in (9.7). The following can be found in [8, Lemma 4] (see also [58, Appendix 2]):

$$\mathbb{E}(\tilde{S}(f)) = \int_{\mathbb{R}^{d+1}} f(u)\mathrm{d}u \tag{9.8}$$

$$\mathbb{E}(\tilde{S}(f_1, f_2)) = \int_{\mathbb{R}^{d+1}} f_1(x)\mathrm{d}x \int_{\mathbb{R}^{d+1}} f_2(x)\mathrm{d}x. \tag{9.9}$$

(9.9) implies that

$$\begin{aligned}\tilde{\mu}\left(\tilde{\Phi}_\nu(\tilde{\mathscr{L}}) > 1\right) &\le \mathbb{E}\left(\tilde{\Phi}^2_\nu(\tilde{\mathscr{L}}) - \tilde{\Phi}_\nu(\tilde{\mathscr{L}})\right) \\ &= \mathbb{E}(\tilde{S}(\tilde{\phi}_\nu, \tilde{\phi}_\nu)) = \left(\int_{\mathbb{R}^{d+1}} \tilde{\phi}_\nu(u)\mathrm{d}u\right)^2 \le \frac{C}{M^2(\ln M)^{2s}}.\end{aligned}$$

Proposition 9.17 then follows by a Borel–Cantelli argument exactly as in the regular lattices case.

For the proof of (Mov) in the affine case we write for $\tau \ge 1$

$$\tilde{E}^\tau_\nu = \left\{(x,y) \in \mathbb{R}^d \times \mathbb{R} \mid 2^{-\tau}|x| \in [1,2], |x|^d|y| \in [0,\nu]\right\}$$

and denote by $\tilde{\phi}^\tau_\nu$ the indicator function of $\tilde{E}^\tau_\nu$. We then have

$$\tilde{\mu}(\Omega_\rho \cap g_{-\tau}\Omega_\rho) \le \mathbb{E}\left(\tilde{\Phi}_\rho\left(\tilde{\Phi}_\rho \circ g_\tau\right)\right) = \mathbb{E}(\tilde{S}(\tilde{\phi}_\nu, \tilde{\phi}^\tau_\nu)).$$

Next, (9.9) implies

$$\mathbb{E}(\tilde{S}(\tilde{\phi}_\nu, \tilde{\phi}^\tau_\nu)) = \int_{\mathbb{R}^{d+1}} \tilde{\phi}_\rho(u)\mathrm{d}u \int_{\mathbb{R}^{d+1}} \tilde{\phi}^\tau_\rho(u)\mathrm{d}u = \left(\int_{\mathbb{R}^{d+1}} \tilde{\phi}_\rho(u)\mathrm{d}u\right)^2 \le C\tilde{\mu}(\Omega_\rho)^2.$$

which is stronger than the required (Mov). □

9.10. **Multiple recurrence for toral translations.**

Proof of Theorem 4.7. PROOF OF PART (a). We begin with several reductions. Let $z = x - y$. Then $d(x, y + k\alpha) = d(z, k\alpha)$. Accordingly denoting $\hat{d}^{(r)}_n(z,\alpha)$ to be the r-th smallest among $\{d(z,k\alpha)\}_{k=0}^{n-1}$ we need to show that for almost every $(z,\alpha) \in (\mathbb{T}^d)^2$ we have

$$\limsup_{n\to\infty} \frac{|\ln \hat{d}^{(1)}_n(z,\alpha)| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{d}, \tag{9.10}$$

$$\limsup_{n\to\infty} \frac{|\ln \hat{d}^{(r)}_n(z,\alpha)| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{2d} \quad \text{for } r \ge 2. \tag{9.11}$$

Next we claim that it suffices to prove (9.11) only for $r = 2$. Indeed, since $\hat{d}^{(r)}_n$ is non decreasing in r, (9.11) with $r = 2$ implies that for $r > 2$,

$$\limsup_{n\to\infty} \frac{|\ln \hat{d}^{(r)}_n(z,\alpha)| - \frac{1}{d}\ln n}{\ln\ln n} \le \frac{1}{2d}.$$

To get the upper bound, suppose that $\hat{d}^{(2)}_n(z,\alpha) \le \varepsilon$. Then there are $0 \le k_1 < k_2 < n$ such that $k_j\alpha \in B(z,\varepsilon)$. Let $k = k_2 - k_1$. Then

$$(k_1 + sk)\alpha \in B(z, (1+2s)\varepsilon)$$

for $s = 0, \cdots, r-2$. Thus $\hat{d}^{(r)}_{(r-1)n}(z,\alpha) \le (2r-1)\hat{d}^{(2)}_n(z,\alpha)$. Taking limit superior, we obtain that if (9.11) holds for $r = 2$ then it holds for arbitrary r. In summary, we only need to show (9.10) and

$$\limsup_{n\to\infty} \frac{|\ln \hat{d}^{(2)}_n(z,\alpha)| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{2d}. \tag{9.12}$$

The proofs of (9.10) and (9.12) are similar to but easier than the proof of Theorem 9.5 so we only explain the changes. First, it suffices to take limit superior, for n of the form 2^M since for $2^{M-1} \le n \le 2^M$ we have

$$\hat{d}^{(r)}_{2^M}(z,\alpha) \le \hat{d}^{(r)}_n(z,\alpha) \le \hat{d}^{(r)}_{2^{M-1}}(z,\alpha).$$

Define for $\nu > 0$,

$$\hat{E}_\nu = \{e = (e', e'') \in \mathbb{R}^d \times \mathbb{R} : ||e'|| \le \nu, e'' \in (0,1]\}, \tag{9.13}$$

and denote by $\tilde{\phi}_\nu$ the indicator function of $\tilde{E}_\nu$. Let $\tilde{\Phi}_\nu = \tilde{S}(\tilde{\phi}_\nu)$. A direct inspection shows that

$$\hat{d}^{(r)}_{2^M}(z,\alpha) \le \nu_M \Leftrightarrow \tilde{S}(\tilde{\phi}_\nu)(\hat{g}_M \hat{\Lambda}_{\alpha,z}) \ge r,$$

where $\tilde{S}$ is defined by (9.7), $\hat{g}_M = g_{-M/d}$ for g given by (9.2), and $\hat{\Lambda}_{\alpha,z}$ is defined by $\hat{\Lambda}_{\alpha,z} = (\hat{\Lambda}_\alpha, (z,0))$ for

$$\hat{\Lambda}_\alpha = \begin{pmatrix} \mathrm{Id}_d & \alpha \\ 0 & 1 \end{pmatrix}.$$

Recall that $\widetilde{\mathscr{M}}$ denotes the space of $(d+1)$-dimensional unimodular affine lattices and $\tilde{\mu}$ the Haar measure on $\widetilde{\mathscr{M}}$. As in the proof of Theorem 9.5 one can show that $\tilde{\Phi}_\nu(\hat{g}_M \hat{\Lambda}_{\alpha,z}) \ge r$ infinitely often for almost every (z,α) if and only if $\tilde{\Phi}_\nu(\hat{g}_M \tilde{\mathscr{L}}) \ge r$ infinitely often for almost every $\tilde{\mathscr{L}} \in \widetilde{\mathscr{M}}$. Thus, to prove (9.10) and (9.12), we need to show that for almost every $\tilde{\mathscr{L}} \in \widetilde{\mathscr{M}}$ and for $\nu_M = M^{-s}$, $s > 0$ that

$$\tilde{\Phi}_{\nu_M}(\hat{g}_M \tilde{\mathscr{L}}) \ge 1 \text{ infinitely often if } s \le \frac{1}{d}, \tag{9.14}$$

$$\tilde{\Phi}_{\nu_M}(\hat{g}_M \tilde{\mathscr{L}}) \ge 1 \text{ finitely often if } s > \frac{1}{d}, \tag{9.15}$$

$$\tilde{\Phi}_{\nu_M}(\hat{g}_M \tilde{\mathscr{L}}) \ge 2 \text{ infinitely often if } s \le \frac{1}{2d}, \tag{9.16}$$

$$\tilde{\Phi}_{\nu_M}(\hat{g}_M \tilde{\mathscr{L}}) \ge 2 \text{ finitely often if } s > \frac{1}{2d}. \tag{9.17}$$

To prove (9.14)–(9.17), we need the following fact:

LEMMA 9.19.

(a) $\tilde{\mu}\left(\tilde{\Phi}_\nu = 1\right) = c_d \nu^d (1 + \mathscr{O}(\nu^{2d}))$,
(b) $c' \nu^{2d} \le \tilde{\mu}\left(\tilde{\Phi}_\nu \ge 2\right) \le c'' \nu^{2d}$.

Before we prove the lemma, we see how it allows to obtain (9.14)–(9.17) and finish the proof of Theorem 4.7(a).

Indeed, taking $\nu = M^{-s}$, Lemma 9.19 shows that

$$\sum_M \tilde{\mu}\left(\tilde{\Phi}_{\nu_M}) = 1\right) = \infty \iff s \le \frac{1}{d}, \quad \sum_M \tilde{\mu}\left(\tilde{\Phi}_{\nu_M}) \ge 2\right) = \infty \iff s \le \frac{1}{2d}.$$

From there, (9.14)–(9.17) follow from the the classical Borel–Cantelli Lemma, that is, from the case $r = 1$ in our Theorem 2.6.[15] For this, observe that the verification of the conditions of Definition 3.1 and Definition 3.2 for the targets

$$\Omega_\rho = \{\tilde{\mathscr{L}} : \tilde{\Phi}_\rho(\tilde{\mathscr{L}}) \ge 1\} \quad \text{and} \quad \hat{\Omega}_\rho = \{\tilde{\mathscr{L}} : \tilde{\Phi}_\rho(\tilde{\mathscr{L}}) \ge 2\}$$

is very similar to the proof of Theorem 9.5 so we omit it.

