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

DMITRY DOLGOPYAT, BASSAM FAYAD, AND SIXU LIU

ABSTRACT. A classical Borel Cantelli Lemma gives conditions for deciding whether an infinite number of rare events will almost surely happen. 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, Diophantine approximations and extreme value theory for dynamical systems.

CONTENTS

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 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 $\{\Omega_n\}$ 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 [4] for a comprehensive discussion of Poisson clustering.

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 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 for large n, $\mathbb{P}(\Omega_\rho^n) \approx \sigma(\rho)$ for some function $\sigma(\rho)$. Let N_ρ^n be the number of Ω_ρ^k, $k \leqslant n$ which has occurred. We fix a sequence ρ_n such that $n\sigma(\rho_n) \to 0$ as $n \to \infty$ and $r \in \mathbb{N}$, and ask if infinitely many events

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

occur. Even if the events Ω_ρ^n are independent for different ρ, 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 those variables are weakly dependent since conditioned on $N_{\rho_{n_2}}^{n_2} \neq 0$ it is very likely that all the events $\Omega_{\rho_{n_2}}^k$ occur for $k > n_1$. Using this, one can show under appropriate monotonicity assumptions (see [118]) that $N_{\rho_n}^n = r$ infinitely often if and only if

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

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 $N^n_{\rho_n} = r$ occur if and only if

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

The 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 more flexible mixing conditions for the multiple Borel Cantelli Lemma. 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. Here we describe sample applications to dynamics, geometry, and number theory.

MultiLog Law for recurrence. Let f be a map preserving a measure μ. Given two points x, y let $d_n^{(r)}(x, y)$ be the r closest distance among $d(x, f^k y)$ for $1 \leqslant k \leqslant n$. In particular, $d_n^{(1)}(x, y)$ is the closest distance the orbit of y comes to x up to time n. It is shown in [61] that for systems with superpolynomial decay for Lipschitz observables, for all x and μ-almost all y $\lim_{n\to\infty} \frac{|\ln d_n^{(1)}(x, y)|}{\ln n} = \frac{1}{\mathbf{d}}$, 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 and y we have

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

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 (6.4). We shall also show that there is G_δ–dense set $\mathcal{H}$ such that for all $x \in \mathcal{H}$, Lebesgue almost all y and all $r \geqslant 1$, 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) it holds that

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

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 compact part and several cusps. *A cusp excursion* is a maximal time segment such that the geodesic stays in a cusp for the whole segment. Let

$$H^{(1)}(T) \geqslant H^{(2)}(T) \geqslant \ldots H^{(r)}(T) \geqslant \ldots$$

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

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

The proof of (1.1) relies on Sullivan's Borel-Cantelli Lemma and it actually also shows that for almost every geodesic

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

We obtain a multiple version of this result by showing that for almost every geodesic

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

MULTIPLE KHINCHINE 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 Khinchine Groshev Theorem ([69, 93, 132]) says that for almost all $\alpha \in \mathbb{R}^d$ there are infinitely many solutions to

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

if and only if

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

In particular the inequality

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

where $|k| = \sqrt{\sum k_i^2}$, has infinitely many solutions for almost every α if and only if $s \leqslant 1$. We now replace the above inequality by

$$(1.4) \qquad |k|^d|\langle k, \alpha \rangle + m| \leqslant \frac{1}{\ln N(\ln \ln N)^s}$$

and say that α is (r, s) approximable if there are infinitely N for which (1.4) has r positive solutions (that is, solutions with $k_1 > 0$). (Our interest in smallness of

$$(1.5) \qquad |k|^d|\langle k, \alpha \rangle + m|$$

is motivated by [46] where the discrepancy of Kronecker sequences with respect to convex sets is studied. Indeed the set of k where (1.5) is small are small denominators of the discrepancy and they determine its growth rate.) We show in Section 9 that almost every $\alpha \in \mathbb{R}^d$ is (r, s) approximable if and only if $s \leqslant \frac{1}{r}$.

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 systems. The results of Section 3 are then used to obtain MultiLog Laws in various settings. Namely, Section 4 studies hitting and return times for multi-fold 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* $(\mathbf{\Omega}, \mathcal{F}, \mathbb{P})$. *Given* $r \in \mathbb{N}^*$ *and a family of events* $\{\Omega_{\rho_n}^k\}_{(n,k) \in \mathbb{N}^2; 1 \leqslant k \leqslant n}$, *we let* $N_{\rho_n}^n$ *be the number of times* $k \leqslant n$ *such that* $\Omega_{\rho_n}^k$ *occurs.*

Remark 2.2. *In all our applications it will be the case that*

(2.1) $$\Omega^n_{\rho_1} \subset \Omega^n_{\rho_2} \text{ if } \rho_1 \leqslant \rho_2$$

however, part of our results will not require this condition.

Our goal is to give a criterion that allows to tell when almost surely $N^n_{\rho_n} \geqslant r$ 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 $\mathfrak{s} : \mathbb{N} \circlearrowleft$ such that $\mathfrak{s}(n) \leqslant (\ln n)^2$,
- a function $\hat{\mathfrak{s}} : \mathbb{N} \circlearrowleft$ such that $\varepsilon n \leqslant \hat{\mathfrak{s}}(n) < n(1-q)/(2r)$ for some $0 < q < 1$, and some $0 < \varepsilon < (1-q)/(2r)$,

for which the following holds.

For an arbitrary r-tuple $0 \leqslant k_1 < k_2 \cdots < k_r \leqslant n$ we consider the *separation indices*

$$\mathrm{Sep}_n(k_1, \ldots, k_r) = \mathrm{Card}\,\{j \in \{0, \ldots r - 1\} : k_{j+1} - k_j \geqslant \mathfrak{s}(n)\}, \quad k_0 := 0,$$

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

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

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

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

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

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

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

then

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

Definition 2.3. *For $r \in \mathbb{N}^*$, we say that the events of the family $\{\Omega^k_{\rho_n}\}_{(n,k) \in \mathbb{N}^2; 1 \leqslant k \leqslant n}$ 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 Ω^n_ρ are $2r$–almost independent at all scales if $(M1)_{\bar{r}}$, $(M2)_{\bar{r}}$ are satisfied for $\bar{r} \in [1, 2r]$, and $(M3)_{\bar{r}}$ is satisfied for $\bar{r} \in [1, r]$.*

Theorem 2.4. *Given a family of events* $\{\Omega^k_{\rho_n}\}_{(n,k)\in\mathbb{N}^2;1\leqslant k\leqslant n}$, *define*

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

(a) *If* $\mathbf{S}_r < \infty$, (2.1) *holds, and* $\{\Omega^k_{\rho_n}\}_{(n,k)\in\mathbb{N}^2;1\leqslant k\leqslant n}$ *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* $\{\Omega^k_{\rho_n}\}_{(n,k)\in\mathbb{N}^2;1\leqslant k\leqslant n}$ *are* $2r$–*almost independent at all scales then with probability 1, there are infinitely many* n *such that* $N^n_{\rho_n} \geqslant r$.

Observe that since ρ_n is decreasing and σ is an increasing function we have that

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

when (2.1) holds. Hence, the convergence of $\mathbf{S}_r$ is equivalent to the convergence of $\sum_{n=1}^{\infty} \sigma^r(\rho_n)n^{r-1}$.

Remark 2.5. *An analogous statement has been obtained in* [1] *under different mixing conditions.*

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

$\mathcal{U}_m = \{(k_1, \ldots, k_r) \text{ such that } 2^m < k_1 < k_2 < \cdots < k_r \leqslant 2^{m+1} \text{ and } \widehat{\mathrm{Sep}}_{2^{m+1}}(k_1, \ldots k_r) = r\}$,

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

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

The goal of this section is to prove the following estimates from which it will be easy to derive Theorem 2.4.

Proposition 2.6. *Suppose*

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

If $\{\Omega^k_{\rho_n}\}_{(n,k)\in\mathbb{N}^2;1\leqslant k\leqslant n}$ *are* $2r$–*almost independent at a fixed scale, then there exists constants* $C, c > 0$ *such that*

$$(2.3) \qquad\qquad\qquad \mathbb{P}(\mathcal{A}_m) \leqslant C\left(2^{rm}\sigma(\rho_{2^{m+1}})^r + m^{-10}\right)$$

$$(2.4) \qquad\qquad\qquad \mathbb{P}(\mathcal{D}_m) \geqslant c(2^{rm}\sigma(\rho_{2^{m+1}})^r - m^{-10})$$

If $\{\Omega^k_{\rho_n}\}_{(n,k)\in\mathbb{N}^2;1\leqslant k\leqslant n}$ *are* $2r$–*almost independent at all scales, then for* $m' > m+1$ *we have a sequence* $\theta_m \to 0$ *such that if* $m' - m \geqslant b$ (*given in* $(M3)_r$)

$$(2.5) \qquad \mathbb{P}(\mathcal{D}_m \cap \mathcal{D}_{m'}) \leqslant (\mathbb{P}(\mathcal{D}_m) + m^{-10})(\mathbb{P}(\mathcal{D}_{m'}) + m'^{-10})(1 + \theta_m)$$

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

$$A_{\rho_n}^{k_1,\ldots,k_r} := \bigcap_{j=1}^{r} \Omega_{\rho_n}^{k_j}.$$

With these notations

(2.6) $$\mathcal{A}_m = \bigcup_{0 < k_1 < k_2 < \cdots < k_r \leqslant 2^{m+1}} A_{\rho_{2^m}}^{k_1,\ldots,k_r},$$

(2.7) $$\mathcal{D}_m = \bigcup_{(k_1,\ldots,k_r) \in \mathcal{U}_m} A_{\rho_{2^{m+1}}}^{k_1,\ldots,k_r}.$$

Lemma 2.7. *Fix* $0 < a_1 < a_2 \leqslant 2$. *If* $(M1)_r$ *and* $(M2)_r$ *hold then there exists two sequences* $\delta_n \to 0$, $\eta_n \to 0$ *such that*

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

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

(2.9) $$\sum_{\substack{a_1 n < k_1 < k_2 < \cdots < k_r \leqslant a_2 n \\ \widehat{\mathrm{Sep}}_n(k_1,\ldots k_r)=r}} \mathbb{P}(A_{\rho_n}^{k_1,\ldots,k_r}) \geqslant c_r (n\sigma(\rho_n))^r.$$

Proof. For $m \leqslant r$, denote

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

Note that S_r includes $\frac{n^r}{r!}(1 + \delta_n')$ terms for some sequence $\delta_n' \to 0$ as $n \to \infty$, hence $(M1)_r$ yields

(2.10) $$S_r = \frac{(n\sigma(\rho_n))^r}{r!}(1 + \delta_n'').$$

where $\delta_n'' \to 0$ as $n \to \infty$.

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

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

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 $\hat{\mathfrak{s}}(n) < n(1 - q)/(2r)$. $\qquad \square$

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

$$I_m = \sum_{(k_1,...k_r)\in\mathcal{U}_m} \mathbb{P}(A^{k_1,...,k_r}_{\rho_{2m+1}})$$

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

From (2.7) and Bonferroni inequalities we get that

$$(2.12) \qquad\qquad I_m - J_m \leqslant \mathbb{P}(\mathcal{D}_m) \leqslant I_m$$

Now, (2.9) implies that

$$(2.13) \qquad\qquad I_m \geqslant c_r 2^{r(m+1)}\sigma(\rho_{2m+1})^r.$$

On the other hand, since

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

we get that

$$J_m \leqslant C_r \sum_{l=r+1}^{2r}\sum_{k_1<\cdots<k_l}\mathbb{P}(A^{k_1,...,k_l}_{\rho_{2m+1}}),$$

and (2.8) then implies that

$$(2.14) \qquad\qquad J_m \leqslant C_r(2^{(r+1)m}\sigma(\rho_{2m+1})^{r+1} + m^{-10}).$$

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

Finally, observe that

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

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

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

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

$$\mathbb{P}(\mathcal{D}_m \cap \mathcal{D}_{m'}) \leqslant I_m I_{m'}(1+\varepsilon_m)$$

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

2.3. **Convergent case. Proof of Theorem 2.4 (a).** Suppose that $\mathbf{S}_r < \infty$. Then by monotinicity of $\sigma(\rho_n)$, we have that $n\sigma(\rho_n) \to 0$. By (2.3) of Proposition 2.6 we have that $\sum_m \mathbb{P}(\mathcal{A}_m) < \infty$. By Borel-Cantelli Lemma, with probability one, $\mathcal{A}_m$ happen only finitely many times. Observe that for $n \in (2^m, 2^{m+1}]$, $\{N^n_{\rho_n} \geqslant r\} \subset \mathcal{A}_m$ because $\Omega_{\rho_n} \subset \Omega_{\rho_{2m}}$ for $n \geqslant 2^m$ due to (2.1). Hence with probability one $\{N^n_{\rho_n} \geqslant r\}$ happen only finitely many times.   □

2.4. **Divergent case. Proof of Theorem 2.4 (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 which will be explained at the end of this section.

Claim 2.8. *Let $Z_n = \sum\limits_{m=1}^{n} 1_{\mathcal{D}_m}$. Then there exists 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 $\mathcal{D}_m$ happen. Note that $\mathcal{D}_m \subset \{N^{2^{m+1}}_{\rho_{2^m+1}} \geqslant r\}$, which completes the proof of Theorem 2.4 (b) in case (2.2) holds.

Proof of Claim 2.8. 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

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

Note that

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

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

(2.17) $\mathbb{P}(\mathcal{D}_i \cap \mathcal{D}_j) - \mathbb{P}(\mathcal{D}_i)\mathbb{P}(\mathcal{D}_j) \leqslant \delta\mathbb{P}(\mathcal{D}_i)\mathbb{P}(\mathcal{D}_j) + 2i^{-10}\mathbb{P}(\mathcal{D}_j) + 2j^{-10}\mathbb{P}(\mathcal{D}_i) + 2(ij)^{-10}.$

Split (2.16) into two parts:

(a) Due to (2.17), the terms where $i \geqslant m(\delta)$, $j - i \geqslant m(\delta)$ contribute at most
$$\sum_{i\geqslant m(\delta), j-i\geqslant m(\delta)} \left[\delta\mathbb{P}(\mathcal{D}_i)\mathbb{P}(\mathcal{D}_j) + 2i^{-10}\mathbb{P}(\mathcal{D}_j) + 2j^{-10}\mathbb{P}(\mathcal{D}_i) + 2(ij)^{-10} \right] \leqslant \delta(\mathbb{E}(Z_n))^2 + 8\mathbb{E}(Z_n) + 8.$$

(b) The terms where $i \leqslant m(\delta)$ or $j - i \leqslant m(\delta)$ contribute at most
$$[2m(\delta) + 1] \sum_{j=1}^{n} \mathbb{P}(\mathcal{D}_j) = [2m(\delta) + 1]\mathbb{E}(Z_n).$$

Since $\mathbb{E}(Z_n) \to \infty$, the case (a) dominates for large n giving
$$\limsup_{n\to\infty} \frac{\mathrm{Var}(Z_n)}{(\mathbb{E}(Z_n))^2} \leqslant \delta.$$

Since δ is arbitrary, (2.15) follows.

Let $n_k = \inf\{n : (\mathbb{E}(Z_n))^2 \geqslant k^2\mathrm{Var}(Z_n)\}$. Then by Chebyshev inequality
$$\mathbb{P}\left(|Z_{n_k} - \mathbb{E}(Z_{n_k})| > \delta\mathbb{E}(Z_{n_k}) \right) \leqslant \frac{1}{\delta^2 k^2}.$$

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

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 $\displaystyle\lim_{n\to\infty}n\tilde{\sigma}(\rho_n)=0$

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

Next, we define for each $n\in\mathbb{N}$ and for each $k\leqslant n$, a sequence of events $\{\tilde{\Omega}^k_{\rho_n}\}_{k\leqslant n}$ 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 $\{\tilde{\Omega}^k_{\rho_n}\}_{(n,k)\in\mathbb{N}^2;1\leqslant k\leqslant n}$ thus satisfy $(M1)_r$, $(M2)_r$, and $(M3)_r$ the same way as the events $\{\Omega^k_{\rho_n}\}_{(n,k)\in\mathbb{N}^2;1\leqslant k\leqslant n}$, 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 $\displaystyle\sum_{j=1}^{\infty}(2^j\tilde{\sigma}(\rho_{2^j}))^r=\infty$, we get that, with probability one, more than r events among the events $\{\tilde{\Omega}^k_{\rho_n}\}_{k\leqslant n}$ occurs for infinitely many n. By definition, this implies that with probability one, more than r events among the events $\{\Omega^k_{\rho_n}\}_{k\leqslant n}$ occurs for infinitely many n. The proof of Theorem 2.4 (b) is thus completed. $\square$

2.5. Prescribing some details.
In the remaining part of Section 2 we describe some extensions of Theorem 2.4(b).

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

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

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 $(k_1,\ldots,k_r)$, $(l_1,\ldots,l_r)$ with

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

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

$$\widetilde{(M3)}_r\qquad\qquad\mathbb{P}\left(\left[\bigcap_{\alpha=1}^{r}\Omega^{k_\alpha,i_\alpha}_{\rho_{2^i}}\right]\bigcap\left[\bigcap_{\beta=1}^{r}\Omega^{l_\beta,j_\beta}_{\rho_{2^j}}\right]\right)$$

---

[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 2.4 (b) condition (2.1) is not needed.

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

Theorem 2.9. [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 2.9 is similar to the proof of Theorem 2.4(b). Without the loss of generality we may assume that I_j does not contain 0. Then we consider the following modification of $\mathcal{D}_m$

$$\tilde{\mathcal{D}}_m := \Big\{ \exists 2^m < k_1 < \dots < k_r \leqslant 2^{m+1} \text{ such that } \frac{k_\alpha}{2^{m+1}} \in I_\alpha,$$

$$\Omega_{\rho_{2^{m+1}}}^{k_\alpha, i_\alpha} \text{ happens and } k_{\alpha+1} - k_\alpha \geqslant \hat{\mathfrak{s}}(2^{m+1}), \, 0 \leqslant \alpha \leqslant r - 1 \Big\}.$$

Arguing as in Proposition 2.6 we conclude that $\tilde{\mathcal{D}}_{m_1}$ and $\tilde{\mathcal{D}}_{m_2}$ 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 2.4(b).