Proof of Lemma 9.19. We get by Rogers identities (9.8) and (9.9) that

$$\mathbb{E}(\tilde{\Phi}_\nu) = c_d \nu^d, \quad \mathbb{E}(\tilde{\Phi}_\nu^2 - \tilde{\Phi}_\nu) = \left(c_d \nu^d\right)^2.$$

It follows that

$$\tilde{\mu}(\tilde{\Phi}_\nu \ge 2) \le \mathbb{E}(\tilde{\Phi}_\nu^2 - \tilde{\Phi}_\nu)/2 \le C\nu^{2d}$$

proving the upper bound of part (b).

In addition

$$\mathbb{E}\left(\tilde{\Phi}_\nu \mathbf{1}_{\tilde{\Phi}_\nu \ge 2}\right) \le \left(c_d \nu^d\right)^2$$

so that

$$\tilde{\mu}(\tilde{\Phi}_\nu = 1) = \mathbb{E}(\tilde{\Phi}_\nu) - \mathbb{E}(\tilde{\Phi}_\nu \mathbf{1}_{\tilde{\Phi}_\nu \ge 2}) = c_d \nu^d + \mathscr{O}\left(\nu^{2d}\right). \tag{9.18}$$

This proves part (a).

To prove the lower bound in part (b) we need the following estimate. Denote by $\mathscr{L}_{primitive}$ the set of primitive vectors in $\mathscr{L}$ for $\mathscr{L} \in \mathscr{M} = SL_{d+1}(\mathbb{R})/SL_{d+1}(\mathbb{Z})$. Let

$$\bar{E}_1 = \left\{(e', e'') \in \mathbb{R}^d \times \mathbb{R} : |e'| \in \left[\frac{\nu}{10}, \frac{\nu}{5}\right], |e''| \le \frac{1}{10}\right\},$$
$$\bar{E}_2 = \left\{(e', e'') \in \mathbb{R}^d \times \mathbb{R} : |e'| \le \frac{\nu}{5}, |e''| \le \frac{1}{10}\right\},$$
$$\mathscr{A} = \left\{\mathscr{L} \in \mathscr{M} : \mathrm{Card}\left(\mathscr{L}_{primitive} \cap \bar{E}_1\right) = \mathrm{Card}\left(\mathscr{L}_{primitive} \cap \bar{E}_2\right) = 1\right\}.$$

CLAIM. *We have*

$$\mu(\mathscr{A}) = c\nu^d + \mathscr{O}\left(\nu^{2d}\right). \tag{9.19}$$

Assume the claim holds. For $\mathscr{L} \in \mathscr{A}$, the fundamental domain of $\mathbb{R}^{d+1}/\mathscr{L}$ can be chosen to contain

$$\bar{E}_3 = \left\{(e', e'') \in \mathbb{R}^d \times \mathbb{R} : |e'| \le \frac{\nu}{100}, |e''| \le \frac{1}{100}\right\}.$$

[15] We note that in case $r = 1$ Theorem 2.6 is a minor variation of standard dynamical Borel–Cantelli Lemmas such as the Borel–Cantelli Lemma of [114].

Note that if $\mathscr{L} \in \mathscr{A}$ then $\mathscr{L} \cap \bar{E}_1$ contains a non-zero vector w and hence if $\tilde{z} \in \bar{E}_3$ then $(\tilde{z}+\mathscr{L}) \cap E_v$ contains at least two vectors: $\tilde{z}$ and $\tilde{z}+w$. We thus have

$$\mu\left((\mathscr{L}+\tilde{z}) : \operatorname{Card}\left((\mathscr{L}+\tilde{z}) \cap \hat{E}_v\right) \geq 2\right) \geq \mu(\mathscr{A})\, \mu\left(\operatorname{Card}\left((\mathscr{L}+\tilde{z}) \cap \hat{E}_v\right) \geq 2 | \mathscr{A}\right)$$
$$\geq \mu(\mathscr{A}) \mu(z \in \bar{E}_3) \geq c' v^{2d}.$$

This gives the lower bound in part (b) of Lemma 9.19. It remains to prove (9.19).

Proof of the claim. We consider the cases $d > 1$ and $d = 1$ separately.

In case $d > 1$, denote $\Psi_j = \tilde{S}(\mathbf{1}_{\bar{E}_j})$ for $j = 1, 2$. By Rogers identities,

$$\mathbb{E}(\Psi_1) = \frac{1}{10} c_d v^d, \quad \mathbb{E}\left(\Psi_1^2 - \Psi_1\right) = \left(\frac{1}{10} c_d v^d\right)^2.$$

Thus arguing as in the proof of (9.18) we conclude that

$$\mu(\Psi_1 = 1) = \frac{1}{10} c_d v^d + \mathscr{O}\left(v^{2d}\right). \tag{9.20}$$

Rogers identities also give

$$\mathbb{E}(\Psi_1(\Psi_2 - \Psi_1)) = \mathscr{O}\left(v^{2d}\right).$$

Hence

(9.21)
$$\mu\left(\operatorname{Card}(\mathscr{L}_{primitive} \cap \bar{E}_1) \geq 1 \text{ and } \operatorname{Card}\left(\mathscr{L}_{primitive} \cap \left(\bar{E}_2 \smallsetminus \bar{E}_1\right)\right) \geq 1\right) = \mathscr{O}\left(v^{2d}\right).$$

Combining (9.20) and (9.21) we obtain (9.19) for $d > 1$.

In case $d = 1$ we still have $\mathbb{E}(\Psi_1) = cv + \mathscr{O}(v^2)$. On the other hand, for $d = 1$ we have $\operatorname{Card}\left(\mathscr{L}_{primitive} \cap \bar{E}_2\right) \leq 1$ since $\mathscr{L}$ is unimodular. Thus

$$\mathbb{E}(\Psi_1) = \mu(\Psi_1 = 1) = \mu(\Psi_1 = 1 \text{ and } \Psi_2 - \Psi_1 = 0) = cv. \qquad \square$$

This completes the proof of Lemma 9.19 and thus of Theorem 4.7(a). $\square$

PROOF OF PART (*b*). It is clear that for any r, if $\bar{\mathscr{E}}_r$ is not empty then it is equal to M. The proof that $\bar{\mathscr{E}}_1 = M$ implies that $\bar{\mathscr{E}}_r = M$ for all r is exactly similar to the implication of (9.11) from (9.12), so we just focus on showing that $\bar{\mathscr{E}}_1 = M$. Adapting the beginning of the proof of part (*a*) to the current homogeneous setting, we see that the proof boils down to showing that for almost every $\mathscr{L} \in \mathscr{M}$, and for $v_M = M^{-s}$, $s > 0$, it holds that

$$\mathscr{S}(\mathbf{1}_{E_{v_n}})(\hat{g}_n \mathscr{L}) \geq 1 \text{ infinitely often if } s \leq \frac{1}{d}, \tag{9.22}$$

$$\mathscr{S}(\mathbf{1}_{E_{v_n}})(\hat{g}_n \mathscr{L}) \geq 1 \text{ finitely often if } s > \frac{1}{d}, \tag{9.23}$$

where E_v is as in (9.13), and $\mathscr{S}$ designates the Siegel transform as in (9.3). By Rogers identity of Lemma 9.13(a) we have $\mathbb{E}\big(\mathscr{S}(\mathbf{1}_{E_{v_M}})\big) = cM^{-sd}$. Hence (9.22) and (9.23) follow by classical Borel–Cantelli Lemma (see for example the Borel–Cantelli Lemma of [114]) or by the case $r = 1$ of our Theorem 2.6.

This completes the proof of Theorem 4.7. $\square$

9.11. **Notes.** A classical Khintchine–Groshev Theorem is given by (1.4)–(1.5). A lot of interest is devoted to extending this result to α lying in a submanifold of $\mathbb{R}^d$ (see, e.g., [14, 19]). In particular, [113] discusses Khintchine–Groshev type results on manifolds using dynamical tools. Surveys on applications of dynamics to other problems in metric Diophantine approximations include [17, 21, 50, 57, 59, 73, 110, 116, 127]. These applications are based on Dani correspondence [43].

The use of Siegel transform as a convenient analytic tool for applying Dani correspondence can be found in [126]. Limit Theorems for Siegel transforms and applications to number theory are discussed in [9, 12, 23, 51, 52].

The Diophantine results presented in our paper require the rotation angle to be random. Much less is known in the case for fixed rotation angle, but some results are available for the circle. Namely, consider the rotation T_α by the irrational angle α. Suppose η to be the Diophantine type of the rotation, that is,

$$\eta = \sup\left\{\beta : \liminf_{j\to\infty} j^\beta \| j\alpha\| = 0\right\}.$$

For $\rho > 0$, let $\tau_\rho(x, y)$ the first time of the trajectory $\{T_\alpha^k(y)\}_{k\geq 0}$ visiting ρ-neighborhood of x. [38] proved that for almost every $x \in \mathbb{T}$,

$$\liminf_{\rho\to 0} \frac{\ln \tau_\rho(x,x)}{|\ln \rho|} = \eta^{-1}, \qquad \limsup_{\rho\to 0} \frac{\ln \tau_\rho(x,x)}{|\ln \rho|} = 1,$$

and [109] proved that for almost every x and y,

$$\liminf_{\rho\to 0} \frac{\ln \tau_\rho(x,y)}{|\ln \rho|} = 1, \qquad \limsup_{\rho\to 0} \frac{\ln \tau_\rho(x,y)}{|\ln \rho|} = \eta.$$

More general Diophantine approximation results valid for a fixed α and almost all x could be found in [60, 106, 108, 120, 147] and references therein.

10. EXTREME VALUES

10.1. **From hitting times to extreme values.** Here we describe applications of our results to extreme value theory.

Let (f, M, μ) be as in Definition 3.1. Recall that the sets $\mathcal{G}_r$ and $\mathcal{H}$ are introduced in Definition 4.3 and Definition 4.5 respectively. Recall also that under the conditions of Theorem 4.4 and Theorem 4.6 $\mu(\mathcal{G}_r) = 1$ and $\mathcal{H}$ contains a residual set.

Given a C^2 function φ and a point $y \in M$, let $\varphi_n^{(r)}(y)$ be the r-th minimum among the values $\{\varphi(f^j y)\}_{j=1}^n$.

THEOREM 10.1.

(a) Suppose f is $(2r+1)$-fold exponentially mixing preserving a smooth measure μ. Then

(i) *There is a set $\mathscr{G}$ of full measure in M such that if φ is a C^2 function with a unique non degenerate minimum at $x \in \mathscr{G}$, then for almost every $y \in M$,*

$$\limsup_{n\to\infty} \frac{\left|\ln\left(\varphi_n^{(r)}(y)-\varphi(x)\right)\right| - \frac{2}{d}\ln n}{\ln\ln n} = \frac{2}{rd}.$$

(ii) *If $\mathscr{G}_1 = M$ and the periodic orbits of f are dense, then there is a dense G_δ set $\mathscr{H} \subset M$, such that if φ is a C^2 function with a unique non degenerate minimum at $x \in \mathscr{H}$, then for almost every $y \in M$,*

$$\limsup_{n\to\infty} \frac{\left|\ln\left(\varphi_n^{(r)}(y)-\varphi(x)\right)\right| - \frac{2}{d}\ln n}{\ln\ln n} = \frac{2}{d}.$$

(b) *If f is an expanding map of $\mathbb{T}$ and μ is a non-conformal Gibbs measure of dimension $\mathbf{d}$, λ is the Lyapunov exponent of μ, then there is a set $\mathscr{G}_\mu$ with $\mu(\mathscr{G}_\mu) = 1$, such that if φ is a C^2 function with a unique non degenerate minimum at $x \in \mathscr{G}_\mu$, then for μ–almost every $y \in M$,*

$$\limsup_{n\to\infty} \frac{\left|\ln\left(\varphi_n^{(r)}(y)-\varphi(x)\right)\right| - \frac{2}{\mathbf{d}}\ln n}{\sqrt{2(\ln n)(\ln\ln\ln n)}} = \frac{2\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}},$$

where σ given by (6.5).

(c) *Part (a) remains valid for the geodesics flow on a compact $(d+1)$-dimensional manifold $\mathscr{Q}$ and C^2 function $\varphi : \mathscr{Q} \to \mathbb{R}$ which has unique non-degenerate minimum at some point on $\mathscr{Q}$. (In this case $\varphi_n^{(r)}(y)$ is the r-th local minimum of the map $t \mapsto \varphi(q(t))$ where $(q(t), v(t))$ is the geodesic starting at q with velocity v.)*

(d) *For toral translations we have that for almost all α and almost all y we have*

$$\limsup_{n\to\infty} \frac{\left|\ln\left(\varphi_n^{(r)}(y)-\varphi(x)\right)\right| - \frac{2}{d}\ln n}{\ln\ln n} = \begin{cases} \frac{2}{d} & \text{if } r=1, \\ \frac{1}{d} & \text{if } r \geq 2. \end{cases}$$

Proof. At a non-degenerate minimum x we have that for y close to x

$$K^{-1}d^2(x,y) \leq \varphi(y)-\varphi(x) \leq Kd^2(x,y) \tag{10.1}$$

so part (i) of (a) holds for $x \in \mathscr{G}_r$ and part (ii) of (a) holds for $x \in \mathscr{H}$ as defined in Theorems 4.4 and 4.6. Part (b) follows from Theorem 6.1. Part (c) follows from Theorem 8.1, and part (d) follows from Theorem 4.7. □

THEOREM 10.2. *Under the assumptions of Theorem 10.1(a) or Theorem 10.1(d) there is a set of points x of full measure such that if φ has a non-degenerate minimum at x then the process*

$$\frac{\varphi_n^{(1)}(y)-\varphi(x)}{\rho^2}, \frac{\varphi_n^{(2)}(y)-\varphi(x)}{\rho^2}, \dots, \frac{\varphi_n^{(r)}(y)-\varphi(x)}{\rho^2}, \dots$$

with $n = [\tau\rho^{-d}]$ converges as $\rho \to 0$ to the Poisson process on $\mathbb{R}^+$ with measure $\gamma(\varphi)\tau\frac{d}{2}t^{\frac{d}{2}-1}\mathrm{d}t$, where $\gamma(\varphi) > 0$ depends on x and φ.

Proof. Note that (10.1) does not provide enough information to deduce the result from (5.1) of Theorem 5.1. However, for any choice of $r_1^- < r_1^+ < r_2^- < r_2^+ < \cdots < r_s^- < r_s^+$, consider the targets

$$\Omega^{n,j} = \left\{ y : \varphi(y) - \varphi(x) \in \left[r_j^- \rho^2, r_j^+ \rho^2 \right] \right\}, \tag{10.2}$$

that satisfy

$$\lim_{\rho \to 0} \tau \rho^{-d} \mu(\Omega^{n,j}) = \tau \gamma(\varphi) \big((r_j^+)^{\frac{d}{2}} - (r_j^-)^{\frac{d}{2}} \big) = \tau \gamma(\varphi) \int_{r_j^-}^{r_j^+} \frac{d}{2} t^{\frac{d}{2}-1} \mathrm{d}t.$$

Conditions $\widetilde{(M1)}_r$ and $(M2)_r$ from §2.5 can easily be checked for the targets $\Omega^{n,j}$ using the results of Section 3. Since (Mov) for targets (10.2) follows from (Mov) for balls, only (Appr) needs to be checked but the latter follows immediately from Lemma 3.4. We can thus apply Theorem 2.14 and conclude the Poisson limit. □

Next, we consider functions of the form

$$\psi(y) = \frac{c}{[d(x,y)]^s} + \widetilde{\psi}(y), \quad \text{where } c < 0 \text{ and } \widetilde{\psi} \in \mathrm{Lip}(M). \tag{10.3}$$

THEOREM 10.3. *Let f be $(2r+1)$-fold exponentially mixing. Then*

(a) *There is a set $\mathcal{G}$ or full measure such that if ψ satisfies (10.3) with $x \in \mathcal{G}$ then for almost all y*

$$\limsup_{n\to\infty} \frac{\ln |\psi_n^{(r)}(y)| - \frac{s}{d} \ln n}{\ln \ln n} = \frac{s}{rd}.$$

(b) *There is a G_δ set $\mathcal{H}$ such that if ψ satisfies (10.3) with $x \in \mathcal{H}$ then for almost all y*

$$\limsup_{n\to\infty} \frac{\ln |\psi_n^{(r)}(y)| - \frac{s}{d} \ln n}{\ln \ln n} = \frac{s}{d}.$$

(c) *If $x \in \mathcal{G}$ then*

$$\frac{\rho^s \psi_n^{(1)}(y)}{c}, \frac{\rho^s \psi_n^{(2)}(y)}{c}, \ldots, \frac{\rho^s \psi_n^{(r)}(y)}{c}, \ldots \quad \textit{where} \quad n = \tau \rho^{-d}$$

converges as $\rho \to 0$ to the Poisson process on $\mathbb{R}^+$ with measure

$$\frac{d\tau\gamma(x)}{s} t^{-(d/s)-1} \mathrm{d}t.$$

The proofs of the above results are similar to the proofs of Theorems 10.1 and 10.2, so we will leave them to the readers.

The next result is an immediate consequence of Theorems 10.2 and 10.3(c).

COROLLARY 10.4.

(a) (FRÉCHET LAW FOR SMOOTH FUNCTIONS) *If f is $(2r+1)$-fold exponentially mixing, φ is a smooth function with non-degenerate minimum at some $x \in \mathcal{G}$ then there is $\sigma = \sigma(x)$ such that for each $t > 0$*

$$\lim_{n\to\infty} \mu(y : \varphi_n^{(1)}(y) > n^{-2/d} t) = e^{-\sigma t^{d/2}}.$$

(b) (WEIBULL LAW FOR UNBOUNDED FUNCTIONS) *If f is $(2r+1)$-fold exponentially mixing, φ is given by (10.3) with $x \in \mathcal{G}$ then there is $\sigma = \sigma(x)$ such that for each $t > 0$*

$$\lim_{n\to\infty} \mu(y : \varphi_n^{(1)}(y) > -n^{-s/d} t) = e^{-\sigma t^{-d/s}}.$$

10.2. **Notes.** A classical Fisher–Tippett–Gnedenko theorem says that for independent identically distributed random variables the only possible limit distributions of normalized extremes are the Gumbel distribution, the Fréchet distribution, or the Weibull distribution. Corollaries 7.3 and 10.4(a) and (b) provide typical examples where one can encounter each of these three types. We refer to [121] for the proof of Fisher–Tippett–Gnedenko theorem as well as for extensions of this theorem to weakly dependent random variables. The weak dependence conditions used in the book have a similar sprit to our conditions $(M1)_r$ and $(M2)_r$. More discussions about relations of extreme value theory to Poisson limit theorems in the context of dynamical systems can be found in [63, 66]. The book [125] discusses extreme value theory for dynamical systems and lists various applications. One application of extreme value theory, is that for non-integrable functions, such as described in Theorem 10.3 above, the growth of ergodic sums are dominated by extreme values, see [1, 28, 45, 101, 102, 130] and references therein.

APPENDIX A. MULTIPLE EXPONENTIAL MIXING

A.1. **Basic properties.** Let f be a smooth map of a compact manifold M preserving a smooth probability measure μ. In the dynamical system literature, for $r \geq 1$, f is called $(r+1)$-fold exponentially mixing if there are constants $s, \bar{C}$ and $\bar{\theta} < 1$ such that for any C^s functions $A_0, A_1, \dots, A_r$ for any r tuple $k_1 < k_2 < \cdots < k_r$

$$\left| \int \prod_{j=0}^{r} \left(A_j \circ f^{k_j} \right) \mathrm{d}\mu - \prod_{j=0}^{r} \int A_j \mathrm{d}\mu \right| \leq \bar{C}\bar{\theta}^m \prod_{j=0}^{r} \|A_j\|_{C^s}, \tag{A.1}$$

where $m = \min_j (k_j - k_{j-1})$ with $k_0 = 0$.