2.6. **Poisson regime.**

Theorem 2.10. *Suppose that* $(M1)_r$ *and* $(M2)_r$ *hold for all* r *and that* $\lim_{n \to \infty} n\sigma(\rho_n) = \lambda$. *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 $\mathcal{X}$ denote the Poisson random variable with parameter λ. Below $\binom{m}{r}$ denotes the binomial coefficient $\frac{m!}{r!(m-r)!}$. Since (see e.g. [126] formula (3.4) in section 7.3)

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

Lemma 2.7 implies for each r

$$(2.18) \qquad \lim_{n \to \infty} \mathbb{E}\left(\binom{N_{\rho_n}^n}{r} \right) = \frac{\lambda^r}{r!} = \mathbb{E}\left(\binom{\mathcal{X}}{r} \right).$$

Since this holds for all r we also have that for all r, $\lim_{n \to \infty} \mathbb{E}((N_{\rho_n}^n)^r) = \mathbb{E}(\mathcal{X}^r)$. Since the Poisson distribution is uniquely determined by its moments the result follows. $\square$

Similarly to Borel-Cantelli Lemma, we also have the following extension of Theorem 2.10 in the setting of §2.5. Denote $N_I^{n,i}$ the number of times event $\Omega_{\rho_n}^{k,i}$ occurs with $k/n \in I$. Write $N^{n,i} := N_{[0,1]}^{n,i}$.

---

[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 2.9 to obtain some analogues of the Functional Law of Iterated Logarithm for heavy tailed random variables.

Theorem 2.11. *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 $\{N^{n,i}_{\rho_n}\}^p_{i=1}$ converge in law as $n \to \infty$ to the independent Poisson random variables with parameter λ_i.

Moreover if $I_1, I_2, \ldots I_s$ are disjoint intervals then $\{N^{n,i}_{I_j}\}$, $i = 1\ldots p$, $j = 1\ldots s$ 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 2.10. Namely, similarly to (2.18) we show that for each set $r_{ij} \in \mathbb{N}$ we have

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

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

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. [140, §12.15], [135, §1]. The connection between Borel-Cantelli Lemma and Poisson Limit Theorem is discussed in [51, 57]. The multiple Borel Cantelli Lemma for independent events is proven in [118]. [1] obtains multiple Borel Cantelli Lemma for systems admitting good symbolic dynamics. Extending multiple Borel Cantelli Lemma for more general sequences allows to obtain many new applications, see Sections 4–10 of this paper. We note that separation conditions similar to our have been used in [40, 130] 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 $\rho' \leqslant \rho$, 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

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

Given a sequence $\{\rho_n\}$, we recall that $N_{\rho_n}^n$ denotes the number of times $k \leqslant n$ such that $\Omega_{\rho_n}^k$ (or $\bar{\Omega}_{\rho_n}^k$) occurs. We want to give conditions on the system (f, X, μ) and on the family $\{\Omega_{\rho_n}^k\}_{(n,k)\in\mathbb{N}^2;1\leqslant k\leqslant n}$ or $\{\bar{\Omega}_{\rho_n}^k\}_{(n,k)\in\mathbb{N}^2;1\leqslant k\leqslant n}$, that imply the validity of the dichotomy of Theorem 2.4 for the number of hits $N_{\rho_n}^n$. 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_ρ^k and $\mathbf{v}_j = \bar{\sigma}(\rho_{2^j})$ if we are considering targets of the type $\bar{\Omega}_\rho^k$.

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 $r \geqslant 1$). *Let $\mathbb{B}$ be a space of real valued functions defined over X^{r+1}, with a norm $\|\cdot\|_{\mathbb{B}}$. For $r \geqslant 1$, 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*

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

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

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

$$\left|\int_X A(x, f^{k_1}x, \cdots, f^{k_r}x)d\mu(x) - \int_{X^{r+1}} A(x_0, \cdots, x_r)d\mu(x_0)\cdots d\mu(x_r)\right| \leqslant C\theta^m \|A\|_{\mathbb{B}}.$$

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

Definition 3.2 (Simple admissible targets). *Let Ω_ρ, $\rho \in \mathbb{R}_+^*$, 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 $A_\rho^\pm \in \mathbb{B}$ and*
  (i) $\|A_\rho^\pm\|_\infty \leqslant 2$ *and* $\|A_\rho^\pm\|_{\mathbb{B}} \leqslant \rho^{-\tau}$;
  (ii) $A_\rho^- \leqslant 1_{\Omega_\rho} \leqslant A_\rho^+$;
  (iii) $\mu(A_\rho^+) - \mu(A_\rho^-) \leqslant \sigma(\rho)^{1+\eta}$,

where $\sigma(\rho) = \mu(\Omega_\rho)$.

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

(Poly) $$\rho_n \geqslant n^{-u}, \quad \sigma(\rho_n) \geqslant n^{-u},$$

and

(Mov) $$\forall R, \exists \bar{C} : \ \forall k \in (0, R\ln n), \quad \mu(\Omega_{\rho_n} \cap f^{-k}\Omega_{\rho_n}) \leqslant \bar{C}\sigma(\rho_n)(\ln n)^{-1000r}.$$

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

$$\mu(A_\rho^+) - \mu(\Omega_\rho) \leqslant \mu(\Omega_\rho)^{1+\eta}, \quad \mu(\Omega_\rho) - \mu(A_\rho^-) \leqslant \mu(\Omega_\rho)^{1+\eta}.$$

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

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

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.

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) and (iii) of (Appr), with $\mu(A_\rho^-)$ replaced by $\mu(\Omega_\rho)$. 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

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

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

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

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

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 $1_{\Omega_\rho} \leqslant A_\rho^+$. On the other hand $\|A_\rho^+\|_\infty \leqslant 2$ and $\|A_\rho^+\|_{\mathbb{B}} \leqslant C(\Phi)\rho^{-b\alpha\eta}$, and (i) holds for $\tau = b\alpha\eta + 1$. We turn now to (iii). We observe that with $J_1 = [a_1(\rho) - \rho^b(a_2(\rho) - a_1(\rho)), a_1(\rho)]$ and $J_2 = [a_2(\rho), a_2(\rho) + \rho^b(a_2(\rho) - a_1(\rho))]$,

$$\mu(A_\rho^+) - \mu(1_{\Omega_\rho}) \leqslant 2\mu\left(\{\Phi(x) \in J_1 \cup J_2\}\right)$$

$$\leqslant 4\rho^{\xi'(b+\alpha')}.$$

Hence, if b is chosen sufficiently large we have $\rho > 0$ sufficiently small that $\mu(A_\rho^+) - \mu(1_{\Omega_\rho}) \leqslant \sigma(\rho)^2$.

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. $\square$

---

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

To deal with recurrence, the following definition is useful.

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

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

   (i) $\|\bar{A}_\rho^\pm\|_\infty \leqslant 2$ *and* $\|\bar{A}_\rho^\pm\|_\mathbb{B} \leqslant \rho^{-\tau}$;

   (ii) $\bar{A}_\rho^- \leqslant 1_{\bar{\Omega}_\rho} \leqslant \bar{A}_\rho^+$;

   (iii) *For any fixed x,*

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

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

The sequence $\bar{\Omega}_{\rho_n}$ is said to be composite admissible if

$(\overline{\text{Poly}})$ $$\rho_n \geqslant n^{-u}, \quad \bar{\sigma}(\rho_n) \geqslant n^{-u},$$

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

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

and

$(\overline{\text{Mov}})$ $$\forall k \neq 0, \quad \mu(\bar{\Omega}_{a\rho_n}^k) \leqslant \bar{C}(\ln n)^{-1000r}.$$

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

$$(3.1) \qquad \bar{C}^{-1}\mu\left(\bar{\Omega}_\rho^n\right) \leqslant \mu\left(\bar{A}_\rho^-(x, f^n x)\right) \leqslant \mu\left(\bar{A}_\rho^+(x, f^n x)\right) \leqslant \bar{C}\mu\left(\bar{\Omega}_\rho^n\right).$$

The typical composite targets we will deal with are of the type $d(x,y) < \rho$ or $d(x,y) < \gamma(x)\rho$, where $\gamma(x)$ is related to the local dimension of a smooth measure 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 exists constants $C, \xi, \xi', \xi'' > 0$ and $\Phi : X \times X \to \mathbb{R}$ a (uniformly) Lipschitz function such that*

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

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

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 $\overline{(\mathrm{Appr})}$ *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

$$\int \bar{A}_\rho^+(x,y)d\mu(y) - \mu\left(\{y \in X : \Phi(x,y) \in I\}\right) \leqslant 2\mu\left(\{y \in X : \Phi(x,y) \in J_1 \cup J_2\}\right)$$

$$\leqslant \bar{\sigma}(\rho)^2$$

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

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

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

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

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

which proves $\overline{(\mathrm{Appr})}(iv)$. $\qquad\square$

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 $\{\Omega_{\rho_n}^k\}_{(n,k)\in\mathbb{N}^2; 1\leqslant k\leqslant n}$ (or $\{\bar{\Omega}_{\rho_n}^k\}_{(n,k)\in\mathbb{N}^2; 1\leqslant k\leqslant n}$), that imply the validity of the dichotomy of Theorem 2.4 for the number of hits $N_{\rho_n}^n$. 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_ρ^k and $\mathbf{v}_j = \bar{\sigma}(\rho_{2^j})$ if we are considering targets of the type $\bar{\Omega}_\rho^k$.

Theorem 3.7. *Assume a system* $(f, X, \mu, \mathbb{B})$ *is* $(2r+1)$-*fold exponentially mixing.*[4]
Then
a) If $\{\Omega_{\rho_n}\}$ *is a sequence of simple admissible targets as in Definition 3.2, then the events of the family* $\{\Omega_{\rho_n}^k\}_{(n,k)\in\mathbb{N}^2; 1\leqslant k\leqslant n}$ *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* $\{\bar{\Omega}_{\rho_n}^k\}_{(n,k)\in\mathbb{N}^2; 1\leqslant k\leqslant n}$ *are* $2r$-*almost independent at all scales.*

Hence, Theorem 2.4 implies

---

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

Corollary 3.8. *If the system $(f, X, \mu, \mathbb{B})$ is $(2r + 1)$-fold exponentially mixing, and if $\{\Omega_{\rho_n}\}$ (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 $N^n_{\rho_n} \geqslant r$.*

In fact Theorem 3.7 is a direct consequence of the following Proposition. We accept a convention that $(EM)_k$ for $k \leqslant 0$ is an always satisfied.

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

    (i) If $(\mathrm{EM})_{r-1}$ holds, then $(M1)_r$ is satisfied with the function $\mathfrak{s} : \mathbb{N} \circlearrowleft : \mathfrak{s}(n) = R \ln n$, 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.

    (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 $\{\bar{\Omega}_{\rho_n}\}$ such that (Prod), $(\overline{\mathrm{Poly}})$ and $(\overline{\mathrm{Appr}})$ hold, then, with the function $\bar{\sigma}(\cdot) := \mu \times \mu(\bar{\Omega}.)$:

    (i) If $(\mathrm{EM})_r$ holds, then $(M1)_r$ is satisfied with the function $\mathfrak{s} : \mathbb{N} \circlearrowleft : \mathfrak{s}(n) = R \ln n$, 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}, $\bar{\Omega}_{\rho_n}$, the order of iteration of f, etc.

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

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

$$\leqslant \left(\mu(\Omega_{\rho_n}) + C\mu(\Omega_{\rho_n})^{1+\eta} \right)^r + C\rho_n^{-r\tau}\theta^{\sqrt{R}\ln n},$$

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

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

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

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

and get

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

which yields the RHS of $(M1)_r$ due to $(\overline{\text{Poly}})$ 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 $\text{Sep}(k_1, \ldots, k_r) = r - 1$ 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 $k_i - k_{i-1} \geqslant R \ln n$ 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

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

Therefore (Mov) implies :

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

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

$$\mu\left(\prod_{i=1}^r 1_{\Omega_{\rho_n}} \left(f^{k_i} x\right)\right) \leqslant \mu\left(\prod_{i=1}^r A_{\rho_n}^+ \left(f^{k_i} x\right)\right) = \mu\left(\prod_{i \neq j-1, j} A_{\rho_n}^+ \left(f^{k_i} x\right) B(f^{k_{j-1}} x)\right)$$

$$\leqslant \mu\left(A_{\rho_n}^+\right)^{r-1} \mu(B) + C\rho_n^{-r\tau} L^{\sqrt{R} \ln n} \theta^{R \ln n} \leqslant C\mu(\Omega_{\rho_n})^{r-1} (\ln n)^{-1000r}$$

proving $(M2)_r$.

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

$$\tilde{B}_r(x_0, \cdots, x_{j-1}, x_{j+1}, \cdots, x_r)$$

$$= 1_{\bar{\Omega}_{\rho_n}}(x_0, x_1) \cdots 1_{\bar{\Omega}_{\rho_n}}(x_0, x_{j-1}) 1_{\bar{\Omega}_{a\rho_n}^{k_j - k_{j-1}}}(x_{j-1}) 1_{\bar{\Omega}_{\rho_n}}(x_0, x_{j+1}) \cdots 1_{\bar{\Omega}_{\rho_n}}(x_0, x_r),$$

$$\hat{B}_r(x_0, \cdots, x_{j-1}, x_{j+1}, \cdots, x_r)$$

$$= \bar{A}_{\rho_n}^+(x_0, x_1) \cdots \bar{A}_{\rho_n}^+(x_0, x_{j-1}) \bar{A}_{a\rho_n}^+(x_{j-1}, f^{k_j - k_{j-1}} x_{j-1}) \bar{A}_{\rho_n}^+(x_0, x_{j+1}) \cdots \bar{A}_{\rho_n}^+(x_0, x_r).$$

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

$$\mu\left(\bigcap_{j=1}^r \bar{\Omega}_{\rho_n}^{k_j}\right) \leqslant \mu\left(\tilde{B}_r(x, \cdots, f^{k_{j-1}} x, f^{k_{j+1}} x, \cdots, f^{k_r} x)\right)$$

$$\leqslant \mu\left(\hat{B}_r(x, \cdots, f^{k_{j-1}} x, f^{k_{j+1}} x, \cdots, f^{k_r} x)\right)$$

$$\leqslant \mu^r(\hat{B}_r) + \bar{C}\rho_n^{-r\tau} L^{\sqrt{R} \ln n} \theta^{R \ln n}.$$

Integrating with respect to all variables except for 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

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

which by (3.1) gives

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

Therefore, $(M2)_r$ follows from $(\overline{\text{Mov}})$, provided R is sufficiently large.

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 $2^i < k_1 < \cdots < k_r \leqslant 2^{i+1}$, we obtain from (Prod), (Gr), (Appr), (Poly) and $(\mathrm{EM})_r$, that $\|B\|_{\mathbb{B}} \leqslant CL^{r2^{i+1}}$. Thus

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

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

$$= \mu\left(B(x)\left(\prod_{\beta=1}^{r} A^+_{\rho_{2^j}}(f^{l_\beta}x)\right)\right) \leqslant \mu(B)\mu\left(A^+_{\rho_{2^j}}\right)^r + CL^{r2^{i+1}}\rho_{2^i}^{-r\tau}\rho_{2^j}^{-r\tau}\theta^{2^j\varepsilon}.$$

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 $\bar{\Omega}_{\rho_n}$. Consider

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

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

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

$$\leqslant \mu\left(\prod_{\alpha=1}^{r}\bar{A}^+_{\rho_{2^i}}(x, f^{k_\alpha}x)\right)\left(\bar{\sigma}(\rho_{2^j}) + \bar{\sigma}(\rho_{2^j})^{1+\eta}\right)^r + CL^{r2^{i+1}}\rho_{2^i}^{-r\tau}\rho_{2^j}^{-r\tau}\theta^{2^j\varepsilon}.$$

Using $(M1)_r$ again we observe that

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

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

Remark 3.10. *In fact, analyzing the proof of Theorem 3.7 we see that the composite targets $(\overline{\mathrm{Appr}})$(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*

$$(3.2a) \qquad \int\ldots\int\left(\prod_{j=1}^{r}\bar{A}^+(x, y_j)d\mu(y_j)\right)d\mu(x) = \sigma_r(\rho)(1 + O(\sigma^\eta(\rho)),$$

$$(3.2b) \qquad \int\ldots\int\left(\prod_{j=1}^{r}\bar{A}^-(x, y_j)d\mu(y_j)\right)d\mu(x) = \sigma_r(\rho)(1 + O(\sigma^\eta(\rho)).$$

We shall call the composite targets satisfying $(\overline{\mathrm{Mov}})$*,* $(\overline{\mathrm{Sub}})$*,* $(\overline{\mathrm{Poly}})$ *as well as* $(\overline{\mathrm{Appr}})$ *with condition* (iii) *replaced by* (3.2a)–(3.2b) *weakly admissible.*

3.3. **Notes.** There is also a vast literature on Borel-Cantelli Lemmas for dynamical systems starting with [123]. Some representative examples dealing with hyperbolic systems are [5, 35, 56, 68, 71, 73, 74, 80, 88, 102] while [29, 30, 88, 95, 96, 97, 108, 136] deal with systems of zero entropy. The later cases are more complicated as counterexamples in [55, 64] show. Survey [7] reviews the results obtained up to 2009 and contains many applications, some of which parallel the results obtained in Sections 4–9 of the present paper. We refer the reader to Appendix A for more background on multiple exponential mixing and for examples of dynamical systems which enjoy this property. We note that limit theorems for smooth systems which are only assumed to be multiply exponentially mixing (but without any additional assumptions) are considered in [21, 33, 133]. [61] 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 will assume that f is a smooth diffeomorphism of a compact $d-$dimensional Riemannian manifold M preserving a smooth measure μ. From now on, we take $\mathbb{B}$ in Definition 3.1 to be the space of Lipschitz observables defined over M^{d+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

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

$$\text{(4.1b)} \qquad \limsup_{n\to\infty} \frac{|\ln d_n^{(1)}(x, y)|}{\ln n} = \frac{1}{d}.$$

In particular, the following results are known.

Theorem 4.1. *(a) If a smooth system* (f, M, μ) *has superpolynomial decay of correlations for Lipschitz observables, that is,*

$$|\mu(A(x)B(f^n x)) - \mu(A)\mu(B)| \leqslant a(n)\|A\|_{Lip}\|B\|_{Lip} \quad where \quad \forall s \lim_{n\to\infty} n^s a(n) = 0,$$

then for **all** x *(4.1b) holds for a.e.* y*. If in addition,* f *has positive entropy, then* (4.1a) *holds for a.e.* x*.*

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

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

In part (a), (4.1a) is proven in [128, Theorem 1] and (4.1b) is proven in [61, Theorem 4]. Part (b) is proven in [44, Theorem 7].