In this paper we need to consider a larger class of functions, namely we need that there are constants s, C and $\theta < 1$ such that for any $B \in C^s(M^{r+1})$ we have

$$\left| \int B(x_0, f^{k_1} x_0, \cdots, f^{k_r} x_0) \mathrm{d}\mu(x_0) - \mu^{r+1}(B) \right| \leq C_s \theta^m \|B\|_{C^s} \tag{A.2}$$

where μ^{r+1} is defined by (3.1).

In this section we show equivalence of (A.1) and (A.2). We use the following fact.

REMARK A.1. If (A.1) holds for some s then it holds for all s (with different $\bar{\theta}$). The same applies for (A.2).

Indeed suppose that (A.2) holds for some C^s functions. Pick some $\alpha < s$. We claim that it also holds for C^α functions. Indeed pick a small ε and approximate a C^α function B with $\|B\|_{C^\alpha} = 1$ by a C^s function $\bar{B}$, so that (assuming that m is large)

$$\|B - \bar{B}\|_{C^0} \le e^{-\varepsilon \alpha m}, \quad \|\bar{B}\|_{C^s} \le e^{\varepsilon s m}.$$

Then

$$\begin{aligned}\int B(x_0, f^{k_1}x_0, \cdots, f^{k_r}x_0)\mathrm{d}\mu(x_0) &= \int \bar{B}(x_0, f^{k_1}x_0, \cdots, f^{k_r}x_0)\mathrm{d}\mu(x_0) + O\left(e^{-\varepsilon\alpha m}\right)\\ &= \mu^{r+1}(\bar{B}) + O\left(e^{-\varepsilon\alpha m}\right) + O\left(\theta^m e^{\varepsilon s m}\right)\\ &= \mu^{r+1}(B) + O\left(e^{-\varepsilon\alpha m}\right) + O\left(\theta^m e^{\varepsilon s m}\right).\end{aligned}$$

and the second error term is exponentially small if ε is small enough. The argument for (A.1) is identical.

We now ready to show that (A.1) implies (A.2).

THEOREM A.2. *Suppose that* (A.1) *holds and s is sufficiently large. Then* (A.2) *holds.*

Proof of Theorem A.2. Since $B \in C^s(M^{r+1})$ it also belongs to Sobolev space $H^s(M^{r+1})$. Hence we can decompose

$$B = \sum_\lambda b_\lambda \phi_\lambda$$

where ϕ_λ are eigenfunctions of Laplacian on M^{r+1} with eigenvalues λ^2 and $\|\phi_\lambda\|_{L^2} = 1$. The eigenfunctions ϕ_λ are of the form

$$\phi_\lambda(x_0, x_1, \ldots, x_r) = \prod_{j=0}^r \psi_j(x_j)$$

where $\Delta_M \psi_j = \zeta_j^2 \psi_j$ and $\lambda^2 = \sum_j \zeta_j^2$. Recall that by Sobolev Embedding Theorem for compact manifolds, $H^s(M) \subset C^{s-\frac{d}{2}-1-\varepsilon}(M)$ for any $\varepsilon > 0$. Since $\|\psi_j\|_{H^s} = \zeta_j^s$, we have

$$\|\psi_j\|_{C^1} \le C_u \zeta_j^u \le C_u \lambda^u \quad \text{if} \quad u > 1 + \frac{d}{2}.$$

It follows from (A.1) that if $\phi \not\equiv 1$ then

$$\left|\int \phi_\lambda(x, f^{k_1}x, \ldots, f^{k_r}x)\mathrm{d}\mu(x) - \prod_{j=0}^r \int \psi_j \mathrm{d}\mu\right| \le C\lambda^{u(r+1)}\bar{\theta}^m.$$

Therefore

$$\begin{aligned}&\left|\int B(x, f^{k_1}x, \ldots, f^{k_r}x)\mathrm{d}\mu(x) - \int B(x_0, \cdots, x_r)\mathrm{d}\mu(x_0)\cdots\mathrm{d}\mu(x_r)\right|\\ &\qquad \le C\theta^m \sum_\lambda b_\lambda \lambda^{u(r+1)} \le C\theta^m \|B\|_{H^{u(r+1)}(M^{r+1})}.\end{aligned}$$

This proves the result if $s > \left(1+\frac{d}{2}\right)(r+1)$. □

A.2. **Mixing for Gibbs measures.**

Proof of Proposition 6.2. The proof consists of three steps.

Step 1. By the same argument as in [155, Proposition 3.8], we have that for $\hat{\psi}_1 \in \mathrm{Lip}(\mathbb{T})$, $\hat{\psi}_2 \in L^1(\mu)$,

$$\left|\int \hat{\psi}_1(\hat{\psi}_2\circ f^n)\mathrm{d}\mu - \int \hat{\psi}_1 \mathrm{d}\mu \int \hat{\psi}_2 \mathrm{d}\mu\right| \le C\|\hat{\psi}_1\|_{Lip}\|\hat{\psi}_2\|_{L^1}\bar{\theta}^n,\ n\ge 0. \tag{A.3}$$

Step 2. We proceed to show inductively that for each $r>0$ and $\psi_i \in \mathrm{Lip}(\mathbb{T})$ for $i=1,\dots,r$,

$$\left|\int\left(\prod_{i=1}^r \psi_i\circ f^{k_i}\right)\mathrm{d}\mu - \prod_{i=1}^r\int \psi_i \mathrm{d}\mu\right| \le C\bar{\theta}^m \prod_{i=1}^r ||\psi_i||_{Lip}, \tag{A.4}$$

where $m = \min_{1\le i\le r-1}(k_{i+1}-k_i)$, $k_0=0$.

By invariance of μ we may assume that $k_1=0$. Applying (A.3) with $\hat{\psi}_1=\psi_1$ and $\hat{\psi}_2 = \prod_{j=2}^r \psi_j\circ f^{k_j-k_2}$ we get

$$\begin{aligned}&\left|\int\left(\prod_{i=1}^r \psi_i\circ f^{k_i}\right)\mathrm{d}\mu - \left(\int\psi_1\mathrm{d}\mu\right)\left[\int\left(\prod_{i=2}^r \psi_i\circ f^{k_i}\right)\mathrm{d}\mu\right]\right|\\ &\qquad\le C\bar{\theta}^m\|\psi_1\|_{Lip}\left\|\left(\prod_{i=2}^r \psi_i\circ f^{k_i}\right)\right\|_{L_1}\le C\bar{\theta}^m\prod_{i=1}^r||\psi_i||_{Lip}.\end{aligned}$$

Applying inductive estimate to

$$\int\left(\prod_{i=2}^r \psi_i\circ f^{k_i}\right)\mathrm{d}\mu$$

we obtain (A.4).

Step 3. Applying the same argument as in the proof of Theorem A.2, we get $(\mathrm{EM})_r$. □

A.3. **Examples of exponentially mixing systems.** There are many results about double (=2-fold) exponential mixing. Many examples of those systems are partially hyperbolic. §3.3 describes the the main examples of smooth exponentially mixing systems. In particular, they expand an invariant foliation W^s by unstable manifolds. The next result allows to promote double mixing to r fold mixing.

THEOREM A.3. ([48, Theorem 2]) Suppose that for each subset D in a single unstable leaf of bounded geometry[16] and any Hölder probability density ρ on D we have

$$\left|\int_D A(f^n x)\rho(x)\mathrm{d}x - \int A\mathrm{d}\mu\right| \le C\theta^n\|A\|_{C^s}\|\rho\|_{C^\alpha}$$

for $A\in C^s$. Then f is r-fold exponentially mixing for all $r\ge 2$.

[16] We refer the reader to [48] for precise requirements on D since those requirements are not essential for the present discussion.

We also note the following fact.

THEOREM A.4. *A product of exponentially mixing maps is exponentially mixing.*

The proof of this theorem is very similar to the proof of Theorem A.2 so we leave it to the reader. We also note that instead of direct products one can also consider certain skew products such that both base and fibers are exponentially mixing and the skewing function satisfies suitable growth conditions. We refer the reader to [49, Section 4] for precise statements and to [53, Section 10] for examples of the skew products satisfying the conditions of [49].

Another source of exponential mixing is a spectral gap for transfer operators (cf. §A.2 as well as [133, 155]). This allows to handle non-uniformly hyperbolic systems admitting Young tower with exponential tails [159] as well as piecewise expanding maps [155].

We note that the maps described in the last paragraph do not fit in the framework of the present paper due to either lack of smoothness or lack of smooth invariant measure. It is interesting to extend the result of the paper to cover those systems as well as some slower mixing system and this is a promising direction for a future work.

APPENDIX B. GIBBS MEASURES FOR EXPANDING MAPS ON THE CIRCLE

B.1. **Some notation.** Recall that we assume $P(g) = 0$, so we have

$$\ln \mu(B_n(x,\varepsilon)) = \sum_{j=0}^{n-1} g(f^j x) + O(1). \tag{B.1}$$

Denote

$$r_n = \sup_{r>0}\{r \mid B(x,r) \subset B_n(x,\varepsilon)\}, \quad \bar{r}_n = \inf_{r>0}\{r \mid B(x,r) \supset B_n(x,\varepsilon)\}.$$

By bounded distortion property, there exist constants $C_0 > 0$ and $\alpha > 0$ such that if $d(f^n y, f^n x) < \varepsilon$ then

$$\left(C_0 \exp \varepsilon^\alpha\right)^{-1} \le \frac{|Df^n(y)|}{|Df^n(x)|} \le C_0 \exp \varepsilon^\alpha.$$

Recalling (6.2),

$$\exp\left[\left(\sum_{j=0}^{n-1} f_u(f^j x)\right) - \varepsilon^\alpha\right] \frac{d(x,y)}{C_0} \le d(f^n x, f^n y)$$
$$\le C_0 \exp\left[\left(\sum_{j=0}^{n-1} f_u(f^j x)\right) + \varepsilon^\alpha\right] d(x,y).$$

Hence

$$\varepsilon C_0^{-1} \exp\left[\left(-\sum_{j=0}^{n-1} f_u(f^j x)\right) - \varepsilon^\alpha\right] \le r_n \le \bar{r}_n$$
$$\le \varepsilon C_0 \exp\left[\left(-\sum_{j=0}^{n-1} f_u(f^j x)\right) + \varepsilon^\alpha\right].$$

It follows that

$$\text{(B.2)} \qquad \ln r_n = \sum_{j=0}^{n-1} -f_u(f^j x) + O(1), \quad \ln \bar{r}_n = \sum_{j=0}^{n-1} -f_u(f^j x) + O(1).$$

Next define

$$N(r) = \max(n : B(x,r) \subset B_n(x,\varepsilon)), \quad \overline{N}(r) = \min(n : B(x,r) \supset B_n(x,\varepsilon)).$$

Then, similarly to (B.2) we obtain

$$\text{(B.3)} \qquad \ln r = \sum_{j=0}^{N(r)-1} -f_u(f^j x) + O(1) = \sum_{j=0}^{\overline{N}(r)-1} -f_u(f^j x) + O(1).$$

B.2. **Proof of (6.8) and (6.9).** Note that

$$\text{(B.4)} \qquad \mu(B_{\overline{N}(r)}(x,\varepsilon)) \le \mu(B(x,r)) \le \mu(B_{N(r)}(x,\varepsilon)).$$

Since f is uniformly expanding, there is a positive constant C such that for each x, $1/C \le f_u(x) \le C$. Accordingly,

$$\text{(B.5)} \qquad \frac{N(r)}{C} \le |\ln r| \le CN(r), \quad \frac{\overline{N}(r)}{C} \le |\ln r| \le C\overline{N}(r).$$

On the other hand, since $P(g) = 0$, [133, Chapter 3] shows that there is a function a Hölder function $\hat{g}(x)$ such that $\hat{g} = g + h - h \circ f$ for a Hölder function h and moreover

$$\sum_{f(y)=x} e^{\hat{g}(y)} = 1.$$

In particular, $\hat{g}(y)$ is negative and, since it is continuous, there are constants $\hat{C}_1 > \hat{\varepsilon} > 0$ such that for any $x \in \mathbb{T}$ we have $\hat{g}(x) \in (-\hat{C}_1, -\hat{\varepsilon})$. Using the estimate

$$\sum_{n=0}^{N-1} g(f^n x) = \sum_{n=0}^{N-1} \hat{g}(f^n x) + O(1)$$

we conclude that for some constant $\hat{C}_2 > 0$ we have for every $x \in \mathbb{T}$,

$$\text{(B.6)} \qquad -\hat{C}_1 N - \hat{C}_2 \le \sum_{n=0}^{N-1} g(f^n x) \le -\hat{\varepsilon} N + \hat{C}_2$$

Combining (B.1), (B.4), (B.5), and (B.6) we obtain (6.8).

Next, (B.3) shows that $N(4r) - \overline{N}(r) = O(1)$. Then (6.9) follows from (B.1) and (B.4).

B.3. **Proof of Lemma 6.3(b).** Observe that (B.1) and (B.2) give

$$\ln \mu(B_n(x,\varepsilon)) - \mathbf{d} \ln r_n = \sum_{j=0}^{n-1} \psi(f^j x) + O(1),$$

$$\ln \mu(B_n(x,\varepsilon)) - \mathbf{d} \ln \bar{r}_n = \sum_{j=0}^{n-1} \psi(f^j x) + O(1)$$

where ψ is defined by (6.4).

By Law of Iterated Logarithm [93],

$$\limsup_{n\to\infty} \frac{\sum_{j=0}^{n-1}\psi(f^j x)}{\sqrt{2n\ln\ln n}} = \sigma, \quad \liminf_{n\to\infty} \frac{\sum_{j=0}^{n-1}\psi(f^j x)}{\sqrt{2n\ln\ln n}} = -\sigma.$$

Since $B(x, r_n) \subset B_n(x,\varepsilon) \subset B(x, \bar{r}_n)$

$$\limsup_{n\to\infty} \frac{|\ln\mu\big(B(x,\bar{r}_n)\big)| - \mathbf{d}|\ln \bar{r}_n|}{\sqrt{2n\ln\ln n}} \le \sigma \le \limsup_{n\to\infty} \frac{|\ln\mu(B(x,r_n))| - \mathbf{d}|\ln r_n|}{\sqrt{2n\ln\ln n}}.$$

Using (B.2) again, we conclude that for every sufficiently small δ, there exist $n(\delta)$ and k independent of δ and $n(\delta)$ such that $\bar{r}_{n+k} \le \delta \le r_n$. Then

$$\begin{aligned}\sigma &\le \limsup_{\delta\to 0} \frac{|\ln\mu(B(x,r_{n(\delta)}))| - \mathbf{d}|\ln r_{n(\delta)}|}{\sqrt{2n(\delta)\ln\ln n(\delta)}} \le \limsup_{\delta\to 0} \frac{\big|\ln\mu(B(x,\delta))\big| - \mathbf{d}|\ln\delta|}{\sqrt{2n(\delta)\ln\ln n(\delta)}} \\ &\le \limsup_{\delta\to 0} \frac{|\ln\mu(B(x,\bar{r}_{n(\delta)}))| - \mathbf{d}|\ln \bar{r}_{n(\delta)}|}{\sqrt{2n(\delta)\ln\ln n(\delta)}} \le \sigma.\end{aligned}$$

It follows that all inequalities above are in fact equalities. In particular,

$$\limsup_{\delta\to 0} \frac{|\ln\mu(B(x,\delta))| - \mathbf{d}|\ln\delta|}{\sqrt{2n(\delta)\ln\ln n(\delta)}} = \sigma.$$

On the other hand, by (B.2) and the ergodic theorem we see that for μ-a.e. $x \in \mathbb{T}$, it holds that $\lim_{n\to\infty} \frac{|\ln r_n|}{n} = \lambda$. For such x we have $\lim_{n\to\infty} \frac{|\ln r_n|(\ln\ln|\ln r_n|)}{n\ln\ln n} = \lambda$. Since $r_n/C \le \delta \le r_n$ we have

$$\lim_{\delta\to 0} \sqrt{\frac{n(\delta)\ln\ln n(\delta)}{|\ln\delta|(\ln\ln|\ln\delta|)}} = \frac{1}{\sqrt{\lambda}}.$$

Multiplying the last two displays we obtain for μ-a.e. $x \in \mathbb{T}$

$$\limsup_{\delta\to 0} \frac{|\ln\mu(B(x,\delta))| - \mathbf{d}|\ln\delta|}{\sqrt{2|\ln\delta|(|\ln\ln|\ln\delta|)}} = \frac{\sigma}{\sqrt{\lambda}},$$

and likewise

$$\liminf_{\delta\to 0} \frac{|\ln\mu(B(x,\delta))| - \mathbf{d}|\ln\delta|}{\sqrt{2|\ln\delta|(|\ln\ln|\ln\delta|)}} = -\frac{\sigma}{\sqrt{\lambda}}.$$

This proves part (b) of Lemma 6.3. □

B.4. **Proof of Lemma 6.3(a).** Suppose that $\sigma^2 = 0$. Since we also have that $\int \psi\, d\mu = 0$, [133, Proposition 4.12] shows that ψ is a coboundary, that is, there exists a Hölder function η such that $\psi(x) = \eta(x) - \eta(fx)$. Thus $\sum_{k=0}^{n-1} \psi(f^k x) = \eta(x) - \eta(f^n x)$ is uniformly bounded with respect to both n and x. Recalling the definition of ψ we see that in this case

$$\sum_{k=0}^{n-1} g(f^k x) = -\left[\mathbf{d}\sum_{k=0}^{n-1} f_u(f^k x)\right] + O(1).$$

Now (B.1) and (B.2) show that μ is conformal.

APPENDIX C. GEOMETRY OF TARGETS IN THE CONFIGURATION SPACE

C.1. **Geometry of spheres. Proof of Lemma 8.7.** Denote $\gamma(t) = \phi^t(q, v)$. The Jacobi field of γ are defined by the solution of the linear equation

$$J''(t) + R(J(t), \gamma'(t))\gamma'(t) = 0,$$

where $J' = \frac{d}{dt}J$ and $R(X, Y)Z$ denotes the curvature tensor, which is equivalent to

$$(J^i)''(t) + \sum_{j=1}^{n} A^i_j(t)J^j(t) = 1, i = 1, \dots, n,$$

where the matrix $A(t) = (A^i_j(t))_{i,j=1,\dots,n}$ is symmetric. Since $\mathcal{Q}$ has negative curvature, the spectrum of $A(t)$ lies between $-K_1^2$ and $-K_2^2$ for some K_1 and K_2.

Recall the following fact (see [119, Lemma 1.1]).

PROPOSITION C.1. *The differential*

$$D\phi^t(v) : T_{\pi v}\mathcal{Q} \times T_{\pi v}\mathcal{Q} \to T_{\pi\phi^t(v)}\mathcal{Q} \times T_{\pi\phi^t(v)}\mathcal{Q}$$

is given by $D\phi^t(v)(x, y) = (J(t), J'(t))$, *where* $J(0) = x$, $J'(0) = y$.

We are interested in the case

$$J(0) = 0, \quad \|J'(0)\| = 1. \tag{C.1}$$

Now Lemma 8.7 follows combining Proposition C.1 with Lemma C.2 below.