Question 4.2. Suppose that (f, μ) is exponentially mixing then (4.2) 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 of (4.2) 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*

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

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

Theorem 4.4. *Suppose that $(f, M, \mu, \mathbb{B})$ is $(2r + 1)$-fold exponentially mixing.[5] Then*
(a) $\mu(\mathcal{G}_r) = 1$; (b) $\mu(\bar{\mathcal{G}}_r) = 1$.

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*

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

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

Theorem 4.6. *Suppose that the periodic points of f are dense. Then*
(a) If $\mathcal{G}_1 = M$, then $\mathcal{H}$ contains a G_δ dense set.
(b) For any $\zeta : \mathbb{N} \to \mathbb{N}^*$, $\bar{\mathcal{H}}_\zeta$ contains a G_δ dense set.

Thus for $r \geqslant 2$ topologically typical points do **not** belong to $\mathcal{G}_r$ or $\bar{\mathcal{G}}_r$.

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.

---

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

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

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

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

Theorem 4.7. *For λ-a.e. $\alpha \in \mathbb{T}^d$, the system $(T_\alpha, \mathbb{T}^d, \lambda)$, satisfies*
a) $\lambda(\mathcal{G}_1) = 1$ and $\lambda(\mathcal{E}_r) = 1$ for $r \geq 2$;
b) $\bar{\mathcal{E}}_r = M$ for all $r \geq 1$.

The proof requires different techniques from the rest of the results of this section, that are related to homogeneous dynamics on the space of lattices, so it will be given in Section 9 after we introduce the necessary tools.

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 $i \geq 0$, 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 $t \in [t_t^-, t_i^+]$. Let $d_n^{(r)}(x,y)$ be the r-th minimum of

$$(4.3) \qquad d(x, \phi^{t_1}(y)), \ldots, d(x, \phi^{t_k}(y)), \quad t_k \leq n < t_{k+1}.$$

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 $d+1$.

Theorem 4.8. *Suppose that the smooth system $(\phi, M, \mu, \mathbb{B})$ is $(2r+1)$-fold exponentially mixing. Then*
a) $\mu(\mathcal{G}_r) = 1$;
b) $\mu(\bar{\mathcal{G}}_r) = 1$.
    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 $\zeta : \mathbb{N} \to \mathbb{N}^*$, $\bar{\mathcal{H}}_\zeta$ contains a G_δ dense set.

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

$$(4.4) \qquad \mu(B(x, \rho)) = \gamma(x)\rho^d + O\left(\rho^{d+1}\right),$$

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

Given $x \in M$, let

$$(4.5) \qquad \Omega_{x,\rho} = \{y : d(x, y) \leq \rho\}$$

and[6]

$$(4.6) \qquad \bar{\Omega}_\rho = \left\{ (x, y) : d(x, y) \leqslant \frac{\rho}{(\gamma(x))^{1/d}} \right\}$$

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

For $s \geqslant 0$, we let $\rho_n = n^{-1/d} \ln^{-s} n$, and recall that $N_{\rho_n}^n$ denotes the number of times $k \leqslant n$ such that Ω_{x,ρ_n}^k (or $\bar{\Omega}_{\rho_n}^k$) occurs.

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

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

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

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

(b) If $s \leqslant \frac{1}{rd}$, then for μ-a.e. x, there are infinitely many n such that $N_{\rho_n}^n \geqslant r$.

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 (4.4) that $\mathbf{S}_r = \infty$ if and only if $s \leqslant \frac{1}{rd}$.

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

Proposition 4.9. *(a) For μ-a.e. x the targets $\{\Omega_{x,\rho_n}\}$ are simple admissible targets.*
(b) The targets $\{\bar{\Omega}_{\rho_n}\}$ are composite admissible targets.

The rest of this section is devoted to the
Proof of Proposition 4.9.

Observe first that with the definition of ρ_n and (4.4), we have that (Poly) and ($\overline{\text{Poly}}$) hold for every x for the target sequences $\{\Omega_{x,\rho_n}\}$ 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 4.10. *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 (4.4).

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

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

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 4.10 is proved. $\qquad\square$

---

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

Next we prove of 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 4.11 (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 $\rho < \rho_0$ for all $n \leqslant K|\ln \rho|$ we have*

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

Definition 4.12 (Slowly recurrent system). *Call the system (f, M, μ) slowly recurrent if for each $A > 0$ $\exists \rho_0$ such that for all $\rho < \rho_0$ for all $n \in \mathbb{N}^*$ we have*

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

Lemma 4.13. *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 $k \geqslant B \ln|\ln \rho|$, take $\hat{\rho} = |\ln \rho|^{-A}$. By 2-fold exponential mixing, we get

$$(4.7) \quad \mu(x : d(x, f^k x) \leqslant \rho) \leqslant \mu(x : d(x, f^k x) \leqslant \hat{\rho})$$
$$\leqslant \mu\left(\bar{A}^+_{\hat{\rho}}(x, f^k x)\right) \leqslant C\left(\hat{\rho}^d + \hat{\rho}^{d+d\eta} + \hat{\rho}^{-\tau}\theta^k\right) \leqslant |\ln \rho|^{-2A},$$

provided ρ is sufficiently small.

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

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

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

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

provided ρ is sufficiently small. But $kL \geqslant B \ln|\ln \sqrt{\rho}|$, hence (4.7) applies and we get

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

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

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

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

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

For such j we get from Markov inequality

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

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

Finally, (a) and (b) clearly follow from ii) and i) respectively. Lemma 4.13 is thus proved. $\qquad\square$

With Lemmas 4.10 and 4.13, the proof of Proposition 4.9 is finished. $\qquad\square$

Proof of Theorem 4.4. Theorem 4.4 directly follows from Proposition 4.9 and Corollary 3.8. $\qquad\square$

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,

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

Since $d_n^{(r)}(x,y) \geqslant d_n^{(1)}(x,y)$, it follows that for any $x \in M$, any $r \geqslant 1$, and almost every y

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

To prove the opposite inequality let

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

We have that

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

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 (4.8), there exists $n \geqslant \exp \circ \exp(\Lambda + pr)$ 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

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

Then

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

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

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

because we took $n \geqslant \exp \circ \exp(\Lambda + pr)$. 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 $\bar{\mathcal{H}}_\zeta \subset \bigcap_l \bigcup_m \mathcal{A}_{m,l}$. But $\mathcal{A}_{m,l}$ is open and $\bigcup_m \mathcal{A}_{m,l}$ clearly contains the periodic points. Part (b) is thus proved. $\square$

4.4. The case of flows. Proof of Theorem 4.8.

The proof proceeds in the same way as for diffeomorphisms with minimal modifications that we now explain. First, we need to modify the targets

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

and

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

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

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

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

To prove a) and b) of Theorem 4.8 we can apply Corollary 3.8 to the smooth system $(\phi, M, \mu, \mathbb{B})$ and to the targets $\Omega_{x,\rho}^n$ and $\bar{\Omega}_\rho^n$. For this, we just need to see that the targets are admissible targets. This can be checked as in the proof of Proposition 4.9, with very minor differences. Let us check for instance that $(\overline{\text{Sub}})$ holds for $\bar{\Omega}_\rho^n$. Note that when $x \in \bar{\Omega}_\rho^{n_1} \cap \bar{\Omega}_\rho^{n_2}$ for $n_1 < n_2$, we have some s_1, $s_2 \in [0, 1]$ such that

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

Hence

$$d(\phi^{n_1}x, \phi^{n_2+s_2-s_1}x) \leqslant \max_{s\in[-1,0]} \|\phi^s\|_{C^1} d(\phi^{n_1+s_1}x, \phi^{n_2+s_2}x)$$

$$\leqslant \max_{s\in[-1,1]} \|\phi^s\|_{C^1} \frac{2\rho}{(\gamma(x))^{1/d}} \leqslant \frac{a\rho}{(\gamma(\phi^{n_1}x))^{1/d}}$$

for some $a > 0$. It follows that $\bar{\Omega}_\rho^{n_1} \cap \bar{\Omega}_\rho^{n_2} \subset \phi^{-n_1}\bar{\Omega}_{a\rho}^{n_2-n_1}$, which is $(\overline{\text{Sub}})$. [7] 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 $n \leqslant K|\ln\rho|$ we have

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

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

$$\mu\left(\bar{\Omega}_\rho^n\right) \leqslant |\ln\rho|^{-A}.$$

The same proof of Lemma 4.13 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.

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 $\mathcal{H}_r$ and $\bar{\mathcal{H}}_r$ and then use the genericity argument. ▫

4.5. **Notes.** Many authors obtain Logarithm Law (4.1b) for hitting times as a consequence of dynamical Borel-Cantelli Lemmas. See [35, 44, 60, 80] and references wherein. [61] also studies return times. We note that [61] works under much weaker conditions than those imposed in the present paper, however, his results are valid only for $r = 1$ (the first visit).

[88, 94, 99] study the recurrence problem when the lim sup in (4.1b) is replaced by lim inf. In particular, [99] proves that for several expanding maps the

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

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 $r \geqslant 2$. Logarithm Laws for unipotent flows were obtained in [9, 10, 65, 88]. It is not known which kind of MultiLog Laws hold for such flows.

---

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

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 (4.5), for almost every x, and for the targets $\bar{\Omega}_\rho$ given by (4.6). Moreover, we have that $\lim_{\rho \to 0} \rho^{-d}\sigma(\rho) = \gamma(x)$ and $\lim_{\rho \to 0} \rho^{-d}\bar{\sigma}(\rho) = 1$, where $\sigma(\rho) = \mu(\Omega_{\rho,x})$, $\bar{\sigma}(\rho) = (\mu \times \mu)(\bar{\Omega}_\rho)$. Accordingly Theorem 2.10 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*

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

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

(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 $B\left(x, \dfrac{\rho}{\gamma^{1/d}(x)}\right)$ converges to a Poisson distribution with parameter τ as $\rho \to 0$.

Proof. All the results except for Poisson limit for (5.1) follows from Theorem 2.10. 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 $d(x, f^k y) \in [r_j^-\rho, r_j^+\rho]$ are converging to independent Poisson random variables with parameters

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

But this follows from Theorem 2.11. The latter theorem can be applied since $\widetilde{(M1)}_r$ follows from Property (Appr) of the targets

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

that holds due to Lemma 3.4. $\qquad\square$

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 to remove an annoying factor $\gamma^{1/d}(x)$ from the denominator. Regarding the first question we have

Conjecture 5.2. *If f is exponentially mixing then the conclusion of Proposition 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

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

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 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 $r \geqslant 2$ then*

(a) For almost all x we have that if $\tau_\varepsilon(y)$ is an 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

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

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). $\square$

Proof of Theorem 5.3. Consider the targets

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

and let $\hat{\Omega}_\rho^k = \{x : (x, f^k x) \in \hat{\Omega}_\rho\}$. Note that $(M2)_r$ for $\bar{\Omega}_\rho^k$ 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 $k_1, k_2, \ldots, k_r$,

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

Namely, note that

$$\int 1_{\hat{\Omega}_\rho}(x_0, x_1) \ldots 1_{\hat{\Omega}_\rho}(x_0, x_r) d\mu(x_0) d\mu(x_1) \ldots d\mu(x_r)$$

$$= \int \mu^r(B(x_0, \rho)) d\mu(x_0) = \rho^{rd}(1 + O(\rho)) \int_M \gamma^r(x_0) d\mu(x_0).$$

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 $0 \leqslant j \leqslant r - 1$, then

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

$$\leqslant \left(\rho^d + C\rho^{d(1+\eta)}\right)^r \int_M \gamma^r(z) d\mu(z) + C\rho^{-r\sigma}\theta^{R|\ln \rho|},$$

and, likewise,

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

Taking R large we obtain (5.3).

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

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

where

$$N_{\rho,\tau,x} = \operatorname{Card}\left\{ k \leqslant \tau\rho^{-d} : d(x, f^k x) \leqslant \rho \right\}.$$

Since the RHS coincides with factorial moments of the Poisson mixture from (5.2), the result follows. $\square$

5.1. **Notes.** Early works on Poisson Limit Theorems for dynamical systems include [36, 42, 81, 82, 83, 124]. [31, 75, 79, 125] 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. [44] obtains Poisson Limit Theorem for partially hyperbolic systems. Some of those papers, including [28, 44, 75, 78] show that in various settings is 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, 76]. The Poisson limit theorems for flows are obtained in [116, 120]. 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 [44, 58, 59]. A mixed exponential distribution for a return time for dynamical systems similar to Corollary 5.4 has been obtained in [38] in a symbolic setting. For more discussion of the distribution of the entry times to small measure sets we refer the readers to [37, 87, 129, 143] and references wherein. We also refer to Section 10 for the 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

$$(6.1) \qquad P(g) = \int g d\mu + h_\mu(f) = 0.$$

This means (see [131] 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

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

then

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

We denote

(6.2) $$f_u = \ln|f'|,$$

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

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

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 [110] that the limit exists for μ-a.e. x and

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

where the last step relies on (6.1).

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

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

It is known (see e.g. [121]) that μ is conformal if and only if g can be represented in the form

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

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

Denote

(6.3) $$\psi(x) = g(x) + \mathbf{d}f_u(x),$$

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

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

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*

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

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

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

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

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 $r \geqslant 2$.*

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,

$$(6.7) \qquad\qquad \rho^a \leqslant \mu(B(x, \rho)) \leqslant \rho^b.$$

We also need the fact that Gibbs measures are Alhfors regular, that is there is a
constant R such that for each x, ρ we have

$$(6.8) \qquad\qquad \mu(B(x, 4\rho)) \leqslant R\mu(B(x, \rho)).$$

We recall the proofs of (6.7) and (6.8) 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

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

We use the notation $\Omega_{x,\rho}^k$ for the event $1_{\Omega_{x,\rho}} \circ f^k$. We also recall the notation $\bar{\Omega}_\rho^k = \{x : (x, f^k x) \in \bar{\Omega}_\rho\}$. In the sequel we will always assume that $\{\rho_n\}$ 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 4.13
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 $\mu(B(y, 1/2^j)) \leqslant C\mu(B(x, 1/2^j))$ for any $x, y \in M$, while this was used in the proof of Lemma 4.13.

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 (6.7) and (6.8), 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 4.13 to overcome the fact that it does not hold anymore that $\mu(B(y, 1/2^j)) \leqslant C\mu(B(x, 1/2^j))$ 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

(6.9)
$$\mu(B(x, \rho) \cap f^{-k}B(x, \rho)) \leqslant \mu(B(x, \rho))^{1+\eta}.$$

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 $B(x, \rho)$, $\int A_\rho^+ d\mu \leqslant 2\mu(B(x, \rho))$ and $\|A_\rho^+\|_{Lip} \leqslant C\rho^{-\tau}$ for some $\tau = \tau(\mu)$. Let $\hat{\rho} = \rho^\sigma$ where σ is a small constant. Then (A.3) gives

$$\mu(B(x, \rho) \cap f^{-k}B(x, \rho)) \leqslant \int A_{\hat{\rho}}^+ (A_\rho^+ \circ f^k) d\mu$$

$$\leqslant 4\mu(B(x, \rho))\mu(B(x, \hat{\rho})) + 2C\bar{\theta}^k \hat{\rho}^{-\tau}\mu(B(x, \rho)) \leqslant C\mu(B(x, \rho)) \left(\rho^{\sigma b} + \rho^{\varepsilon|\ln \bar{\theta}|} \rho^{-\tau\sigma} \right)$$

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) $k \leqslant \varepsilon|\ln \rho|$. 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. $\qquad\square$

Proposition 6.5. *([15, 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 $\rho \leqslant \rho_0(x)$, and all $0 < k \leqslant \frac{1}{2L}|\ln \rho|$, we have $T^{-k}B(x, \rho) \cap B(x, \rho) = \varnothing$.*

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 4.9 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 (4.6) are weakly admissible in the sense of Remark 3.10.*

6.4. **The conformal case.**

Proof of Theorem 6.1 (a). We take $\rho_n = n^{-1/\mathbf{d}} \ln^{-s} n$. Due to 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. $\qquad\square$

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

6.5.1. *The iterated logarithm law for hitting times : Proof of (6.6) of Theorem 6.1 (b).* For $\varepsilon > 0$ and $c > 0$ arbitrary let

$$(6.10) \qquad \rho_n = \rho_n(c) = \frac{1}{n^{1/\mathbf{d}}} \exp\left(-c\sqrt{2(\ln n)(\ln\ln\ln n)}\right).$$

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

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

then

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

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 $n \geqslant n(x)$, we have

$$(6.11) \qquad\qquad \mu\left(\Omega_{x,\rho_n}\right) \leqslant \tilde{\vartheta}_{\varepsilon,c}^+(n).$$

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

$$(6.12) \qquad\qquad \mu\left(\Omega_{x,\rho_{n_l}}\right) \geqslant \tilde{\vartheta}_{\varepsilon,c}^-(n_l).$$

Now it follows that for any $r \geqslant 1$, $S_r = \sum_{k=1}^{\infty}\left(2^k\mu(\Omega_{x,\rho_{2^k}})\right)^r$ 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 (6.6) follows from Proposition 6.2, Lemma 6.4 and Corollary 3.8 $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

6.5.2. *The iterated logarithm law for return times: Proof of the upper bound in (6.5).* Now we turn to the proof of

$$(6.13) \qquad\qquad \limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,x)| - \frac{1}{\mathbf{d}}\ln n}{\sqrt{2(\ln n)(\ln\ln\ln n)}} \leqslant \frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}.$$

Since $d_n^{(r)}(x,x) \geqslant d_n^{(1)}(x,x)$, we only need to show (6.13) 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

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

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

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

$$\mathcal{A}_k = \{x : x \text{ is } n\text{−good for } n \geqslant N_k \text{ but } d(x, f^m x) \leqslant r_{N_k} \text{ for some } m = \bar{\varepsilon} \ln N_k, \ldots, N_k\}.$$

Let $\mathcal{X}_k = \{x_{j,k}\}_{j=1}^{l_k}$ 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 $m \leqslant K \ln N_k$, (6.9) is telling us that

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

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

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

Summing those estimate for we obtain

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

for some $\kappa = \kappa(\bar{\varepsilon}) > 0$. Since $B(x_{j,k}, r_{N_k}/2)$ are disjoint for different j, by (6.8) we conclude that

$$\sum_j \mu\left(B(x_{j,k}, 2r_{N_k})\right) \leqslant R \sum_j \mu\left(B(x_{j,k}, r_{N_k}/2)\right) \leqslant R.$$

It follows that

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

Now the result follows from the classical Borel Cantelli Lemma.