LEMMA C.2. *If* (C.1) *holds then for each* t_0 *there is a constant* $C > 0$ *such that*

$$\|J'(t)\| \le C\|J(t)\| \quad \text{for } t > t_0. \tag{C.2}$$

Proof. Denote $S(t) = \langle J(t), J'(t)\rangle$, $N(t) = \|J'(t)\|^2$ and $|||J|||^2 = \|J\|^2 + \|J'\|^2$. Then

$$\begin{aligned}\frac{d}{dt}S(t) &= \|J'(t)\|^2 + \langle J(t), J''(t)\rangle \\ &= \|J'(t)\|^2 + \langle J(t), -K(t)J(t)\rangle \ge C_1|||J|||^2\end{aligned} \tag{C.3}$$

for some $C_1 > 0$. It follows that $S(t) > 0$ for $t > 0$. Once we know that $S(t)$ is positive we can also conclude from (C.3) that $\dfrac{d}{dt}S(t) > \dfrac{C_1 S(t)}{2}$, whence

$$S(t) > S(u)e^{C_1(t-u)/2} \quad \text{for } t > u. \tag{C.4}$$

Next $N(t) \ge N(0)e^{-K_2^2 t} = e^{-K_2^2 t}$ which together with (C.3) gives

$$S(t) \ge e^{-K_2^2 t}t \quad \text{for } t \in [0, 1]. \tag{C.5}$$

Combining this with (C.4) we get

$$S(t) > e^{-K_2^2}e^{C_1(t-1)/2} \quad \text{for } t > 1. \tag{C.6}$$

Combining (C.5) and (C.6) with a trivial bound

$$N(t) \le |||J(t)||| \le N(0)e^{K_2^2 t} = e^{K_2^2 t} \tag{C.7}$$

proves (C.2) for small t. To prove this estimate for large t we shall use the fact, proven in [6, Lecture 6] that J can be decomposed as $J = c_+ J_+ + c_- J_-$, where

$$\max(|c_+|, |c_-|) \le C_3, \quad |||J_-||| \le C_4 e^{-K_1 t}$$

and

$$J_+ = R(t) J_+'(t) \tag{C.8}$$

where R is a symmetric matrix with spectrum between K_1 and K_2. It follows that

$$|||J(t)||| \le c_+ |||J_+(t)||| + C_3 C_4 e^{-K_1 t} \le \sqrt{1+K_2^2}\, \|c_+ J_+(t)\| + C_3 C_4 e^{-K_1 t} \tag{C.9}$$

On the other hand (C.6) gives a uniform lower bound

$$|||J||| \ge 2 e^{-K_2^2/2} e^{C_1(t-1)/4}. \tag{C.10}$$

Combining (C.9) and (C.10) we obtain

$$\|J(t)\| \ge \|c_+ J^+(t)\| - c_- \|J^-(t)\| \ge \frac{2}{1+K_2^2} e^{-K_2^2/2} e^{C_1(t-1)/4} - 2 C_3 C_4 e^{-K_1 t}$$

which proves (C.2) for large t. □

C.2. **Volume of the targets in the configuration space.**

Proof of Lemma 8.2. If $(q, v) \in \hat{B}_\rho(a)$, denote

$$L(q, v) = L^+(q, v) + L^-(q, v)$$

where

$$L^\pm(q, v) = \sup\{t : \phi^{\pm s}(q, v) \in \hat{B}_\rho(a) \text{ for } 0 \le s \le t\}.$$

Then we have the following estimate

$$\mu\left(\Omega_{a,\rho}\right) = \varepsilon \left(\int_{\hat{B}_\rho(a)} \frac{1}{L(q, v)} \mathrm{d}\mu \right) \left(1 + O(\rho)\right)$$

(see, e.g., [36]). Note that μ is of the form $d\mu(q, v) = \frac{\mathrm{d}\lambda(q)\mathrm{d}\sigma(v)}{\lambda(\mathcal{Q})}$ where λ is the Riemann volume on $\mathcal{Q}$ and σ is normalized volume on the d-dimensional sphere. If ρ is small then the integral in parenthesis equals to $\rho^d \gamma(1 + O(\rho))$ where

$$\gamma = \frac{1}{\lambda(\mathcal{Q})} \int_{\mathscr{B} \times \mathbb{S}^d} \frac{1}{\mathscr{L}(x, v)} \mathrm{d}x \mathrm{d}\sigma(v) \tag{C.11}$$

where $\mathscr{B}$ is the unit ball in $\mathbb{R}^{d+1}$ and $\mathscr{L}(\cdot)$ is defined similarly $L(\cdot)$ with geodesics in $\mathcal{Q}$ replaced by geodesics in $\mathbb{R}^{d+1}$. Specifically, an elementary plane geometry gives $\mathscr{L}(x, v) = \sqrt{1 - r_{min}^2}$ where r_{min} is the minimal distance between the line $x + tv$ and the origin. Thus $r_{min} = r \sin\theta$ where r is the distance from x to 0, θ is the angle between v and the segment from x to 0. This proves (8.1) with γ given by (C.11). □

REFERENCES

[1] J. Aaronson and H. Nakada, Trimmed sums for non-negative, mixing stationary processes, *Stochastic Process. Appl.*, **104** (2003), 173–192.

[2] M. Abadi, Poisson approximations via Chen-Stein for non-Markov processes, *Progr. Probab.*, **60** (2008), 1–19.

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

[4] R. Aimino, M. Nicol and M. Todd, Recurrence statistics for the space of interval exchange maps and the Teichmuller flow on the space of translation surfaces, *Ann. Inst. Henri Poincaré Probab. Stat.*, **53** (2017), 1371–1401.

[5] D. Aldous, Probability approximations via the Poisson clumping heuristic, *Applied Math. Sci.*, **77** (1989), 269 pp.

[6] D. V. Anosov and Y. G. Sinai, Certain smooth ergodic systems, *Russian Math. Surveys*, **22** (1967), 103–167.

[7] J. S. Athreya, Logarithm laws and shrinking target properties, *Proc. Indian Acad. Sci. Math. Sci.*, **119** (2009), 541–557.

[8] J. S. Athreya, Random affine lattices, *Contemp. Math.*, **639** (2015), 169–174.

[9] J. S. Athreya, A. Ghosh and J. Tseng, Spiraling of approximations and spherical averages of Siegel transforms, *J. Lond. Math. Soc.*, **91** (2015), 383–404.

[10] J. S. Athreya and G. A. Margulis, Logarithm laws for unipotent flows-I, *J. Mod. Dyn.*, **3** (2009), 359–378.

[11] J. S. Athreya and G. A. Margulis, Logarithm laws for unipotent flows-II, *J. Mod. Dyn.*, **11** (2017), 1–16.

[12] J. S. Athreya, A. Parrish and J. Tseng, Ergodic theory and Diophantine approximation for translation surfaces and linear forms, *Nonlinearity*, **29** (2016), 2173–2190.

[13] M. Babillot and M. Peigne, Asymptotic laws for geodesic homology on hyperbolic manifolds with cusps, *Bull. Soc. Math. France*, **134** (2006), 119–163.

[14] D. Badziahin, V. Beresnevich and S. Velani, Inhomogeneous theory of dual Diophantine approximation on manifolds, *Adv. Math.*, **232** (2013), 1–35.

[15] L. Barreira, Y. Pesin and J. Schmeling, Dimension and product structure of hyperbolic measures, *Ann. of Math.*, **149** (1999), 755–783.

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

[17] M. B. Bekka and M. Mayer, Ergodic theory and topological dynamics of group actions on homogeneous spaces, *London Math. Soc. Lecture Note Ser.*, **269** (2000), 200 pp.

[18] R. Benedetti and C. Petronio, *Lectures on Hyperbolic Geometry*, Springer, Berlin, 1992.

[19] V. I. Bernik and M. M. Dodson, Metric Diophantine approximation on manifolds, Cambridge Tracts in Mathematics, **137** (1999), 172 pp.

[20] M. Björklund, M. Einsiedler and A. Gorodnik, Quantitative multiple mixing, *JEMS*, **22** (2020), 1475–1529.

[21] M. Björklund and A. Gorodnik, Central limit theorems in the geometry of numbers, *Electron. Res. Announc. Math. Sci.*, **24** (2017), 110–122.

[22] M. Björklund and A. Gorodnik, Central limit theorems for group actions which are exponentially mixing of all orders, *Journal d'Analyse Mathematiques*, **141** (2020), 457–482.

[23] M. Björklund and A. Gorodnik, Central limit theorems for Diophantine approximants, *Math. Ann.*, **374** (2019), 1371–1437.

[24] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms. 2d revised ed.*, Springer Lecture Notes in Math, **470** (2008), 75 pp.

[25] H. Bruin and M. Todd, Return time statistics of invariant measures for interval maps with positive Lyapunov exponent, *Stoch. Dyn.*, **9** (2009), 81–100.

[26] H. Bruin and S. Vaienti, Return time statistics for unimodal maps, *Fund. Math.*, **176** (2003), 77–94.

[27] M. Carney, M. Holland and M. Nicol, Extremes and extremal indices for level set observables on hyperbolic systems, *Nonlinearity*, **34** (2021), 1136–1167.

[28] M. Carney and M. Nicol, Dynamical Borel–Cantelli lemmas and rates of growth of Birkhoff sums of non-integrable observables on chaotic dynamical systems, *Nonlinearity*, **30** (2017), 2854–2870.

[29] M. Carvalho, A. C. M. Freitas, J. M. Freitas, M. Holland and M. Nicol, Extremal dichotomy for uniformly hyperbolic systems, *Dyn. Syst.*, **30** (2015), 383–403.

[30] A. Castro, Fast mixing for attractors with a mostly contracting central direction, *Ergodic Theory Dynam. Systems*, **24** (2004), 17–44.

[31] J. Chaika, Shrinking targets for IETs: Extending a theorem of Kurzweil, *GAFA*, **21** (2011), 1020–1042.

[32] J. Chaika and D. Constantine, Quantitative shrinking target properties for rotations and interval exchanges, *Israel J. Math.*, **230** (2019), 275–334.

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

[34] J.-R. Chazottes and E. Ugalde, Entropy estimation and fluctuations of hitting and recurrence times for Gibbsian sources, *Discrete Contin. Dyn. Syst. B*, **5** (2005), 565–586.

[35] N. I. Chernov, Limit theorems and Markov approximations for chaotic dynamical systems, *Probab. Theory Related Fields*, **101** (1995), 321–362.

[36] N. Chernov, Entropy, Lyapunov exponents, and mean free path for billiards, *J. Statist. Phys.*, **88** (1997), 1–29.

[37] N. Chernov and D. Kleinbock, Dynamical Borel–Cantelli lemmas for Gibbs measures, *Israel J. Math.*, **122** (2001), 1–27.

[38] G. H. Choe and B. K. Seo, Recurrence speed of multiples of an irrational number, *Proc. Japan Acad. Ser. A Math. Sci.*, **77** (2001), 134–137.

[39] Z. Coelho, Asymptotic laws for symbolic dynamical processes, *London Math. Soc. Lecture Notes*, **279** (2000), 123–165.

[40] P. Collet, Statistics of closest return for some non-uniformly hyperbolic systems, *Ergodic Theory Dynam. Systems*, **21** (2001), 401–420.

[41] P. Collet, A. Galves and B. Schmitt, Repetition times for Gibbsian sources, *Nonlinearity*, **12** (1999), 1225–1237.

[42] I. P. Cornfeld, S. V. Fomin and Y. G. Sinai, *Ergodic Theory*, Grundlehren der Mathematischen Wissenschaften, 245, Springer, New York, 1982.

[43] S. G. Dani, Divergent trajectories of flows on homogeneous spaces and Diophantine approximation, *J. Reine Angew. Math.*, **359** (1985), 55–89.

[44] M. Denker and N. Kan, Om Sevastyanov's theorem, *Stat. Probab. Lett.*, **77** (2007), 272–279.

[45] H. G. Diamond and J. D. Vaaler, Estimates for partial sums of continued fraction partial quotients, *Pacific J. Math.*, **122** (1986), 73–82.

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

[47] D. Dolgopyat, On dynamics of mostly contracting diffeomorphisms, *Comm. Math. Phys.*, **213** (2000), 181–201.

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

[49] D. Dolgopyat, C. Dong, A. Kanigowski and P. Nandori, Mixing properties of generalized T, T^{-1} transformations, *Israel J. Math.*, **247** (2022), no. 1, 21–73.

[50] D. Dolgopyat and B. Fayad, Limit theorems for toral translations, *Hyperbolic dynamics, fluctuations and large deviations*, **89** (2015), 227–277.

[51] D. Dolgopyat and B. Fayad, Deviations of ergodic sums for toral translations II: Squares, *Publ. IHES*, **132** (2020), 293–352.

[52] D. Dolgopyat, B. Fayad and I. Vinogradov, Central limit theorems for simultaneous Diophantine approximations, *J. Éc. Polytech. Math.*, **4** (2017), 1–35.

[53] D. Dolgopyat, A. Kanigowski and F. R. Hertz, Exponential mixing implies Bernoulli, arXiv:2106.03147.

[54] D. Dolgopyat and O. Sarig, Quenched and annealed temporal limit theorems for circle rotations, *Asterisque*, **415** (2020), 57–83.

[55] L. E. Dubins and D. A. Freedman, A sharper form of the Borel–Cantelli lemma and the strong law, *Ann. Math. Statist.*, **36** (1965), 800–807.

[56] N. Enriquez and Y. L. Jan, Statistic of the winding of geodesics on a Riemann surface with finite area and constant negative curvature, *Rev. Mat. Iberoamericana*, **13** (1997), 377–401.

[57] M. L. Einsiedler, D. A. Ellwood, A. Eskin, D. Kleinbock, E. Lindenstrauss, G. Margulis, S. Marmi and J.-C. Yoccoz (editors): *Homogeneous Flows, Moduli Spaces and Arithmetic*, Clay Mathematics Proceedings, 10, 2010.

[58] D. El-Baz, J. Marklof and I. Vinogradov, The distribution of directions in an affine lattice: Two-point correlations and mixed moments, *IMRN*, **2015** (2015), 1371–1400.

[59] A. Eskin, Counting problems and semisimple groups, *Doc. Math.*, **II** (1998), 539–552.

[60] B. Fayad, Mixing in the absence of the shrinking target property, *Bull. London Math. Soc.*, **38** (2006), 829–838.

[61] J. L. Fernandez, M. V. Melian and D. Pestana, Expanding maps, shrinking targets and hitting times, *Nonlinearity*, **25** (2012), 2443–2471.

[62] D. Freedman, Another note on the Borel–Cantelli Lemma and the strong law with the Poisson approximation as a by-product, *Ann. Probability*, **1** (1973), 910–925.

[63] A. C. M. Freitas, J. M. Freitas, M. Magalhães, Convergence of marked point processes of excesses for dynamical systems, *J. Eur. Math. Soc.*, **20** (2018), 2131–2179.

[64] A. C. M. Freitas, J. M. Freitas and M. Magalhães, Complete convergence and records for dynamically generated stochastic processes, *Trans. Amer. Math. Soc.*, **373** (2020), 435–478.

[65] A. C. M. Freitas, J. M. Freitas, F. Rodrigues and J. V. Soares, Rare events for Cantor target sets, *Comm. Math. Phys.*, **378** (2020), 75–115.

[66] A. C. M. Freitas, J. M. Freitas and M. Todd, Hitting time statistics and extreme value theory, *Probab. Theory Related Fields*, **147** (2010), 675–710.

[67] S. Galatolo, Dimension via waiting time and recurrence, *Math. Res. Lett.*, **12** (2005), 377–386.

[68] S. Galatolo, Dimension and hitting in rapidly mixing system, *Math. Res. Lett.*, **14** (2007), 797–805.

[69] S. Galatolo, Hitting time in regular sets and logarithm law for rapidly mixing dynamical systems, *Proc. Amer. Math. Soc.*, **138** (2010), 2477–2487.

[70] S. Galatolo and I. Nisoli, Shrinking targets in fast mixing flows and the geodesic flow on negatively curved manifolds, *Nonlinearity*, **24** (2011), 3099–3113.

[71] S. Galatolo and P. Peterlongo, Long hitting time, slow decay of correlations and arithmetical properties, *Discrete Contin. Dyn. Syst.*, **27** (2010), 185–204.

[72] A. Ghosh and D. Kelmer, Shrinking targets for semisimple groups, *Bull. Lond. Math. Soc.*, **49** (2017), 235–245.

[73] A. Gorodnik and A. Nevo, Quantitative ergodic theorems and their number-theoretic applications, *Bull. Lond. Math. Soc.*, **52** (2015), 65–113.

[74] A. Gorodnik and R. Spatzier, Exponential mixing of nilmanifold automorphisms, *J. Anal. Math.*, **123** (2014), 355–396.

[75] S. Gouëzel, A Borel–Cantelli lemma for intermittent interval maps, *Nonlinearity*, **20** (2007), 1491–1497.

[76] A. Groshev, A theorem on a system of linear forms, *Dokl. Akad. Nauk SSSR*, **19** (1938), 151–152.

[77] Y. Guivarch and Y. L. Jan, Asymptotic winding of the geodesic flow on modular surfaces and continued fractions, *Ann. Sci. École Norm. Sup.*, **26** (1993), 23–50; **29** (1996), 811–814.

[78] C. Gupta, M. Nicol and W. Ott, A Borel–Cantelli lemma for nonuniformly expanding dynamical systems, *Nonlinearity*, **23** (2010), 1991–2008.

[79] N. Haydn, The central limit theorem for uniformly strong mixing measures, *Stoch. Dyn.*, **12** (2012), 1250006, 31 pp.

[80] N. Haydn, Entry and return times distribution, *Dyn. Syst.*, **28** (2013), 333–353.

[81] N. Haydn and S. Vaienti, Limiting entry and return times distribution for arbitrary null sets, *Comm. Math. Phys.*, **378** (2020), 149–184.

[82] 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.

[83] N. Haydn, M. Nicol, S. Vaienti and L. Zhang, Central limit theorems for the shrinking target problem, *J. Stat. Phys*,**153** (2013), 864–887.

[84] N. T. A. Haydn and Y. Psiloyenis, Return times distribution for Markov towers with decay of correlations, *Nonlinearity*, **27** (2014), 1323–1349.

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

[86] N. Haydn and S. Vaienti, Fluctuations of the metric entropy for mixing measures, *Stoch. Dyn.*, **4** (2004), 595–627.

[87] N. Haydn and S. Vaienti, Limiting entry times distribution for arbitrary null sets, *Comm. Math. Phys.* , **378** (2020), 149–184.

[88] N. T. A. Haydn and K. Wasilewska, Limiting distribution and error terms for the number of visits to balls in nonuniformly hyperbolic dynamical systems, *Discrete Contin. Dyn. Syst.*, **36** (2016), 2585–2611.

[89] R. Hill and S. Velani, The ergodic theory of shrinking targets, *Invent. Math.*, **119** (1995), 175–198.

[90] M. Hirata, Poisson law for Axiom A diffeomorphisms, *Ergodic Theory Dynam. Systems*, **13** (1993), 533–556.

[91] M. Hirata, Poisson law for the dynamical systems with self-mixing conditions, in *Dynamical Systems and Chaos, Vol. 1 (Hachioji, 1994)*, World Sci. Publ., River Edge, NJ, 1995, 87–96.

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

[93] F. Hofbauer and G. Keller, Ergodic properties of invariant measures for piecewise monotonic transformations, in *The Theory of Chaotic Attractors*, B. R. Hunt, T. Y. Li, J. A. Kennedy, H. E. Nusse (eds), Springer, New York, 1982.

[94] M. Holland, M. Kirsebom, P. Kunde and T. Persson, Dichotomy results for eventually always hitting time statistics and almost sure growth of extremes, `arXiv:2109.06314`.

[95] I. A. Ibragimov, Some limit theorems for stationary processes, (Russian) *Teor. Verojatnost. i Primenen.*, **7** (1962), 361–392.

[96] J. Jaerisch, M. Kessebohmer, B. O. Stratmann, A Fréchet law and an Erdös–Philipp law for maximal cuspidal windings, *Ergodic Theory Dynam. Systems*, **33** (2013), 1008–1028.

[97] G. Keller, Rare events, exponential hitting times and extremal indices via spectral perturbation, *Dyn. Syst.*, **27** (2012), 11–27.