6.5.3. *The law of iterated logarithm for return times: Proof of the lower bound in* (6.5). Here we prove that

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

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

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

and

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

Let

$$\mathcal{A}_j = \left\{x : \text{Card}\left\{k_{j-1}(x) \leqslant k \leqslant k_j(x) : f^k x \in P_{n_j}(x)\right\} \geqslant r\right\}.$$

Then

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

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

Denote by $\mathbb{P}(\cdot|\cdot)$ the conditional probability and $\mathcal{F}_j = \mathcal{B}\left(\mathcal{P}_{k_1}, \cdots, \mathcal{P}_{k_j}\right)$ 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. *([140, §12.15]) If* $\sum_j \mathbb{P}\left(\mathcal{A}_{j+1}|\mathcal{F}_j\right) = \infty$ *a.s. then* $\mathcal{A}_j$ *happen infinitely many times almost surely.*

Hence (6.14) 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* $\mathbb{P}\left(\mathcal{A}_{j+1}|\mathcal{F}_j\right) \geqslant c^*$ *for all* $j \geqslant j_0$.

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

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

by bounded distortion property. Note that

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

By construction

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

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 2.10 to get

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

proving the lemma. $\qquad\qquad\square$

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 [24, 25, 32, 38, 77, 85, 119, 127, 137]. In particular, [72] 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 [14], while fluctuations are determined only for a limited class of systems [111]. 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 $\mathcal{Q}$ be a finite volume non-compact $(d+1)$-dimensional manifold of curvature -1. Let $S\mathcal{Q}$ denote the unit tangent bundle to $\mathcal{Q}$. For $(q,v) \in S\mathcal{Q}$, 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, that is $g^t(\gamma(0), \dot{\gamma}(0)) = (\gamma(t), \dot{\gamma}(t))$. g^t preserves the Liouville measure μ. Fix a reference point $O \in \mathcal{Q}$ and let

$D(q, v, t) = \text{dist}(O, \gamma(t))$. According to Sullivan's Logarithm Law for excursions [135] for μ-a.e. $(q, v) \in S\mathcal{Q}$, it holds that

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

In fact, the Borel Cantelli Lemma of [135] also shows that

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

Here we present a multiple excursions version of (7.2). Recall ([17, Proposition D.3.12]) that $\mathcal{Q}$ admits a decomposition $\mathcal{Q} = \mathcal{K} \bigcup \left(\bigcup_{j=1}^{p} \mathcal{C}_j \right)$ where $\mathcal{K}$ is a compact set and $\mathcal{C}_j$ are cusps. Moreover 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 $\mathcal{K}$ in between. We note that[8] for each $q_0 = (x_0, y_0) \in \mathcal{C}_i$ 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 $\gamma(q, v, t) \in \mathcal{K}$ and $h(q, v, t) = \ln y(t)$ if $\gamma(q, v, t) = (x(t), y(t)) \in \mathcal{C}_i$. It is easy to see using the triangle inequality that there exists a constant C such that

$$|D(q, v, t) - h(q, v, t)| \leqslant C.$$

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

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

Note that (7.2) is implied by

$$(7.3) \qquad \limsup_{T \to \infty} \frac{H^{(1)}(q, v, T) - \frac{1}{d} \ln T}{\ln \ln T} = \frac{1}{d}.$$

Here we prove the following multiple excursions version of (7.3).

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.

---

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

Corollary 7.2. *There is a constant a_i such that for each $\mathfrak{h}$ the following holds. Suppose that (q, v) is uniformly distributed on SQ. Then the number of excursions in the cusp 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 $r \geqslant 1$, we have

$$(7.4) \qquad \lim_{T \to \infty} \mu\left(H_i^{(r)}(q, v, T) < \frac{\ln T}{d} + \mathfrak{h} \right) = \sum_{l=0}^{r-1} \frac{(a_i e^{-d\mathfrak{h}})^l}{l!} \exp\left(-a_i e^{-d\mathfrak{h}} \right).$$

In particular, taking $r = 1$ in (7.4) we obtain

Corollary 7.3. *(Gumbel distribution for the maximal excursion) If (q, v) is uniformly distributed on SQ. Let $H_i^{(1)}(q, v, T)$ denote the maximal height reached by $\gamma(q, v, t)$ up to time T inside cusp 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. **MultiLog law for geodesic excursions.**

In this section we prove Theorem 7.1. We first need to discuss the probability of having an excursion reaching a given level. To this end 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 $R \geqslant R_0$ is equivalent to the condition $\sin \theta \leqslant \frac{y(n^*)}{R_0}$.

Definition 7.4. *Given H we consider the set $\mathcal{A}_{i,H}$ which consists of points $(q, v) \in C_i$ 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 $(q, v) \in C_i$ 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 $h \geqslant H$, consuming for this a time comparable to h. Moreover condition (ii) on the angle is a necessary and sufficient condition for the excursion to reach height H.

We also introduce

$$(7.5) \qquad\qquad\qquad \mathcal{A}_H = \bigcup_i \mathcal{A}_{i,H}.$$

It is a basic fact (e.g. see the proof of Theorem 6 in [135])

$$(7.6) \qquad\qquad\qquad \mu(\mathcal{A}_{i,H}) = a_i e^{-dH}(1 + o(1)).$$

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

$$(7.7) \qquad \Omega_\rho^k = g^{-k} \mathcal{A}_{-\ln \rho}.$$

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 $s \geqslant 0$, we let $\rho_n = n^{-1/d} \ln^{-s} n$, and 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 the following:

(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 $s \leqslant \frac{1}{rd}$, then for μ-a.e. (q, v), there are infinitely many n such that $N_{\rho_n}^{\frac{n}{2}} \geqslant r$.[9]

Observe that by (7.6), we have that

$$(7.8) \qquad \mu(\Omega_{\rho_n}) \in \left[C^{-1} n^{-1} \ln^{-sd} n, C n^{-1} \ln^{-sd} n \right]$$

With the notation $\mathbf{S}_r = \sum_{j=1}^\infty (2^j \mathbf{v}_j)^r$ where $\mathbf{v}_j = \sigma(\rho_{2^j})$ we see from (7.8) that $\mathbf{S}_r = \infty$ if and only if $s \leqslant \frac{1}{rd}$. We want thus to apply Corollary 3.8, but first we need to verify its conditions.

The system $(g^1, S\mathcal{Q}, \mu, \mathbb{B})$ is r-fold exponentially mixing for every $r \geqslant 2$ in the sense of Definition 3.1. Indeed (Prod) an (Gr) are clear, while $(EM)_r$ follows from Theorem 1.1 of [19] (see also Theorem 1.2 of [103]) 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 $\limsup$ of Theorem 7.1 follows from Corollary 3.8 and the fact that $\mathbf{S}_r = \infty$ if and only if $s \leqslant \frac{1}{rd}$. $\qquad \square$

Proof of Proposition 7.5. First, the definition of ρ_n and (7.8) imply (Poly). Next, the first time $\bar{t}(q, v) \geqslant 0$ 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 benign extra condition that $\bar{t}(q, v) \in [0, 1]$.

It remains to prove (Mov). We denote $\mathcal{A}_H^n = g^{-n} \mathcal{A}_H$. (Mov) is an immediate consequence from the following quasi-independence result on the excursions. Similar quasi-independence results are obtained in [135, 122], and we will give a proof adapted to our setting for completeness.

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

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

---

[9]We introduce a factor $1/2$ to make sure the last excursion that starts before $n/2$ finishes before n. Here, we are using the control on the excursion time that is comparable to the excursions height $\ln n \ll n/2$.

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

The rest of this section is devoted to the

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

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

Thus $\mathcal{B}_{H,n_1,\bar{n}}$ 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 $\mathcal{A}_H^{n_1} = \bigcup_{\bar{n} > n_1} \mathcal{B}_{H,n_1,\bar{n}}$. Note that $\bar{n} - n_1 \geqslant H$.

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

$$(7.9) \qquad\qquad \sum_{\bar{n}} \mu(\tilde{\mathcal{B}}_{H,n_1,\bar{n}}) \leqslant 3\mu(\mathcal{A}_{H-1}).$$

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

$$(7.10) \qquad\qquad \mu(\mathcal{A}_H^{n_1} \cap \mathcal{A}_H^{n_2}) \leqslant \sum_{n_1 < \bar{n} < n_2} \mu\left(\tilde{\mathcal{B}}_{H,n_1,\bar{n}} \cap \mathcal{A}_H^{n_2}\right).$$

We claim that for some constant $C \geqslant 1$ for each $n_1 < \bar{n} < n_2$

$$(7.11) \qquad\qquad \mu\left(\mathcal{B}_{H,n_1,\bar{n}} \cap \mathcal{A}_H^{n_2}\right) \leqslant C\mu(\tilde{\mathcal{B}}_{H,n_1,\bar{n}})\mu(\mathcal{A}_H).$$

Now, (7.6), (7.9), (7.10) and (7.11), imply the estimate of Lemma 7.6.

It remains to establish (7.11). To this end, fix a large $\bar{H}$ and partition a small neighborhood $\mathcal{U}$ of $\tilde{\mathcal{A}}_{\bar{H}}$ into unstable cubes of size δ (these are nice unstable cubes around points that are close to the compact region $\mathcal{K}$). For $H \geqslant \bar{H}$, let

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

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

$$(7.12) \qquad\qquad \mathcal{B}_{H,n_1,\bar{n}} \subset g^{-\bar{n}}\hat{\mathcal{B}}_{H,n_1,\bar{n}} \subset \tilde{\mathcal{B}}_{H,n_1,\bar{n}}.$$

Thus

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

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

$$(7.13) \qquad\qquad \mu\left(\hat{\mathcal{B}}_{H,n_1,\bar{n}} \cap \mathcal{A}_H^{n^*}\right) \leqslant C\mu(\hat{\mathcal{B}}_{H,n_1,\bar{n}})\mu(\mathcal{A}_H).$$

Indeed (7.12) and (7.13) imply (7.11). By construction, $\hat{\mathcal{B}}_{H,n_1,\bar{n}}$ is partitioned into nice unstable cubes of size δ. It suffices to show that for any such cube $\mathcal{W}$ we have

$$(7.14) \qquad\qquad \mu(\mathcal{A}_H^{n^*}|\mathcal{W}) \leqslant Ce^{-dH}$$

where $\mu(\cdot|\cdot)$ denotes the conditional expectation. Let $Q = \bigcup_{x \in \mathcal{W}} \bigcup_{|t| < \delta} \mathcal{W}^s(g^t x, \delta)$, where $\mathcal{W}^s(y, \delta)$ 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 $x \in \mathcal{W}$ and $t \in [-\delta, \delta]$ such that $y \in \mathcal{W}^s(g^t x, \delta)$. In addition if $g^{n^*} x \in \mathcal{A}_H$ then $g^{n^*} y \in \mathcal{A}_{H-1}$. Since the measure of Q is bounded from below uniformly in $\mathcal{W} \subset \mathcal{U}$, it follows that

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

This establishes (7.14) and, hence (7.13) completing the proof of Lemma 7.6. □

7.3. 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 $\mathcal{A}_{i,H}$ defined as in Definition 7.4, consider the targets

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

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 (7.6)

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

Therefore Corollary 7.2 follows from Theorem 2.10. □

7.4. Notes. The logarithm law for the highest excursion was proven in [135]. The extensions for infinite volume hyperbolic manifolds is studied in [134]. Corollary 7.3 for surfaces is obtained in [86] where the authors also consider infinite volume surfaces. Papers [12, 52] 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. [48, 50]. 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 convergent of the continued fraction expansion of the geodesic endpoint [70], 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, [102] studies partially hyperbolic flows on homogenous spaces and presents applications to metric number theory, cf. Section 9 of the present paper. Logarithm Law for unipotent flows is considered in [9, 10, 65, 88]. In the next section 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 $S\mathcal{Q}$ for the unitary tangent bundle over $\mathcal{Q}$, $\pi : S\mathcal{Q} \to \mathcal{Q}$ the canonical projection, ϕ 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 $t \to d(a, \pi(\phi^t(q,v)))$ has a local minima such that $d_j := d(a, \pi(\phi^{t_j}(q, v)) \leqslant \bar{\rho}$. Let $d_n^{(r)}(a, (q, v))$ be the r-th minima among the numbers $\{d_j\}_{t_j \leqslant n}$.

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

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

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

Lemma 8.2. The following limits exist and does not dependent on $a \in \mathcal{Q}$:

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

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 $r \geqslant 1$, we have

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

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

8.2. **MultiLog Law. Proof of Theorem 8.1.** We fix $r \in \mathbb{N}$ and consider the system $(f, S\mathcal{Q}, \mu, \text{Lip})$, where $f = \phi^\varepsilon$ for a small $\varepsilon > 0$. We note that it follows from [112][Theorem 2.4], [44][Theorem 2], Remark A.1 and Theorem A.2 that $(f, S\mathcal{Q}, \mu, \text{Lip})$ is r-fold exponentially mixing for every $r \geqslant 2$ as in Definition 3.1.

For $k \neq 0$, we keep the notations $\Omega_{a,\rho}^k$ for the event $1_{\Omega_{a,\rho}} \circ f^k$, and $\bar{\Omega}_\rho^k$ 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 $s \geqslant 0$, we let $\rho_n = n^{-1/d} \ln^{-s} n$, and recall that $N_{\rho_n}^n$ denotes the number of times $k \leqslant n$ such that Ω_{a,ρ_n}^k (or $\bar\Omega_{\rho_n}^k$) occurs.

The statement of Theorem 4.4 becomes equivalent to the following :

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

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

With the notation $\mathbf{S}_r = \sum_{j=1}^\infty (2^j \mathbf{v}_j)^r$ 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 $s \leqslant \frac{1}{rd}$.

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

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

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 exists $\eta > 0$ [10] and $t_0 > 0$ such that for any $a \in Q$ and ρ sufficiently small,*

$$(8.2) \qquad \mu(\Omega_{a,\rho} \cap \phi^{-t}\Omega_{a,\rho}) \leqslant \mu(\Omega_{a,\rho})^{1+\eta},$$

for all $t > t_0$.

Recall that $S_q\mathcal{Q}$ is the unit tangent bundle at the point q. Denote $A_\varepsilon(q) = \bigcup_{s \in [0,\varepsilon]} \phi^s S_q\mathcal{Q}$, 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 $a \in \mathcal{Q}$*

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

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

Introduce $\Sigma(t, q, \varepsilon) := \phi^t A_\varepsilon(q)$. 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.*

---

[10]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 $\mu(\Omega_{a,\rho} \cap \phi^{-t}\Omega_{a,\rho}) \leqslant C\mu(\Omega_{a,\rho})^2$

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

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

By Lemma 8.7, $\|d\pi^{-1}\|$ is uniformly bounded. Note that $\pi\Sigma(t, a, \varepsilon)$ 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 piece $\Sigma_j(t)$ 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

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

Summing over j in (8.4) we obtain (8.3) which finishes the proof of Lemma 8.6. $\square$

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 4.10 that treats the case of the composite targets of Section 4.2. It is left to verify $\overline{(\text{Mov})}$. Take $x_i \in \mathcal{Q}$, $B_i = B(x_i, \rho)$, $1 \leqslant i \leqslant k$ such that $\mathcal{Q} = \bigcup_{i=1}^{k} B_i$ and $k = O(\rho^{-d})$. By (8.2), for $t > t_0$

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

This completes the proof of Proposition 8.4 and finishes the proof of Theorem 8.1. $\square$
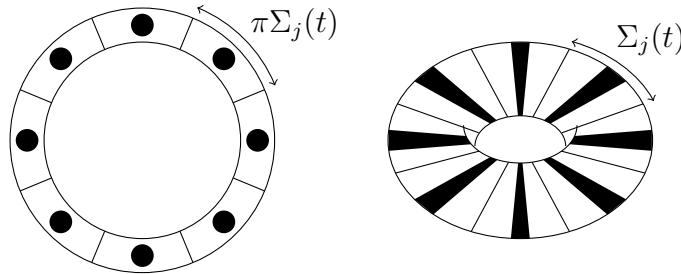

FIGURE 1. Proof of Lemma 8.6

8.3. **Poisson regime. Proof of Corollary 8.3.** Part (a) follows from Theorem 2.10, 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. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

8.4. **Notes.** In [117], Maucourant proved that for all $a \in \mathcal{Q}$ and almost every $(q, v) \in S\mathcal{Q}$

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

[106] generalized Maucourant's result to study a shrinking target problem for time h map. The shrinking target problems for sets with complicated geometry is discussed in [62, 63, 65, 88, 89, 90].

Concerning Poisson Limits we note that visits to sets with complicated geometry naturally appears in Extreme Value Theory, see Section 10 for details. [141] provided a 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.**

Homogenous 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* $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$, $s \geqslant 0$, $c > 0$, *let* $D_N(\alpha, s, c)$ *be the set of* $k = (k_1, \ldots, k_d) \in \mathbb{Z}^d$ *such that*

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

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

Theorem 9.2. *If* $s \leqslant 1/r$ *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 couples$)$. *Given* $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$ *and* $z \in \mathbb{R}$, $s \geqslant 0$ *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*

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

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

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

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

Definition 9.6 ((r,s)-simultaneously approximable vectors). *Given $\alpha = (\alpha_1, \ldots \alpha_d) \in \mathbb{R}^d$, $s \geqslant 0$, $c > 0$, let $D_N(\alpha, s, c)$ be the set of $k \in \mathbb{Z}^*$ such that*

$$k \leqslant N \ \text{ and } \ \exists m \in \mathbb{Z}^d : gcd(k, m_1, \ldots, m_d) = 1$$

$$\text{and for all } i = 1, \ldots, d, \quad k^{\frac{1}{d}} |k\alpha_i + m_i| \leqslant \frac{c}{(\ln N)^{\frac{1}{d}} (\ln \ln N)^{\frac{s}{d}}}.$$

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

Theorem 9.7. *If $s \leqslant 1/r$ 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 9.7 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 $\mathcal{M}$ be the space of $d+1$ dimensional unimodular lattices. We identify $\mathcal{M}$ with $SL_{d+1}(\mathbb{R})/SL_{d+1}(\mathbb{Z})$. Denote Haar measure on $\mathcal{M}$ by μ. Define

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

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

$$(9.1) \qquad\qquad g_t = \begin{pmatrix} 2^{-t} & & & \\ & \ddots & & \\ & & 2^{-t} & \\ & & & 2^{dt} \end{pmatrix}$$

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

Given a function f on $\mathbb{R}^{d+1}$ we consider its Siegel transform $\mathcal{S}(f) : \mathcal{M} \to \mathbb{R}$ defined by

$$(9.2) \qquad\qquad \mathcal{S}(f)(\mathcal{L}) = \sum_{e \in \mathcal{L}, \ e \text{ prime}} f(e).$$

For $a > 0$, let ϕ_a be the indicator of the set[11]

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

---

[11]We added $x_1 > 0$ in the definition of E_a since we will restrict to vectors $k \in \mathbb{Z}^d$ with $k_1 \geqslant 0$.