[98] D. Kelmer, Shrinking targets for discrete time flows on hyperbolic manifolds, *Geom. Funct. Anal.*, **27** (2017), 1257–1287.

[99] D. Kelmer and H. Oh, Exponential mixing and shrinking targets for geodesic flow on geometrically finite hyperbolic manifolds, *J. Mod. Dyn.*, **17** (2021), 401–434.

[100] D. Kelmer and S. Yu, Shrinking target problems for flows on homogeneous spaces, *Trans. Amer. Math. Soc.*, **372** (2019), 6283–6314.

[101] M. Kesseböhmer and T. Schindler, Strong laws of large numbers for intermediately trimmed Birkhoff sums of observables with infinite mean, *Stoch. Process. Appl.*, **129** (2019), 4163–4207.

[102] H. Kesten and R. A. Maller, Ratios of trimmed sums and order statistics, *Ann. Probab.*, **20** (1992), 1805–1842.

[103] A. Y. Khintchine, Einige Sätze über Kettenbruche, mit Anwendungen auf die Theorie der Diophantischen Approximationen, (German) *Math. Ann.*, **92** (1924), 115–125.
[104] A. Y. Khintchine, *Continued Fractions*, Dover, Mineola, NY, 1997.
[105] M. Kirsebom, P. Kunde and T. Persson, Shrinking targets and eventually always hitting points for interval maps, *Nonlinearity*, **33** (2020), no. 2, 892–914.
[106] D. H. Kim, The shrinking target property of irrational rotations, *Nonlinearity*, **20** (2007), 1637–1643.
[107] D. H. Kim, The dynamical Borel–Cantelli lemma for interval maps, *Discrete Contin. Dyn. Syst.*, **17** (2007), 891–900.
[108] D. H. Kim, Refined shrinking target property of rotations, *Nonlinearity*, **27** (2014), 1985–1997.
[109] D. H. Kim and B. K. Seo, The waiting time for irrational rotations, *Nonlinearity*, **16** (2003), 1861–1868.
[110] D. Kleinbock, Some applications of homogeneous dynamics to number theory, *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, **69** (2001), 639–660.
[111] D. Kleinbock, I. Konstantoulas and F. K. Richter, Zero-one laws for eventually always hitting points in mixing systems, `arXiv:1904.08584`.
[112] D. Y. Kleinbock and G. A. Margulis, Bounded orbits of nonquasiunipotent flows on homogeneous spaces, *Amer. Math. Soc. Transl. Ser. 2,* **171** (1996), 141–172.
[113] D. Y. Kleinbock and G. A. Margulis, Flows on homogeneous spaces and Diophantine approximation on manifolds, *Ann. Math.*, **148** (1998), 339–360.
[114] D. Y. Kleinbock and G. A. Margulis, Logarithm laws for flows on homogeneous spaces, *Invent. Math.*, **138** (1999), 451–494.
[115] D. Y. Kleinbock and G. A. Margulis, On effective equidistribution of expanding translates of certain orbits in the space of lattices, *Number Theory, Analysis and Geometry*, Springer, New York, 2012, 385–396.
[116] D. Kleinbock, N. Shah and A. Starkov, Dynamics of subgroup actions on homogeneous spaces of Lie groups and applications to number theory, *Handbook of Dynamical Systems, Vol. 1A* (2002), North-Holland, Amsterdam, 813–930.
[117] D. Kleinbock, R. Shi and B. Weiss, Pointwise equidistribution with an error rate and with respect to unbounded functions, *Math. Ann.*, **36** (2017), 857–879.
[118] D. Y. Kleinbock and X. Zhao, An application of lattice points counting to shrinking target problems, *Discrete Contin. Dyn. Syst.*, **38** (2018), 155–168.
[119] G. Knieper, Hyperbolic dynamics and Riemannian geometry, *Handbook of Dynamical Systems, Vol. 1A,* 453, North-Holland, Amsterdam, 2002.
[120] J. Kurzweil, On the metric theory of inhomogeneous Diophantine approximations, *Studia Math.*, **15** (1955), 84–112.
[121] M. R. Leadbetter, G. Lindgren and H. Rootzen, *Extremes and Related Properties of Random Sequences and Processes*, Springer, New York–Berlin, 1983.
[122] F. Ledrappier and L.-S. Young, The metric entropy of diffeomorphisms. Part II. Relations between entropy, exponents and dimension, *Ann. Math.*, **122** (1985), 540–574.
[123] R. Leplaideur and B. Saussol, Central limit theorem for dimension of Gibbs measures in hyperbolic dynamics, *Stoch. Dyn.*, **12** (2012), 1150019, 24 pp.
[124] C. Liverani, On contact Anosov flows, *Ann. of Math. (2)*, **159** (2004), 1275–1312.
[125] V. Lucarini, D. Faranda, A. de Freitas, J. de Freitas, M. Holland, T. Kuna, M. Nicol, M. Todd and S. Vaienti, *Extremes and Recurrence in Dynamical Systems*, John Wiley, Hoboken, NJ, 2016.
[126] J. Marklof, The n-point correlations between values of a linear form, *Ergodic Theory Dynam. Systems*, **20** (2000), 1127–1172.
[127] J. Marklof, Distribution modulo one and Ratner's theorem, *Equidistribution in Number Theory, an Introduction*, **237** (2007), 217–244.
[128] J. Marklof, Entry and return times for semi-flows, *Nonlinearity*, **30** (2017), 810–824.

[129] F. Maucourant, Dynamical Borel–Cantelli lemma for hyperbolic spaces, *Israel J. Math.*, **152** (2006), 143–155.

[130] T. Mori, The strong law of large numbers when extreme terms are excluded from sums, *Z. Wahrsch. Verw. Gebiete*, **36** (1976), 189–194.

[131] F. Paccaut, Statistics of return times for weighted maps of the interval, *Ann. Inst. H. Poincare Probab. Statist.*, **36** (2000), 339–366.

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

[133] W. Parry and M. Pollicott, Zeta functions and the periodic orbit structure of hyperbolic dynamics, *Astérisque*, **187-188** (1990), 268 pp.

[134] F. Paulin and M. Pollicott, Logarithm laws for equilibrium states in negative curvature, *Comm. Math. Phys.*, **346** (2016), 1–34.

[135] F. Pène and B. Saussol, Back to balls in billiards, *Comm. Math. Phys.*, **293** (2010), 837–866.

[136] 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.

[137] W. Philipp, Some metric theorems in number theory, *Pacific J. Math*, **20** (1967), 109–127.

[138] B. Pitskel, Poisson limit law for Markov chains, *Ergodic Theory Dynam. Sys.*, **11** (1991), 501–513.

[139] S. Ross, *A First Course in Probability*, 8^{th} edition, Macmillan Co., New York; Collier Macmillan Ltd., London, 2010.

[140] J. Rousseau, Recurrence rates for observations of flows, *Ergodic Theory Dynam. Systems*, **32** (2012), 1727–1751.

[141] J. Rousseau, Hitting time statistics for observations of dynamical systems, *Nonlinearity*, **27** (2014), 2377–2392.

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

[143] B. Saussol, Recurrence rate in rapidly mixing dynamical systems, *Discrete Contin. Dyn. Syst.*, **15** (2006), 259–267.

[144] B. Saussol, An introduction to quantitative Poincaré recurrence in dynamical systems, *Rev. Math. Phys.*, **21** (2009), 949–979.

[145] C. Series, The modular surface and continued fractions, *J. London Math. Soc.*, **31** (1985), 69–80.

[146] B. A. A. Sevastjanov, Possion limit law in a scheme of sums of dependent random variables, (Russian) *Teor. Verojatnost. i Primenen.*, **17** (1972), 733–738.

[147] D. Simmons, An analogue of a theorem of Kurzweil, *Nonlinearity*, **28** (2015), 1401–1408.

[148] Y. G. Sinai, Gibbs measures in ergodic theory, *Russian Math. Surveys*, **27** (1972), 21–69.

[149] V. G. Sprindžuk, *Metric theory of Diophantine approximations*, V. H. Winston & Sons, Washington, D.C. (1979).

[150] M. A. Stenlund, Strong pair correlation bound implies the CLT for Sinai billiards, *J. Stat. Phys.*, **140** (2010), 154–169.

[151] B. Stratmann and S. L. Velani, The Patterson measure for geometrically finite groups with parabolic elements, new and old, *Proc. London Math. Soc. (3)*, **71** (1995), 197–220.

[152] D. Sullivan, Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics, *Acta Math.*, **149** (1982), 215–237.

[153] J. Tseng, On circle rotations and the shrinking target properties, *Discrete Contin. Dyn. Syst.*, **20** (2008), 1111–1122.

[154] P. Varandas, Correlation decay and recurrence asymptotics for some robust nonuniformly hyperbolic maps, *J. Stat. Phys.*, **133** (2008), 813–839.

[155] M. Viana, *Stochastic Dynamics of Deterministic Systems*, Rio de Janeiro: IMPA, 1997.

[156] I. Vinogradov, Limiting distribution of visits of several rotations to shrinking intervals, preprint.

[157] D. Williams, *Probability with Martingales*, Cambridge Math. Textbooks. Cambridge University Press, Cambridge, 1991.

[158] F. Yang, Rare event process and entry times distribution for arbitrary null sets on compact manifolds, *Ann. Inst. Henri Poincare Probab. Stat.*, **57** (2021), 1103–1135.
[159] L.-S. Young, Statistical properties of dynamical systems with some hyperbolicity, *Ann. Math.*, **147** (1998), 585–650.
[160] X. Zhang, Note on limit distribution of normalized return times and escape rate, *Stoch. Dyn.*, **16** (2016), 1660014, 15 pp.

DMITRY DOLGOPYAT <dolgop@umd.edu>: Department of Mathematics, University of Maryland, 4176 Campus Drive, College Park, MD 20742–4015, USA

BASSAM FAYAD <bassam@umd.edu>: Department of Mathematics, University of Maryland, 4176 Campus Drive, College Park, MD 20742–4015, USA

SIXU LIU <liusixu@mail.tsinghua.edu.cn>: Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, China