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

$$(9.3) \qquad \nu := \frac{c}{M(\ln M)^s}, \quad \Phi_\nu := \mathcal{S}(\phi_\nu).$$

For $t \geqslant 0$, we then define

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

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

$$(9.4) \qquad \exists m, \quad gcd(k_1, \ldots k_d, m) = 1, \quad |k|^d |\langle k, \alpha \rangle + m| \leqslant \frac{c}{M(\ln M)^s}.$$

If α is such that $\Phi_\nu(g_t \Lambda_\alpha) \leqslant 1$ for every $t \in \mathbb{N}$, then we get that α is (r, s)-approximable if and only if there exists infinitely many M for which there exists $0 < t_1 < t_2 < \ldots < t_r \leqslant M$ satisfying $\alpha \in \bigcap_{j=1}^r A_{t_j}(M)$.

But in general, for α and $t \leqslant M$ such that $\alpha \in A_t(M)$, there may be multiple solutions k such that $2^t < |k| \leqslant 2^{t+1}$ for the same t. Since in Theorem 9.2 we are counting all solutions we have to deal with this issue.

The following proposition proven in §9.3 shows that for a.e. α, multiple solutions do not occur.

Proposition 9.8. *For almost every α, we have that for every M sufficiently large, for every $t \in [0, M]$, it holds that $\Phi_\nu(g_t \Lambda_\alpha) \leqslant 1$*

Hence, Theorem 9.2 is equivalent to the following.

Theorem 9.9. *If $rs \leqslant 1$, then for almost every $\alpha \in \mathbb{T}^d$, there exists infinitely many M for which there exists $0 < t_1 < t_2 < \ldots < t_r \leqslant M$ 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 exists $0 < t_1 < t_2 < \ldots < t_r \leqslant M$ 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 $\Lambda_\alpha : \alpha \in \mathbb{T}^d$. But this horocycle is exactly the full strong unstable direction of the rapidly mixing partially hyperbolic action g_t. Due to the equidistribution of the strong unstable horocycles, it is thus possible and much more convenient to work with Haar measure on $\mathcal{M}$ instead of Haar measure on Λ_α for $\alpha \in \mathbb{T}^d$.

Hence, we define

$$B_t(M) := \{\mathcal{L} \in \mathcal{M} : \Phi_\nu(g_t \mathcal{L}) \geqslant 1\},$$

where we recall that $\nu := \dfrac{c}{M(\ln M)^s}$, $\Phi_\nu := \mathcal{S}(\phi_\nu)$, and ϕ_ν is the indicator of the set $E_\nu = \left\{(x, y) \in \mathbb{R}^d \times \mathbb{R} \mid x_1 > 0, |x| \in [1, 2], |x|^d |y| \in [0, \nu]\right\}$.

Our goal becomes to prove the following.

Proposition 9.10. *For μ-almost every $\mathcal{L} \in \mathcal{M}$, we have that for every M sufficiently large, for every $t \in [0, M]$, it holds that $\Phi_\nu(g_t \mathcal{L}) \leqslant 1$.*

Theorem 9.11. *If $rs \leqslant 1$, then for μ-almost every $\mathcal{L} \in \mathcal{M}$, there exists infinitely many M for which there exists $0 < t_1 < t_2 < \ldots < t_r \leqslant M$ satisfying*

$$\mathcal{L} \in \bigcap_{j=1}^r B_{t_j}(M).$$

If $rs > 1$, then for μ-almost every $\mathcal{L} \in \mathcal{M}$, there exists at most finitely many M for which there exists $0 < t_1 < t_2 < \ldots < t_r \leqslant M$ satisfying

$$\mathcal{L} \in \bigcap_{j=1}^r B_{t_j}(M).$$

Proof that Proposition 9.10 and Theorem 9.11 imply Proposition 9.8 and Theorem 9.9.

Recall that for $M \in \mathbb{N}$ we defined $\nu = \frac{c}{M(\ln M)^s}$. Fix $\eta > 0$ and define $\Phi_\nu^\pm$ as in (9.7) but with $(1 + \eta)c$ and $(1 - \eta)c$ instead of c. Next, define for $\beta \in \mathbb{R}^d$

$$\Lambda_\beta^- = \begin{pmatrix} \mathrm{Id}_d & \beta \\ 0 & 1 \end{pmatrix},$$

and for $B \in SL_d(\mathbb{R})$ we define

$$D_B = \begin{pmatrix} B & 0 \\ 0 & 1 \end{pmatrix},$$

and finally

$$\tilde{\Lambda}_{\alpha,\beta,B} = D_B \Lambda_\beta^- \Lambda_\alpha.$$

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 $\mathcal{M}$ 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

$$\Phi_\nu^-(g_t \tilde{\Lambda}_{\alpha,\beta,B}) \geqslant 1 \implies \Phi_\nu(g_t \Lambda_\alpha) \geqslant 1 \implies \Phi_\nu^+(g_t \tilde{\Lambda}_{\alpha,\beta,B}) \geqslant 1.$$

This shows that Proposition 9.8 and Theorem 9.9 follow from Proposition 9.10 and Theorem 9.11 respectively. $\qquad \square$

9.4. **Rogers identities.** The following identities (see [114, 139]) 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 $\mathbb{R}^{d+1}$.
Let

$$F(\mathcal{L}) = \sum_{e \in \mathcal{L}, \text{ prime}} f(e), \quad \bar{F}(\mathcal{L}) = \sum_{e_1 \neq \pm e_2 \in \mathcal{L}, \text{ prime}} f_1(e_1) f_2(e_2).$$

F is the *Siegel transform of* f that we denoted $\mathcal{S}(f)$.

Lemma 9.12. *We have*

$$(a) \quad \int_{\mathcal{M}} F(\mathcal{L}) d\mu(\mathcal{L}) = \mathbf{c}_1 \int_{\mathbb{R}^{d+1}} f(x) dx,$$

$$(b) \quad \int_{\mathcal{M}} \bar{F}(\mathcal{L}) d\mu(\mathcal{L}) = \mathbf{c}_2 \int_{\mathbb{R}^{d+1}} f_1(x) dx \int_{\mathbb{R}^{d+1}} f_2(x) dx.$$

9.5. **Multiple solutions on the same scale. Proof of Proposition 9.10.** Recall that $\nu = \frac{c}{M \ln M^s}$

Lemma 9.13. *There exists a constant $C > 0$, such that for every M, for every $t \in \mathbb{R}$, it holds that*

$$\mu \left(\Phi_\nu(g_t \mathcal{L}) > 1 \right) \leqslant C c^2 M^{-2} (\ln M)^{-2s}.$$

For $K \geqslant 0$, apply the lemma for $M = 2^K$ and sum over all $t \in [0, M]$, then

$$\mu \left(\exists t \leqslant 2^K, \Phi_{4\nu}(g_t \mathcal{L}) > 1 \right) \leqslant 16 C c^2 2^{-K} K^{-2s}.$$

The straightforward side of Borel Cantelli lemma gives that for almost every $\mathcal{L}$, for K sufficiently large, for any $t \leqslant 2^K, \Phi_{4\nu}(g_t \mathcal{L}) \leqslant 1$. For the same $\mathcal{L}$, it then holds that for M sufficiently large, for any $t \leqslant M, \Phi_\nu(g_t \mathcal{L}) \leqslant 1$.

To finish the proof of Proposition 9.10 we give

Proof of Lemma 9.13. Since g_t preserves Haar measure on $\mathcal{M}$ it suffices to prove the lemma for $t = 0$. But the condition $k_1 \geqslant 0$ implies that

$$\Phi_\nu^2(\mathcal{L}) - \Phi_\nu(\mathcal{L}) = \sum_{e_1 \neq e_2 \in L \text{ prime}} \phi_\nu(e_1) \phi_\nu(e_2) = \sum_{e_1 \neq \pm e_2 \in L \text{ prime}} \phi_\nu(e_1) \phi_\nu(e_2).$$

It then follows from Rogers identity (b) of Lemma 9.12 that

$$\mu \left(\Phi_\nu(\mathcal{L}) > 1 \right) \leqslant \mathbb{E} \left(\Phi_\nu^2(\mathcal{L}) - \Phi_\nu(\mathcal{L}) \right) \leqslant \mathbf{c}_2 \left(\int_{\mathbb{R}^{d+1}} \phi_\nu(u) du \right)^2 \leqslant C c^2 M^{-2} (\ln M)^{-2s}. \quad \square$$

9.6. **Proof of Theorem 9.11.**

We want to apply Corollary 3.8. For the system (f, X, μ) we take $(g_1, \mathcal{M}, \mu)$, where μ is the Haar measure on $\mathcal{M}$. For the targets, we take $\Omega_\rho = \{\mathcal{L} : \Phi_\rho(\mathcal{L}) \geqslant 1\}$ 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 9.11 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 $t \leqslant n$ such that Ω_ρ^t occurs.

Indeed, if we recall 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 $rs \leqslant 1$.

This being said, to be able to apply Corollary 3.8 and finish, we still need to check the conditions of Definition 3.1 and Definition 3.2 for the system $(g_1, \mathcal{M}, \mu, \mathbb{B})$ and for the family of targets given by Ω_ρ and the sequence ρ_M. $(EM)_r$ follows from Theorem 1.1 of [19], Remark A.1 and Theorem A.2. And the approximation condition (Appr) can be checked as follows. Indeed we have:

CLAIM. There exists $\sigma > 0$ such that, for every $\rho > 0$ sufficiently small, there exists $A_\rho^-, A_\rho^+ \in \text{Lip}(\mathcal{M})$ such that

   (i) $\|A_\rho^\pm\|_\infty \leqslant 2$ and $\|A_\rho^\pm\|_{\text{Lip}} \leqslant \rho^{-\sigma}$;
   (ii) $A_\rho^- \leqslant 1_{\Omega_\rho} \leqslant A_\rho^+$;
   (iii) $\mu(A_\rho^+) - \mu(A_\rho^-) \leqslant \rho^2$

Clearly the claim implies (Appr) since $\mu(\Omega_\rho) = \mathcal{O}(\rho)$.

Proof of the claim. Recall that $\Phi_\rho = \mathcal{S}(\phi_\rho)$, 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 (iii) with $\mu(A_\rho^-)$ replaced by $\mu(1_{\Omega_\rho})$. The construction of A_ρ^- is similar.

Pick $f^+ \in \text{Lip}(\mathbb{R}^{d+1}, [0, 2])$ such that for some $\sigma > 0$

   • $\|f^+\|_{\text{Lip}} \leqslant \rho^{-\sigma}$
   • For $z \in E_\rho$, $f^+(z) = 1$
   • For $z \notin E_{\rho+\rho^{10}}$, $f^+(z) = 0$.

As the consequence $\mathcal{S}(f^+) \in \text{Lip}(\mathcal{M})$ and $\Phi_\rho \leqslant \mathcal{S}(f^+)$, and using Rogers identity of Lemma 9.12(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 $\mathcal{E}_\rho \subset \mathcal{M}$ such that $\mu(\mathcal{E}_\rho) \leqslant \rho^3$

   (P1) For $\mathcal{L} \notin \mathcal{E}_\rho$, if $\mathcal{S}(f^+) > 0$, then $\Phi_\rho \geqslant 1$.
   (P2) If $\mathcal{M}_\rho := \{\mathcal{L} : \mathcal{S}(f^+) < 2\}$, then $\|\mathcal{S}(f^+)\|_{\text{Lip}(\mathcal{M}_\rho)} \leqslant \rho^{-\sigma-1}$.

Let now $u : \mathbb{R} \to [0, 1]$ be some increasing C^∞ function such that $u(x) = 0$ for $x \leqslant 0$ and $u(x) = 1$ for $x \geqslant 1$.

Finally, introduce $A_\rho^+ : \mathcal{M} \to \mathbb{R}$ such that for $\mathcal{L} \in \mathcal{M}$

$$A_\rho^+(\mathcal{L}) = u\left(\mathcal{S}(f^+)(\mathcal{L})\right)$$

We now check that A_ρ^+ satisfies the requirements of the claim.

Since $u \in C^\infty(\mathcal{M}, [0,1])$ we get that $A_\rho^+ \in \mathrm{Lip}(\mathcal{M})$ and $\|A_\rho^+\|_\infty \leqslant 2$. To prove the Lipschitz bound, observe that for $\mathcal{L} \notin \mathcal{M}_\rho$ we have that $A_\rho^+(\mathcal{L}) = 1$, while for $\mathcal{L} \in M_\rho$ we have $(P2)$. Hence $\|A_\rho^+\|_{\mathrm{Lip}} \leqslant \rho^{-2\sigma}$. This proves (i) of the claim. To see (ii), just observe that

$$\Phi_\rho(\mathcal{L}) \geqslant 1 \implies \mathcal{S}(f^+)(\mathcal{L}) \geqslant 1 \implies A_\rho^+(\mathcal{L}) = 1.$$

We turn to (iii). If $\mathcal{L} \notin \mathcal{E}_\rho$, then by $(P1)$

$$A_\rho^+(\mathcal{L}) > 0 \implies \mathcal{S}(f^+)(\mathcal{L}) > 0 \implies \Phi_\rho(\mathcal{L}) \geqslant 1 \implies A_\rho^+(\mathcal{L}) = 1.$$

Since $\mu(\mathcal{E}_\rho) \leqslant \rho^3$ and $\|A_\rho^+\|_\infty \leqslant 2$, we get that $\mu(A_\rho^+) - \mu(1_{\Omega_\rho}) \leqslant \rho^2$ and (iii) is proved. $\qquad\square$

Next we show now how Rogers identity of Lemma 9.12(b) implies (Mov). Define

$$E_\nu^\tau = \left\{ (x,y) \in \mathbb{R}^d \times \mathbb{R} \mid x_1 > 0, 2^{-\tau}|x| \in [1,2], |x|^d|y| \in [0,\nu] \right\}$$

and let ϕ_ν^τ be the indicator function of E_ν^τ. Then

$$\mu(\Omega_\rho \cap g_{-\tau}\Omega_\rho) \leqslant \mathbb{E}(\Phi_\rho \Phi_\rho \circ g_\tau) = \int_{\mathcal{M}} \sum_{e_2 \neq \pm e_1 \in \mathcal{L} \text{ prime}} \phi_\rho(e_1)\phi_\rho^\tau(e_2) d\mu(\mathcal{L}),$$

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 9.12 (b) we get that

$$\mu(\Omega_\rho \cap g_{-\tau}\Omega_\rho) \leqslant C\mu(\Omega_\rho)^2$$

which is stronger than the required (Mov).

Finally, the condition (Poly) clearly holds for the sequence $\rho_M = \frac{c}{M(\ln M)^s}$ due to Lemma 9.12 (a). $\qquad\square$

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 $\widetilde{\mathcal{M}}$ be the space of $d+1$ dimensional unimodular affine lattices. We identify $\widetilde{\mathcal{M}}$ with $SL_{d+1}(\mathbb{R}) \ltimes \mathbb{R}^{d+1}/SL_{d+1}(\mathbb{Z}) \ltimes \mathbb{Z}^{d+1}$, where the multiplication rule 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 $\widetilde{\mathcal{M}}$.

For $\alpha \in \mathbb{R}^d$ and $z \in \mathbb{R}$, we define

$$(9.5) \qquad \Lambda_{\alpha,z} = (\Lambda_\alpha, (0, \dots, 0, z))$$

Given a function f on $\mathbb{R}^{d+1}$ we consider its Siegel transform $\mathcal{S}(f) : \widetilde{\mathcal{M}} \to \mathbb{R}$ defined by

$$(9.6) \qquad \tilde{\mathcal{S}}(f)(\tilde{\mathcal{L}}) = \sum_{e \in \tilde{\mathcal{L}}} f(e).$$

Note that, unlike our definition of the Siegel transform in the case of regular lattices, we do not require in this affine setting that the vectors e in the summation be prime. This is because in this affine 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.

For $a > 0$, let $\tilde{\phi}_a$ be the indicator of the set[12]

$$\tilde{E}_a := \left\{ (x, y) \in \mathbb{R}^d \times \mathbb{R} \mid |x| \in [1, 2], |x|^d |y| \in [0, a] \right\}$$

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

(9.7) $$\nu := \frac{c}{M(\ln M)^s}, \quad \tilde{\Phi}_\nu := \tilde{\mathcal{S}}(\phi_\nu).$$

For $t \geqslant 0$, we then define

$$\tilde{A}_t(M) := \{ (\alpha, z) \in \mathbb{R}^d \times \mathbb{R} : \tilde{\Phi}_\nu(g_t \Lambda_{\alpha, z}) \geqslant 1 \}$$

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 $2^t < |k| \leqslant 2^{t+1}$ and that

$$\exists m, \quad |k|^d |z + \langle k, \alpha \rangle + m| \leqslant \frac{c}{M(\ln M)^s}.$$

If α is such that $\tilde{\Phi}_\nu(g_t \Lambda_{\alpha, z}) \leqslant 1$ for every $t \in \mathbb{N}$, then we get that (α, z) is (r, s)-approximable if and only if there exists infinitely many M for which there exists $0 < t_1 < t_2 < \ldots < t_r \leqslant M$ satisfying $(\alpha, z) \in \bigcap_{j=1}^{r} \tilde{A}_{t_j}(M)$.

But in general, for α and $t \leqslant M$ such that $(\alpha, z) \in \tilde{A}_t(M)$, there may be multiple solutions k such that $2^t < |k| \leqslant 2^{t+1}$ 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.

Proposition 9.14. *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 $\tilde{\Phi}_\nu(g_t \Lambda_{\alpha, z}) \leqslant 1$*

Hence, Theorem 9.5 is equivalent to the following.

Theorem 9.15. *If $rs \leqslant 1$, then for almost every $(\alpha, z) \in \mathbb{T}^d \times \mathbb{T}$, there exists infinitely many M for which there exists $0 < t_1 < t_2 < \ldots < t_r \leqslant M$ 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 exists $0 < t_1 < t_2 < \ldots < t_r \leqslant M$ satisfying

$$\alpha \in \bigcap_{j=1}^{r} \tilde{A}_{t_j}(M).$$

---

[12]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 homogenous case of Theorem 9.2 do not appear in the inhomogeneous Diophantine approximation problem of Theorem 9.5.

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 9.14 and Theorem 9.15 into a problem of multiple recurrence of the diagonal action g_t when applied to a random lattice in $\widetilde{\mathcal{M}}$.

We define

$$\tilde{B}_t(M) := \{\tilde{\mathcal{L}} \in \widetilde{\mathcal{M}} : \tilde{\Phi}_\nu(g_t\tilde{\mathcal{L}}) \geqslant 1\}$$

Our goal becomes to prove the following.

Proposition 9.16. *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 $\tilde{\Phi}_\nu(g_t\tilde{\mathcal{L}}) \leqslant 1$*

Theorem 9.17. *If $rs \leqslant 1$, then for $\tilde{\mu}$-almost every $\tilde{\mathcal{L}} \in \widetilde{\mathcal{M}}$, there exists infinitely many M for which there exists $0 < t_1 < t_2 < \ldots < t_r \leqslant M$ 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 exists $0 < t_1 < t_2 < \ldots < t_r \leqslant M$ satisfying

$$\tilde{\mathcal{L}} \in \bigcap_{j=1}^{r} \tilde{B}_{t_j}(M).$$

9.9. **Proofs of Proposition 9.16 and Theorem 9.17.** Again, the proofs of Proposition 9.16 and Theorem 9.17 are very similar to the proofs of their counterpart in the homogeneous case, Proposition 9.10 and Theorem 9.11.

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 $\Omega_\rho = \{\tilde{\mathcal{L}} : \tilde{\Phi}_\rho(\tilde{\mathcal{L}}) \geqslant 1\}$. 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 9.16 and Theorem 9.17 compared to that of Proposition 9.10 and Theorem 9.11, is in the application of Rogers identities to prove Proposition 9.16 as well as in the proof of (Mov) that is part of the proof of Theorem 9.17.

We explain this difference now.

In fact, Rogers identities are slightly simpler in the affine case, where there is no need to pay a special attention to the multiples of a vector in the affine lattice. Recall (9.6) Rogers identities for affine lattices (read [114])

$$\mathbb{E}(\tilde{\mathcal{S}}(f)) = \int_{\mathbb{R}^{d+1}} f(u)du$$

$$\mathbb{E}(\tilde{\mathcal{S}}(f)^2) = \left(\int_{\mathbb{R}^{d+1}} f(u)du\right)^2 + \int_{\mathbb{R}^{d+1}} f^2(u)du.$$

(The idea behind the proof for the second moment identity is that the linear functionals on the space of continuous functions on $\mathbb{R}^{d+1} \times \mathbb{R}^{d+1}$ that are $SL_{d+1}(\mathbb{R}) \ltimes \mathbb{R}^{d+1}$ invariant

can be identified to invariant measures on $\mathbb{R}^{d+1} \times \mathbb{R}^{d+1}$ by the action of $SL_{d+1}(\mathbb{R}) \ltimes \mathbb{R}^{d+1}$. But the orbits of the latter action decompose into pairs of independent vectors and pairs of equal vectors.)

Now for the proof Proposition 9.16, we have that

$$\tilde{\mu}\left(\tilde{\Phi}_\nu(\tilde{\mathcal{L}}) > 1\right) \leqslant \mathbb{E}\left(\tilde{\Phi}_\nu^2(\tilde{\mathcal{L}}) - \tilde{\Phi}_\nu(\tilde{\mathcal{L}})\right) \leqslant \left(\int_{\mathbb{R}^{d+1}} \tilde{\phi}_\nu(u) du\right)^2 \leqslant Cc^2 M^{-2}(\ln M)^{-2s}$$

and Proposition 9.16 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 $\tau \geqslant 1$

$$\tilde{E}_\nu^\tau = \left\{(x, y) \in \mathbb{R}^d \times \mathbb{R} \mid 2^{-\tau}|x| \in [1, 2], |x|^d |y| \in [0, \nu]\right\}$$

and for $\tilde{\phi}_\nu^\tau$ the indicator function of $\tilde{E}_\nu^\tau$, observe that

$$\tilde{\mu}(\Omega_\rho \cap g_{-\tau}\Omega_\rho) \leqslant \mathbb{E}\left(\tilde{\Phi}_\rho\left(\tilde{\Phi}_\rho \circ g_\tau\right)\right) = \int_{\mathcal{M}} \sum_{e_2, e_1 \in \tilde{\mathcal{L}}} \tilde{\phi}_\rho(e_1)\tilde{\phi}_\rho^\tau(e_2) d\tilde{\mu}(\tilde{\mathcal{L}})$$

$$= \int_{\tilde{\mathcal{M}}} \sum_{e_2 \neq e_1 \in \tilde{\mathcal{L}}} \tilde{\phi}_\rho(e_1)\tilde{\phi}_\rho^\tau(e_2) d\tilde{\mu}(\tilde{\mathcal{L}}) = \left(\int_{\mathbb{R}^{d+1}} \tilde{\phi}_\rho(u) du\right)^2 \leqslant C\tilde{\mu}(\Omega_\rho)^2$$

which is stronger than the required (Mov). $\qquad\qquad\qquad\qquad\qquad\square$

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 $\hat{d}_n^{(r)}(z, \alpha)$ 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

$$(9.8) \qquad\qquad \limsup_{n \to \infty} \frac{|\ln \hat{d}_n^{(1)}(z, \alpha)| - \frac{1}{d}\ln n}{\ln \ln n} = \frac{1}{d},$$

$$(9.9) \qquad\qquad \limsup_{n \to \infty} \frac{|\ln \hat{d}_n^{(r)}(z, \alpha)| - \frac{1}{d}\ln n}{\ln \ln n} = \frac{1}{2d}, \text{ for } r \geqslant 2.$$

Next we claim that it suffices to prove (9.9) only for $r = 2$. Indeed, since $\hat{d}_n^{(r)}$ is non decreasing in r, (9.9) with $r = 2$ implies that for $r > 2$,

$$\limsup_{n \to \infty} \frac{|\ln \hat{d}_n^{(r)}(z, \alpha)| - \frac{1}{d}\ln n}{\ln \ln n} \leqslant \frac{1}{2d}.$$

To get the upper bound, suppose that $\hat{d}_n^{(2)}(z, \alpha) \leqslant \varepsilon$. Then there are $0 \leqslant k_1 < k_2 < n$ such that $k_j\alpha \in B(z, \varepsilon)$. Let $k = k_2 - k_1$. Then

$$k_2 + s\alpha \in B(z, (1 + 2s)\varepsilon)$$

for $s = 1, \cdots, r - 2$. Thus $\hat{d}^{(r)}_{(r-1)n}(z, \alpha) \leqslant (2r - 1)\hat{d}^{(2)}_n(z, \alpha)$. Taking limit superior, we obtain that if (9.9) holds for $r = 2$ then it holds for arbitrary r. In summary, we only need to show (9.8) and

$$(9.10) \qquad \limsup_{n \to \infty} \frac{|\ln \hat{d}^{(2)}_n(z, \alpha)| - \frac{1}{d} \ln n}{\ln \ln n} = \frac{1}{2d}.$$

The proofs of (9.8) and (9.10) are similar to but easier than the proof of Theorem 9.5 so we only explain the changes. First, it is suffices to take limit superior, for n of the form 2^M since for $2^{M-1} \leqslant n \leqslant 2^M$ we have

$$\hat{d}^{(r)}_{2^M}(z, \alpha) \leqslant \hat{d}^{(r)}_n(z, \alpha) \leqslant \hat{d}^{(r)}_{2^{M-1}}(z, \alpha).$$

Let $\nu_M = M^{-s}$ for a suitable s and

$$(9.11) \qquad \hat{E}_\nu = \{e = (e', e'') \in \mathbb{R}^d \times \mathbb{R} : ||e'|| \leqslant \nu, e'' \in (0, 1]\}.$$

Then a direct inspection shows that

$$\hat{d}^{(r)}_{2^M}(z, \alpha) \leqslant \nu_M \Leftrightarrow \tilde{S}(\mathbf{1}_{\hat{E}_{\nu_M}})(\hat{g}_M \hat{\Lambda}_{\alpha, z}) \geqslant r,$$

where $\tilde{S}$ is defined by (9.6), $\hat{g}_M = g_{-M/d}$ for g given by (9.1), 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 $\widetilde{\mathcal{M}}$ denoted by the space of $d + 1$ dimensional unimodular affine lattices and $\tilde{\mu}$ the Haar measure on $\widetilde{\mathcal{M}}$. As in the proof of Theorem 9.5 one can show that $\tilde{S}(\mathbf{1}_{\hat{E}_{\nu_M}})(\hat{g}_M \hat{\Lambda}_{\alpha, z}) \geqslant r$ infinitely often for almost every (z, α) if and only if $\tilde{S}(\mathbf{1}_{\hat{E}_{\nu_M}})(\hat{g}_M \tilde{\mathcal{L}}) \geqslant r$ infinitely often for almost every $\tilde{\mathcal{L}} \in \widetilde{\mathcal{M}}$. Thus we need to show that for almost every $\tilde{\mathcal{L}} \in \widetilde{\mathcal{M}}$

$$(9.12) \qquad \tilde{S}(\mathbf{1}_{\hat{E}_{\nu_M}})(\hat{g}_M \tilde{\mathcal{L}}) \geqslant 1 \text{ infinitely often if } s < \frac{1}{d},$$

$$(9.13) \qquad \tilde{S}(\mathbf{1}_{\hat{E}_{\nu_M}})(\hat{g}_M \tilde{\mathcal{L}}) \geqslant 1 \text{ finitely often if } s > \frac{1}{d},$$

$$(9.14) \qquad \tilde{S}(\mathbf{1}_{\hat{E}_{\nu_M}})(\hat{g}_M \tilde{\mathcal{L}}) \geqslant 2 \text{ infinitely often if } s < \frac{1}{2d},$$

$$(9.15) \qquad \tilde{S}(\mathbf{1}_{\hat{E}_{\nu_M}})(\hat{g}_M \tilde{\mathcal{L}}) \geqslant 2 \text{ finitely often if } s > \frac{1}{2d}.$$

To prove (9.12)–(9.15), we need the following fact.

Lemma 9.18. *(a)* $\tilde{\mu}\left(\tilde{S}(1_{\hat{E}_\nu}) = 1\right) = c_d \nu^d (1 + \mathcal{O}(\nu^{2d}))$,

  (b) $c' \nu^{2d} \leqslant \tilde{\mu}\left(\tilde{S}(1_{\hat{E}_\nu}) \geqslant 2\right) \leqslant c'' \nu^{2d}$.

Before we give the proof of the lemma, we see how it allows to obtain (9.12)–(9.15) and finish the proof of part (a) of Theorem 4.7.

Indeed, Lemma 9.18 shows that

$$\sum_M \tilde{\mu}\left(\tilde{S}(1_{\hat{E}_{\nu_M}}) = 1\right) = \infty \iff s \leqslant \frac{1}{d}, \quad \sum_M \tilde{\mu}\left(\tilde{S}(1_{\hat{E}_{\nu_M}}) \geqslant 2\right) = \infty \iff s \leqslant \frac{1}{2d}.$$

From there, (9.12)–(9.15) follow from the the classical Borel Cantelli Lemma, that is, from the case $r = 1$ in our Theorem 2.4.[13] For this, denote $\hat{\Phi}_\nu = \tilde{S}(1_{\hat{E}_\nu})$, and observe that the verification of the conditions of Definitions 3.1, and Definition 3.2 for the targets $\Omega_\rho = \{\tilde{\mathcal{L}} : \hat{\Phi}_\rho(\tilde{\mathcal{L}}) \geqslant 1\}$ is very similar to the proof of Theorem 9.5 so we omit it.

Proof of Lemma 9.18. we get by Rogers

$$\mathbb{E}(\hat{\Phi}_\nu) = c_d \nu^d, \quad \mathbb{E}(\hat{\Phi}_\nu^2 - \hat{\Phi}_\nu) = \left(c_d \nu^d\right)^2.$$

It follows that

$$\tilde{\mu}(\hat{\Phi}_\nu \geqslant 2) \leqslant \mathbb{E}(\hat{\Phi}_\nu^2 - \hat{\Phi}_\nu)/2 \leqslant C\nu^{2d}$$

proving the upper bound of part (b).

In addition

$$\mathbb{E}\left(\hat{\Phi}_\nu \mathbf{1}_{\hat{\Phi}_\nu \geqslant 2}\right) \leqslant \left(c_d \nu^d\right)^2$$

so that

(9.16) $$\tilde{\mu}(\hat{\Phi}_\nu = 1) = \mathbb{E}(\hat{\Phi}_\nu) - \mathbb{E}(\hat{\Phi}_\nu \mathbf{1}_{\hat{\Phi}_\nu \geqslant 2}) = c_d \nu^d + \mathcal{O}\left(\nu^{2d}\right).$$

This proves part (a).

To prove the lower bound in part (b) we need the following estimate. Denote $\mathcal{L}_{prime}$ the set of prime vectors in $\mathcal{L}$ for $\mathcal{L} \in \mathcal{M} = SL_{d+1}(\mathbb{R})/SL_{d+1}(\mathbb{Z})$. Let

$$\bar{E}_1 = \left\{(e', e'') \in \mathbb{R}^d \times \mathbb{R} : |e'| \in \left[\frac{\nu}{10}, \frac{\nu}{5}\right], |e''| \leqslant \frac{1}{10}\right\},$$

$$\bar{E}_2 = \left\{(e', e'') \in \mathbb{R}^d \times \mathbb{R} : |e'| \leqslant \frac{\nu}{5}, |e''| \leqslant \frac{1}{10}\right\},$$

$$\mathcal{A} = \left\{\mathcal{L} \in \mathcal{M} : \text{Card}\left(\mathcal{L}_{prime} \cap \bar{E}_1\right) = \text{Card}\left(\mathcal{L}_{prime} \cap \bar{E}_2\right) = 1\right\}.$$

CLAIM. *We have*

(9.17) $$\mu(\mathcal{A}) = c\nu^d.$$

Assume the claim holds. Denote $\tilde{z} = (z, 0)$. For $\mathcal{L} \in \mathcal{A}$, the fundamental domain of $\mathbb{R}^{d+1}/\mathcal{L}$ can be chosen to contain

$$\bar{E}_3 = \left\{(e', e'') \in \mathbb{R}^d \times \mathbb{R} : |e'| \leqslant \frac{\nu}{100}, |e''| \leqslant \frac{1}{100}\right\}.$$

---

[13]We note that in case $r = 1$ Theorem 2.4 is a minor variation of standard dynamical Borel Cantelli Lemmas such as e.g., the Borel Cantelli Lemma of [102].

We thus have

$$\mu\left((\mathcal{L}+\tilde{z}): \text{Card}\left((\mathcal{L}+\tilde{z})\cap\hat{E}_\nu\right)\geqslant 2\right)\geqslant\mu\left(\mathcal{A}\right)\mu\left(\text{Card}\left((\mathcal{L}+\tilde{z})\cap\hat{E}_\nu\right)\geqslant 2|\mathcal{A}\right)$$

$$\geqslant\mu(\mathcal{A})\mu(z\in\bar{E}_3)\geqslant c'\nu^{2d}.$$

This gives the lower bound in part (b) of Lemma 9.18. To complete the proof, we now give the

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,

$$\mathbb{E}\left(\Psi_1\right)=\frac{1}{10}c_d\nu^d,\quad\mathbb{E}\left(\Psi_1^2-\Psi_1\right)=\left(\frac{1}{10}c_d\nu^d\right)^2.$$

Thus arguing as in the proof of (9.16) we conclude that

(9.18) $$\mu(\Psi_1=1)=\frac{1}{10}c_d\nu^d+\mathcal{O}\left(\nu^{2d}\right).$$

Rogers identities also give

$$\mathbb{E}(\Psi_1(\Psi_2-\Psi_1))=\mathcal{O}\left(\nu^{2d}\right).$$

Hence

(9.19) $$\mu\left(\text{Card}(\mathcal{L}_{prime}\cap\bar{E}_1)\geqslant 1\text{ and Card}\left(\mathcal{L}_{prime}\cap\left(\bar{E}_2\backslash\bar{E}_1\right)\right)\geqslant 1\right)=\mathcal{O}\left(\nu^{2d}\right).$$

Combining (9.18) and (9.19) we obtain (9.17) for $d>1$.

In case $d=1$ we still have $\mathbb{E}(\Psi_1)=c\nu+\mathcal{O}(\nu^2)$. On the other hand, for $d=1$ we have $\text{Card}\left(\mathcal{L}_{prime}\cap\bar{E}_2\right)\leqslant 1$ since $\mathcal{L}$ is unimodular. Thus

$$\mathbb{E}(\Psi_1)=\mu(\Psi_1=1)=\mu(\Psi_1=1\text{ and }\Psi_2-\Psi_1=0)=c\nu.\qquad\square$$

This completes the proof of Lemma 9.18 and thus of part (a) of Theorem 4.7. $\qquad\square$

PROOF OF PART (b). It is clear that for any r, if $\bar{\mathcal{E}}_r$ is not empty then it is equal to M. The fact that $\bar{\mathcal{E}}_1=M$ implies that $\bar{\mathcal{E}}_r=M$ for all r is exactly similar to the implication of (9.9) from (9.10), so we just focus on showing that $\bar{\mathcal{E}}_1=M$. Adapting the beginning of the proof of part (a) to the current homogeneous setting, we see that what we want to prove boils down to showing that for almost every $\mathcal{L}\in\mathcal{M}$

(9.20) $$\mathcal{S}(\mathbf{1}_{E_{\nu_n}})(\hat{g}_n\mathcal{L})\geqslant 1\text{ infinitely often if }s<\frac{1}{d},$$

(9.21) $$\mathcal{S}(\mathbf{1}_{E_{\nu_n}})(\hat{g}_n\mathcal{L})\geqslant 1\text{ finitely often if }s>\frac{1}{d},$$

where E_{ν_n} is as in (9.11), and $\mathcal{S}$ designates the Siegel transform as in (9.2). By Rogers identity, Lemma 9.12(a), we have that $\mathbb{E}\left(\mathcal{S}(\mathbf{1}_{E_{\nu_n}})\right)=cn^{-sd}$, hence (9.20) and (9.21) follow by classical Borel Cantelli Lemma (see for example the Borel Cantelli Lemma of [102]) or by the case $r=1$ of our Theorem 2.4 .

This completes the proof of Theorem 4.7. $\qquad\square$

9.11. **Notes.** A classical Khintchine–Groshev Theorem is given by (1.2)–(1.3). A lot of interest is devoted to extending this result to α lying in a submanifold of $\mathbb{R}^d$ (see e.g. [13, 18]). The applications of dynamics to Diophantine approximation are based on Dani correspondence [39]. In particular, [101] discusses Khintchine–Groshev type results on manifolds using dynamical tools. The use of Siegel transform as a convenient analytic tool for applying Dani correspondence can be found in [114]. Surveys on applications of dynamics to metric Diophantine approximations include [16, 20, 47, 53, 54, 66, 98, 104, 115]. Limit Theorems for Siegel transforms are discussed in [8, 11, 22, 48, 49].

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 Defenitions 4.3 and 4.5 respectively. Recall also that under the conditions of Theorems 4.4 and 4.6 $\mu(\mathcal{G}_r) = 1$ and $\mathcal{H}$ contains a residual set.

Given a function ϕ and a point $y \in M$, let $\phi_n^{(r)}(y)$ be the r-th minimum among the values $\{\phi(f^j y)\}_{j=1}^n$.

Theorem 10.1. *(a) Suppose f is $(2r+1)$-fold exponentially mixing preserving a smooth measure μ. Then*

(i) There is a set $\mathcal{G}$ of full measure in M such that if ϕ is a function with a unique non degenerate minimum at $x \in \mathcal{G}$, then for almost every $y \in M$,

$$\limsup_{n \to \infty} \frac{\left| \ln \left(\phi_n^{(r)}(y) - \phi(x) \right) \right| - \frac{2}{d} \ln n}{\ln \ln n} = \frac{2}{rd}.$$

(ii) If $\mathcal{G}_1 = M$ and the periodic orbits of f are dense, then there is a dense G_δ set $\mathcal{H} \subset M$, such that if ϕ is a function with a unique non degenerate minimum at $x \in \mathcal{H}$, then for almost every $y \in M$,

$$\limsup_{n \to \infty} \frac{\left| \ln \left(\phi_n^{(r)}(y) - \phi(x) \right) \right| - \frac{2}{d} \ln n}{\ln \ln n} = \frac{2}{d}.$$

(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 $\mathcal{G}_\mu$ with $\mu(\mathcal{G}_\mu) = 1$, such that if ϕ is a function with a unique non degenerate minimum at $x \in \mathcal{G}_\mu$, then for μ–almost every $y \in M$,

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

where σ given by (6.4).

(c) Part (a) remains valid for the geodesics flow on a compact $(d+1)-$dimensional manifold $\mathcal{Q}$ and functions $\phi : \mathcal{Q} \to \mathbb{R}$ which have unique non-degenerate minimum at some point on $\mathcal{Q}$. (In this case $\phi_r(y)$ is the r-th local minimum of the map $t \mapsto \phi(q(t))$ 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

$$\limsup_{n\to\infty} \frac{\left|\ln\left(\phi_n^{(r)}(y) - \phi(x)\right)\right| - \frac{2}{d}\ln n}{\ln\ln n} = \begin{cases} \frac{2}{d} & \text{if } r = 1, \\ \frac{1}{d} & \text{if } r \geqslant 2. \end{cases}$$

Proof. At a non-degenerate minimum x we have that for y close to x

$$(10.1) \qquad K^{-1}d^2(x,y) \leqslant \phi(y) - \phi(x) \leqslant Kd^2(x,y)$$

so part (i) of (a) holds for $x \in \mathcal{G}_r$ and part (ii) of (a) holds for $x \in \mathcal{H}$ 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. $\qquad\square$

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*

$$\frac{\phi_n^{(1)}(y) - \phi(x)}{\rho^2}, \frac{\phi_n^{(2)}(y) - \phi(x)}{\rho^2}, \ldots, \frac{\phi_n^{(r)}(y) - \phi(x)}{\rho^2}, \ldots$$

with $n = [\tau\rho^{-d}]$ converges as $\rho \to 0$ to the Poisson process on $\mathbb{R}^+$ with measure $\gamma(\phi)\tau\frac{d}{2}t^{\frac{d}{2}-1}dt$, where $\gamma(\phi) > 0$ 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

$$(10.2) \qquad \Omega^{n,j} = \left\{ y : \phi(y) - \phi(x) \in \left[r_j^-\rho^2, r_j^+\rho^2\right] \right\},$$

that satisfy

$$\lim_{\rho\to 0} \tau\rho^{-d}\mu(\Omega^{n,j}) = \tau\gamma(\phi)((r_j^+)^{\frac{d}{2}} - (r_j^-)^{\frac{d}{2}}) = \tau\gamma(\phi)\int_{r_j^-}^{r_j^+} \frac{d}{2}t^{\frac{d}{2}-1}dt.$$

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 2.11 and conclude the Poisson limit. $\qquad\square$

Next, we consider functions of the form

$$(10.3) \qquad \psi(y) = \frac{c}{d^s(x,y)} + \widetilde{\psi}(y), \quad \text{where} \quad c < 0 \quad \text{and} \quad \widetilde{\psi} \in Lip(M).$$

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

$$\frac{\rho^s \psi_n^{(1)}(y)}{c}, \frac{\rho^s \psi_n^{(2)}(y)}{c}, \dots, \frac{\rho^s \psi_n^{(r)}(y)}{c}, \dots \quad where \quad n = \tau\rho^{-d}$$

converges as $\rho \to 0$ to the Poisson process on $\mathbb{R}^+$ with measure $\frac{d\tau\gamma(x)}{s} t^{-(d/s)-1} dt$.

The proofs of the above results is similar to the proofs of Theorem 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 $t > 0$

$$\lim_{n\to\infty} \mu(y : \phi_n^{(1)}(y) > n^{-2/d}t) = e^{-\sigma t^{d/2}}.$$

(a) (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 $t > 0$

$$\lim_{n\to\infty} \mu(y : \phi_n^{(1)}(y) > -n^{-s/d}t) = e^{-\sigma t^{-d/s}}.$$

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 [109] 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) and (M2). More discussions about relations of extreme value theory to Poisson limit theorems in the context of dynamical systems can be found in [58]. The book [113] 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, 26, 41, 91, 92, 118] and references wherein.

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 $r \geqslant 1$, f is called $(r + 1)$-fold exponentially mixing if there are constant $s, \bar{C}$ and $\bar{\theta} < 1$ such that for any C^s functions $A_0, A_1, \dots, A_r$ for any r tuple $k_1 < k_2 < \cdots < k_r$

$$(A.1) \qquad \left| \int \prod_{j=0}^r (A_j \circ f^{k_j}) \, d\mu - \prod_{j=0}^r \int A_j d\mu \right| \leqslant \bar{C}\bar{\theta}^m \prod_{j=0}^r \|A_j\|_{C^s},$$

where $m = \min\limits_{j}(k_j - k_{j-1})$ 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

(A.2)
$$\left| \int B(x_0, f^{k_1}x_0, \cdots, f^{k_r}x_0)d\mu(x_0) - \int B(x_0, \cdots, x_r)d\mu(x_0)\cdots d\mu(x_r) \right| \leqslant C_s\bar{\theta}^m \|B\|_{C^s}.$$

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) 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)

$$\|B - \bar{B}\|_{C^0} \leqslant e^{-\varepsilon\alpha m}, \quad \|\bar{B}\|_{C^s} \leqslant e^{\varepsilon s m}.$$

Then

$$\int B(x_0, f^{k_1}x_0, \cdots, f^{k_r}x_0)d\mu(x_0) = \int \bar{B}(x_0, f^{k_1}x_0, \cdots, f^{k_r}x_0)d\mu(x_0) + O\left(e^{-\varepsilon\alpha m}\right)$$

$$= \int \bar{B}(x_0, x_1, \ldots x_r)d\mu(x_0)d\mu(x_1)\ldots d\mu(x_r) + O\left(e^{-\varepsilon\alpha m}\right) + O\left(\theta^m e^{\varepsilon s m}\right)$$

$$= \int B(x_0, x_1, \ldots x_r)d\mu(x_0)d\mu(x_1)\ldots d\mu(x_r) + O\left(e^{-\varepsilon\alpha m}\right) + O\left(\theta^m e^{\varepsilon s m}\right).$$

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

$$\|\psi_j\|_{C^1} \leqslant C_u \zeta_j^u \leqslant C_u \lambda^u \quad \text{if} \quad u > 1 + \frac{d}{2}.$$

It follows from (A.1) that if $\phi \not\equiv 1$ then

$$\left| \int \phi_\lambda(x, f^{k_1}x, \ldots, f^{k_r}x)d\mu(x) - \prod_{j=0}^{r} \int \psi_j d\mu \right| \leqslant C\lambda^{u(r+1)}\theta^m.$$

Therefore

$$\left| \int B(x, f^{k_1}x, \ldots, f^{k_r}x)d\mu(x) - \int B(x_0, \cdots, x_r)d\mu(x_0)\cdots d\mu(x_r) \right|$$

$$\leqslant C\theta^m \sum_\lambda b_\lambda \lambda^{u(r+1)} \leqslant C\theta^m ||B||_{H^{u(r+1)}(M^{r+1})}.$$

This proves the result if $s > \left(1 + \frac{d}{2}\right)(r+1)$. $\qquad\square$

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 [138, Proposition 3.8], we have that for $\hat\psi_1 \in \text{Lip}(\mathbb{T})$, $\hat\psi_2 \in L^1(\mu)$,

(A.3) $$\left| \int \hat\psi_1(\hat\psi_2 \circ f^n)d\mu - \int \hat\psi_1 d\mu \int \hat\psi_2 d\mu \right| \leqslant C\|\hat\psi_1\|_{Lip}\|\hat\psi_2\|_{L^1}\bar\theta^n, \ n \geqslant 0.$$

Step 2. We proceed to show inductively that for each $r > 0$ and $\psi_i \in \text{Lip}(\mathbb{T})$ for $i = 1, \ldots, r$,

(A.4) $$\left| \int \left(\prod_{i=1}^{r} \psi_i \circ f^{k_i} \right) d\mu - \prod_{i=1}^{r} \int \psi_i d\mu \right| \leqslant C\bar\theta^m \prod_{i=1}^{r} ||\psi_i||_{Lip},$$

where $m = \min_{1\leqslant i\leqslant r-1}(k_{i+1} - k_i)$, $k_0 = 0$.

By invariance of μ we may assume that $k_1 = 0$. Applying (A.3) with $\hat\psi_1 = \psi_1$, $\hat\psi_2 = \prod_{j=2}^{r} \psi_j \circ f^{k_j - k_2}$ we get

$$\left| \int \left(\prod_{i=1}^{r} \psi_i \circ f^{k_i} \right) d\mu - \left(\int \psi_1 d\mu \right) \left[\int \left(\prod_{i=2}^{r} \psi_i \circ f^{k_i} \right) d\mu \right] \right|$$

$$\leqslant C\bar\theta^m \|\psi_1\|_{Lip} \left\| \left(\prod_{i=2}^{r} \psi_i \circ f^{k_i} \right) \right\|_{L_1} \leqslant C\bar\theta^m \prod_{i=1}^{r} ||\psi_i||_{Lip}.$$

Applying inductive estimate to

$$\int \left(\prod_{i=2}^{r} \psi_i \circ f^{k_i} \right) d\mu$$

we obtain (A.4).

Step 3. Applying the same argument as in proof of Theorem A.2 we get $(EM)_r$. $\qquad\square$

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. 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. ([44, Theorem 2]) Suppose that for each subset D in a single unstable leave of bounded geometry[14] and any Hölder probability density ρ on D we have

$$\left| \int_D A(f^n x)\rho(x)dx - \int A d\mu \right| \leqslant C\theta^n \|A\|_{C^s} \|\rho\|_{C^\alpha}$$

for $A \in C^s$. Then f is r-fold exponentially mixing for all $r \geqslant 2$.

Examples of maps satisfying the conditions of Theorem A.3 include expanding maps, volume preserving Anosov diffeomorphsims [23, 121], time one maps of contact Anosov flows [112], mostly contracting systems [27, 43], partially hyperbolic translations on homogeneous spaces [100], and partially hyperbolic automorphisms of nilmanifolds [67].

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 (so called generalized T, T^{-1} transformations) provided that the skewing function has positive drift. We refer the reader to [45] for more details.

Another source of exponential mixing is spectral gap for transfer operators (cf. §A.2 as well as [121, 138]). This allows to handle non-uniformly hyperbolic systems admitting Young tower with exponential tails [142] as well as piecewise expanding maps [138].

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

(B.1) $$\ln \mu \left(B_n(x, \varepsilon) \right) = \sum_{j=0}^{n-1} g(f^j x) + O(1).$$

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

$$(C_0 \exp \varepsilon^\alpha)^{-1} \leqslant \frac{|Df^n(y)|}{|Df^n(x)|} \leqslant C_0 \exp \varepsilon^\alpha.$$

---

[14]We refer the reader to [44] for precise requirements on D since those requirements are not essential for the present discussion.

Recalling (6.2)

$$\exp\left[\left(\sum_{j=0}^{n-1} f_u(f^j x)\right) - \varepsilon^\alpha\right]\frac{d(x,y)}{C_0} \leqslant d(f^n x, f^n y) \leqslant C_0 \exp\left[\left(\sum_{j=0}^{n-1} f_u(f^j x)\right) + \varepsilon^\alpha\right] d(x,y).$$

Hence

$$\varepsilon C_0^{-1}\exp\left[\left(-\sum_{j=0}^{n-1} f_u(f^j x)\right) - \varepsilon^\alpha\right] \leqslant r_n \leqslant \bar{r}_n \leqslant \varepsilon C_0 \exp\left[\left(-\sum_{j=0}^{n-1} f_u(f^j x)\right) + \varepsilon^\alpha\right].$$

It follows that

$$(B.2) \qquad \ln r_n = \sum_{j=0}^{n-1} -f_u(f^j x) + O(1), \quad \ln \bar{r}_n = \sum_{j=0}^{n-1} -f_u(f^j x) + O(1).$$

Next define

$$N(r) = \max\left(n : B(x,r) \subset B_n(x,\varepsilon)\right), \quad \bar{N}(r) = \min\left(n : B(x,r) \supset B_n(x,\varepsilon)\right).$$

Then, similarly to (B.2) we obtain

$$(B.3) \qquad \ln r = \sum_{j=0}^{N(r)-1} -f_u(f^j x) + O(1) = \sum_{j=0}^{\bar{N}(r)-1} -f_u(f^j x) + O(1).$$

B.2. **Proof of** (6.7) **and** (6.8). Note that

$$(B.4) \qquad\qquad \mu(B_{\bar{N}(r)}(x,\varepsilon)) \leqslant \mu(B(x,r)) \leqslant \mu(B_{N(r)}(x,\varepsilon)).$$

Since f is uniformly expansing there is a positive constant C such that for each x $1/C \leqslant f_u(x) \leqslant C$. Accordingly

$$(B.5) \qquad\qquad \frac{N(r)}{C} \leqslant |\ln r| \leqslant CN(r), \quad \frac{\bar{N}(r)}{C} \leqslant |\ln r| \leqslant C\bar{N}(r).$$

On the other hand, since $P(g) = 0$, [121, 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 $x \in \mathbb{T}$,

$$(B.6) \qquad\qquad -\hat{C}_1 N - \hat{C}_2 \leqslant \sum_{n=0}^{N-1} g(f^n x) \leqslant -\hat{\varepsilon}N + \hat{C}_2$$

Combining (B.1), (B.4), (B.5) and (B.6) we obtain (6.7).

Next (B.3) shows that $N(4r) - \bar{N}(r) = O(1)$. Now (6.8) follows from (B.1) and (B.4).

B.3. **Proof of Lemma 6.3(b).** Observe that (B.1) (B.2) give

$$\ln \mu\left(B_n(x,\varepsilon)\right) - \mathbf{d} \ln r_n = \sum_{j=0}^{n-1} \psi(f^j x) + O(1), \quad \ln \mu\left(B_n(x,\varepsilon)\right) - \mathbf{d} \ln \bar{r}_n = \sum_{j=0}^{n-1} \psi(f^j x) + O(1)$$

where ψ is defined by (6.3).

By Law of Iterated Logarithm [84],

$$\limsup_{n\to\infty} \frac{\sum_{j=0}^{n-1} \psi(f^j x)}{\sqrt{2n \ln \ln n}} = \sigma, \quad \liminf_{n\to\infty} \frac{\sum_{j=0}^{n-1} \psi(f^j x)}{\sqrt{2n \ln \ln n}} = -\sigma.$$

Since $B(x, r_n) \subset B_n(x,\varepsilon) \subset B(x, \bar{r}_n)$

$$\limsup_{n\to\infty} \frac{|\ln \mu\left(B(x,\bar{r}_n)\right)| - \mathbf{d}|\ln \bar{r}_n|}{\sqrt{2n \ln \ln n}} \leqslant \sigma \leqslant \limsup_{n\to\infty} \frac{|\ln \mu\left(B(x,r_n)\right)| - \mathbf{d}|\ln r_n|}{\sqrt{2n \ln \ln n}}.$$

Using (B.2) again, we conclude that for every sufficiently small δ, there exists $n(\delta)$ and k independent of δ and $n(\delta)$ such that $\bar{r}_{n+k} \leqslant \delta \leqslant r_n$. Then

$$\sigma \leqslant \limsup_{\delta\to 0} \frac{|\ln \mu(B(x,r_{n(\delta)}))| - \mathbf{d}|\ln r_{n(\delta)}|}{\sqrt{2n(\delta) \ln \ln n(\delta)}} \leqslant \limsup_{\delta\to 0} \frac{|\ln \mu\left(B(x,\delta)\right)| - \mathbf{d}|\ln \delta|}{\sqrt{2n(\delta) \ln \ln n(\delta)}}$$

$$\leqslant \limsup_{\delta\to 0} \frac{|\ln \mu(B(x,\bar{r}_{n(\delta)}))| - \mathbf{d}|\ln \bar{r}_{n(\delta)}|}{\sqrt{2n(\delta) \ln \ln n(\delta)}} \leqslant \sigma.$$

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 $\lim_{n\to\infty} \dfrac{|\ln r_n|}{n} = \lambda$. For such x we have $\lim_{n\to\infty} \dfrac{|\ln r_n|(\ln \ln |\ln r_n|)}{n \ln \ln n} = \lambda$. Since $r_n/C \leqslant \delta \leqslant r_n$ 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. $\qquad\square$

B.4. **Proof of Lemma 6.3(a).** Suppose that $\sigma^2 = 0$. Since we also have that $\int \psi d\mu = 0$ [121, 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. $\qquad\qquad\square$

APPENDIX C. GEODESIC FLOWS: GEOMETRY OF TARGETS IN THE CONFIGURATION SPACE. PROOF OF LEMMA 8.2 AND LEMMA 8.7.

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

$$(J^i)''(t) + \sum_{j=1}^{n} A_j^i(t) J^j(t) = 1, i = 1, \dots, n,$$

where the matrix $A(t) = (A_j^i(t))_{i,j=1,\dots,n}$ 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 [Lemma 1.1][107]).

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

(C.1) $$J(0) = 0, \quad \|J'(0)\| = 1.$$

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*

(C.2) $$\|J'(t)\| \leqslant C\|J(t)\| \quad \text{for } t > t_0.$$

Proof. Denote $S(t) = \langle J(t), J'(t) \rangle$, $N(t) = \|J'(t)\|^2$ and $|||J|||^2 = \|J\|^2 + \|J'\|^2$. Then

(C.3) $$\frac{d}{dt}S(t) = \|J'(t)\|^2 + \langle J(t), J''(t) \rangle = \|J'(t)\|^2 + \langle J(t), -K(t)J(t) \rangle \geqslant C_1 |||J|||^2$$

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 $\frac{d}{dt}S(t) > \frac{C_1 S(t)}{2}$, whence

(C.4) $$S(t) > S(u)e^{C_1(t-u)/2} \quad \text{for } t > u.$$

Next $N(t) \geqslant N(0)e^{-K_2^2 t} = e^{-K_2^2 t}$ which together with (C.3) gives

(C.5)
$$S(t) \geqslant e^{-K_2^2 t} t \quad \text{for } t \in [0, 1].$$

Combining this with (C.4) we get

(C.6)
$$S(t) > e^{-K_2^2} e^{C_1(t-1)/2} \quad \text{for } t > 1.$$

Combining (C.5) and (C.6) with a trivial bound

(C.7)
$$N(t) \leqslant |||J(t)||| \leqslant N(0)e^{K_2^2 t} = e^{K_2^2 t}$$

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

$$\max(|c_+|, |c_-|) \leqslant C_3, \quad |||J_-||| \leqslant C_4 e^{-K_1 t}$$

and

(C.8)
$$J_+ = R(t)J_+'(t)$$

where R is a symmetric matrix with spectrum between K_1 and K_2. It follows that

(C.9)
$$|||J(t)||| \leqslant c_+ |||J_+(t)||| + C_3 C_4 e^{-K_1 t} \leqslant \sqrt{1 + K_2^2} \, \|c_+ J_+(t)\| + C_3 C_4 e^{-K_1 t}$$

On the other hand (C.6) gives a uniform lower bound

(C.10)
$$|||J||| \geqslant 2 e^{-K_2^2/2} e^{C_1(t-1)/4}$$

Combining (C.9) and (C.10) we obtain

$$\|J(t)\| \geqslant \|c_+ J^+(t)\| - c_- \|J^-(t)\| \geqslant \frac{2}{1 + K_2^2} e^{-K_2^2/2} e^{C_1(t-1)/4} - 2C_3 C_4 e^{-K_1 t}$$

which proves (C.2) for large t. $\qquad \square$

C.2. Volume of the targets in the configuration space.

Proof of Lemma 8.2. If $(q, v) \in \hat{B}_\rho(a)$, denote

$$L(q, v) = L^+(q, v) + L^-(q, v) \text{ where } L^\pm(q, v) = \sup\{t : \phi^{\pm s}(q, v) \in \hat{B}_\rho(a) \text{ for } 0 \leqslant s \leqslant t\}.$$

Then we have the following estimate

$$\mu(\Omega_{a,\rho}) = \varepsilon \left(\int_{\hat{B}_\rho(a)} \frac{1}{L(q, v)} d\mu \right) (1 + O(\rho))$$

(see e.g. [34]). Note that μ is of the form $d\mu(q, v) = \frac{d\lambda(q) d\sigma(v)}{\lambda(Q)}$ where λ is the Riemann volume on 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

(C.11)
$$\gamma = \frac{1}{\lambda(Q)} \int_{\mathcal{B} \times \mathbb{S}^d} \frac{1}{\mathcal{L}(x, v)} dx d\sigma(v)$$

where $\mathcal{B}$ is the unit ball in $\mathbb{R}^{d+1}$ and $\mathcal{L}(\cdot)$ is defined similarly $L(\cdot)$ with geodesics in Q replaced by geodesics in $\mathbb{R}^{d+1}$. Specifically, an elementary plane geometry gives $\mathcal{L}(x, v) = \sqrt{1 - r_{min}^2}$ 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] Aaronson, J., Nakada, H.: Trimmed sums for non-negative, mixing stationary processes. *Stochastic Process. Appl.* **104** (2003), no. 2, 173–192.

[2] Abadi, M.: Poisson approximations via Chen-Stein for non-Markov processes. *Progr. Probab.* **60** (2008), 1–19.

[3] Abadi, M., Vergne N.: Sharp errors for point-wise Poisson approximations in mixing processes. *Nonlinearity* **21** (2008), no. 12, 2871–2885.

[4] Aldous, D.: Probability approximations via the Poisson clumping heuristic. *Applied Math. Sci.* **77** Springer, New York, (1989), xvi+269 pp.

[5] Aimino R., Nicol, M., Todd M.: 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), no. 3, 1371–1401.

[6] Anosov D. V.; Sinai, Ya. G. Certain smooth ergodic systems. *Russian Math. Surveys* **22** (1967), no. 5 103–167.

[7] Athreya, J. S.: Logarithm laws and shrinking target properties. *Proc. Indian Acad. Sci. Math. Sci.* **119** (2009), no. 4, 541–557.

[8] Athreya, J. S., Ghosh A., Tseng J.: Spiraling of approximations and spherical averages of Siegel transforms. *J. Lond. Math. Soc.* **91** (2015), no. 2, 383–404.

[9] Athreya, J. S., Margulis, G. A.: Logarithm laws for unipotent flows-I. *J. Mod. Dyn.* **3** (2009) 359–378.

[10] Athreya, J. S., Margulis, G. A.: Logarithm laws for unipotent flows-II. *J. Mod. Dyn.* **11** (2017) 1–16.

[11] Athreya, J. S., Parrish A., Tseng J.: Ergodic theory and Diophantine approximation for translation surfaces and linear forms. *Nonlinearity* **29** (2016), no. 8, 2173–2190.

[12] Babillot, M., Peigne, M.: Asymptotic laws for geodesic homology on hyperbolic manifolds with cusps. *Bull. Soc. Math. France* **134** (2006), no. 1, 119–163.

[13] Badziahin, D., Beresnevich, V., Velani, S.: Inhomogeneous theory of dual Diophantine approximation on manifolds. *Adv. Math.* **232** (2013), 1–35.

[14] Barreira, L., Pesin, Y., Schmeling, J.: Dimension and product structure of hyperbolic measures. *Ann. of Math.* **149** (1999), no. 3, 755–783.

[15] Barreira, L., Saussol, B.: Hausdorff dimension of measures via Poincaré recurrence. *Comm. Math. Phys.* **219** (2001), no. 2, 443–463.

[16] Bekka, M. B., Mayer, M.: *Ergodic theory and topological dynamics of group actions on homogeneous spaces.* London Math. Soc. Lecture Note Ser. **269** (2000), x+200 pp.

[17] Benedetti, R., Petronio, C.: *Lectures on hyperbolic geometry.* Springer-Verlag, Berlin (1992), xiv+330 pp.

[18] Bernik, V. I., Dodson, M. M.: *Metric Diophantine approximation on manifolds.* Cambridge Tracts in Mathematics **137** (1999), xii+172 pp.

[19] Björklund, M., Einsiedler M, Gorodnik, A.: Quantitative multiple mixing. ArXiv: 1701.00945v2.

[20] Björklund, M., Gorodnik, A.: Central limit theorems in the geometry of numbers. *Electron. Res. Announc. Math. Sci.* **24** (2017) 110–122.

[21] Björklund, M., Gorodnik, A.: Central Limit Theorems for group actions which are exponentially mixing of all orders, to appear in *Journal d'Analyse Mathematiques.*

[22] Björklund, M., Gorodnik, A.: Central limit theorems for Diophantine approximants. *Math. Ann.* **374** (2019), no. 3-4, 1371–1437.

[23] Bowen, R.: *Equilibrium states and the ergodic theory of Anosov diffeomorphisms. 2d revised edition.* Springer Lecture Notes in Math. **470** (2008), viii+75 pp.

[24] Bruin, H., Todd, M.: Return time statistics of invariant measures for interval maps with positive Lyapunov exponent. *Stoch. Dyn.* **9** (2009), no. 1, 81–100.

[25] Bruin, H., Vaienti, S.: Return time statistics for unimodal maps. *Fund. Math.* **176** (2003), no. 1, 77–94.

[26] Carney, M., Nicol, M.: Dynamical Borel-Cantelli lemmas and rates of growth of Birkhoff sums of non-integrable observables on chaotic dynamical systems. *Nonlinearity* **30** (2017), no. 7, 2854–2870.

[27] Castro, A.: Fast mixing for attractors with a mostly contracting central direction. *Ergodic Theory Dynam. Systems* **24** (2004), no. 1, 17–44.

[28] Carvalho M., Freitas, A. C. M., Freitas J. M., Holland M., Nicol M.: Extremal dichotomy for uniformly hyperbolic systems. *Dyn. Syst.* **30** (2015), no. 4, 383–403.

[29] Chaika, J.: Shrinking targets for IETs: extending a theorem of Kurzweil. *Geom. Funct. Anal.* **21** (2011), no. 5, 1020–1042.

[30] Chaika, J., Constantine, D.: Quantitative shrinking target properties for rotations and interval exchanges. *Israel J. Math.* **230** (2019), no. 1, 275–334.

[31] Chazottes, J.-R., Collet, P.: Poisson approximation for the number of visits to balls in non-uniformly hyperbolic dynamical systems. *Ergodic Theory Dynam. Systems* **33** (2013), no. 1, 49–80.

[32] Chazottes, J.-R., Ugalde, E.: Entropy estimation and fluctuations of hitting and recurrence times for Gibbsian sources. *Discrete Contin. Dyn. Syst. B* **5** (2005), no. 3, 565–586.

[33] Chernov, N. I.: Limit theorems and Markov approximations for chaotic dynamical systems, *Probab. Theory Related Fields* **101** (1995), no. 3, 321–362.

[34] Chernov, N.: Entropy, Lyapunov exponents, and mean free path for billiards. *J. Statist. Phys.* **88** (1997), no. 1-2, 1–29.

[35] Chernov, N., Kleinbock, D.: Dynamical Borel-Cantelli lemmas for Gibbs measures, *Israel J. Math.* **122** (2001), 1–27.

[36] Coelho, Z.: Asymptotic laws for symbolic dynamical processes. *London Math. Soc. Lecture Note Ser. 279* (2000), no. 3, 123–165.

[37] Collet, P.: Statistics of closest return for some non-uniformly hyperbolic systems. *Ergodic Theory Dynam. Systems* **21** (2001), no. 2, 401–420.

[38] Collet, P., Galves, A., Schmitt, B.: Repetition times for Gibbsian sources. *Nonlinearity* **12** (1999), no. 4, 1225–1237.

[39] Dani, S. G.: Divergent trajectories of flows on homogeneous spaces and Diophantine approximation. *J. Reine Angew. Math.* **359** (1985), 55–89.

[40] Denker, M., Kan, N.: Om Sevast'yanov's theorem. *Statist. Probab. Lett.* **77** (2007), no. 3, 272–279.

[41] Diamond, H. G., Vaaler, J. D.: Estimates for partial sums of continued fraction partial quotients. *Pacific J. Math.* **122** (1986), no. 1, 73–82.

[42] Doeblin W.: Remarques sur la théorie métrique des fractions continues. *Compositio Math.* **7** (1940), 353–371.

[43] Dolgopyat, D.: On dynamics of mostly contracting diffeomorphisms. *Comm. Math. Phys.* **213** (2000), no. 1, 181–201.

[44] Dolgopyat, D.: Limit theorems for partially hyperbolic systems. *Trans. Amer. Math. Soc.* **356** (2004), no. 4, 1637–1689.

[45] Dolgopyat, D., Dong, C., Kanigowski, A., Nandori, P.: Mixing properties of generalized T, T^{-1} transformations, arXiv:2004.07298.

[46] Dolgopyat, D., Fayad, B.: Deviations of ergodic sums for toral translations I: Convex bodies. *Geom. Funct. Anal.* **24** (2014), no. 4, 85–115.

[47] Dolgopyat, D., Fayad, B.: Limit theorems for toral translations. *Hyperbolic dynamics, fluctuations and large deviations,* Proc. Sympos. Pure Math. **89** (2015), 227–277.

[48] Dolgopyat, D., Fayad, B.: Deviations of Ergodic sums for Toral Translations II: Squares. ArXiv:1211.4323v1.

[49] Dolgopyat, D., Fayad, B., Vinogradov, I.: Central limit theorems for simultaneous Diophantine approximations. *J. Éc. polytech. Math.* **4** (2017), 1–35.

[50] Dolgopyat, D., Sarig, O.: Quenched and annealed temporal limit theorems for circle rotations, to appear in *Asterisque.*

[51] Dubins, L. E., Freedman, D. A.: A sharper form of the Borel-Cantelli lemma and the strong law. *Ann. Math. Statist.* **36** (1965), 800–807.

[52] Enriquez, N., Le Jan, Y.: Statistic of the winding of geodesics on a Riemann surface with finite area and constant negative curvature. *Rev. Mat. Iberoamericana* **13** (1997), no. 2, 377–401.

[53] Einsiedler, M. L., Ellwood, D. A., Eskin, A., Kleinbock, D., Lindenstrauss, E., Margulis, G., Marmi, S., Yoccoz, J.-C. (editors): *Homogeneous flows, moduli spaces and arithmetic.* Clay Mathematics Proceedings, **10** (2010), xii+438 pp.

[54] Eskin, A.: Counting problems and semisimple groups. *Doc. Math.* Extra Vol. **II** (1998), 539–552.

[55] Fayad, B.: Mixing in the absence of the shrinking target property. *Bull. London Math. Soc.* **38** (2006), no. 5, 829–838.

[56] Fernandez, J. L., Melian, M. V., Pestana, D.: Expanding maps, shrinking targets and hitting times. *Nonlinearity* **25** (2012), no. 9, 2443–2471.

[57] Freedman, D.: 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.

[58] Freitas, A., Freitas, J., Magalhães, M.: Convergence of marked point processes of excesses for dynamical systems. *J. Eur. Math. Soc.* **20** (2018), no. 9, 2131–2179.

[59] Freitas, A., Freitas, J., Magalhães, M.: Complete convergence and records for dynamically generated stochastic processes. *Trans. Amer. Math. Soc.* **373** (2020), no. 1, 435–478.

[60] Galatolo, S.: Dimension via Waiting Time and Recurrence. *Math. Res. Lett.* **12** (2005), no. 2-3, 377–386.

[61] Galatolo, S.: Dimension and Hitting in Rapidly Mixing System. *Math. Res. Lett.* **14** (2007), no. 5, 797–805.

[62] Galatolo, S.: Hitting time in regular sets and logarithm law for rapidly mixing dynamical systems. *Proc. Amer. Math. Soc.* **138** (2010), no. 7, 2477–2487.

[63] Galatolo, S., Nisoli, I.: Shrinking targets in fast mixing flows and the geodesic flow on negatively curved manifolds. *Nonlinearity* **24** (2011), no. 11, 3099–3113.

[64] Galatolo, S., Peterlongo, P.: Long hitting time, slow decay of correlations and arithmetical properties. *Discrete Contin. Dyn. Syst.* **27** (2010), no. 1, 185–204.

[65] Ghosh, A., Kelmer, D.: Shrinking targets for semisimple groups. *Bull. Lond. Math. Soc.* **49** (2017), no. 2, 235–245.

[66] Gorodnik, A., Nevo, A.: Quantitative ergodic theorems and their number-theoretic applications, *Bull. Lond. Math. Soc.* **52** (2015), no. 1, 65–113.

[67] Gorodnik, A., Spatzier, R.: Exponential mixing of nilmanifold automorphisms. *J. Anal. Math.* **123** (2014), 355–396.

[68] Gouëzel, S.: A Borel-Cantelli lemma for intermittent interval maps. *Nonlinearity* **20** (2007), no. 6, 1491–1497.

[69] Groshev, A.: A theorem on a system of linear forms. *Dokl. Akad. Nauk SSSR* **19** (1938), 151–152.

[70] Guivarch, Y., Le Jan, Y.: Asymptotic winding of the geodesic flow on modular surfaces and continued fractions. *Ann. Sci. École Norm. Sup.* **26** (1993), no. 1, 23–50; **29** (1996), no. 6, 811–814.

[71] Gupta, C., Nicol, M., Ott, W.: A Borel-Cantelli lemma for nonuniformly expanding dynamical systems, *Nonlinearity* **23** (2010), no. 8, 1991–2008.

[72] Haydn, N.: The central limit theorem for uniformly strong mixing measures. *Stoch. Dyn.* **12** (2012), no. 4, paper 1250006, 31 pp.

[73] Haydn, N., Nicol, M., Persson, T., Vaienti, S.: A note on Borel-Cantelli lemmas for non-uniformly hyperbolic dynamical systems. *Ergodic Theory Dynam. Systems* **33** (2013), no. 2, 475–498.

[74] Haydn, N., Nicol, M., Vaienti, S., Zhang, L.: Central limit theorems for the shrinking target problem. *J. Stat. Phys.* **153** (2013), no. 5, 864–887.

[75] Haydn, N. T. A., Psiloyenis, Y.: Return times distribution for Markov towers with decay of correlations. *Nonlinearity* **27** (2014), no. 6, 1323–1349.

[76] Haydn, N., Vaienti, S.: The limiting distribution and error terms for return times of dynamical systems. *Discrete Contin. Dyn. Syst.* **10** (2004), no. 3, 589–616.

[77] Haydn, N., Vaienti, S.: Fluctuations of the metric entropy for mixing measures. *Stoch. Dyn.* **4** (2004), no. 4, 595–627.

[78] Haydn, N., Vaienti, S.: Limiting entry times distribution for arbitrary null sets, arXiv:1904.08733.

[79] Haydn, N. T. A., Wasilewska, K.: Limiting distribution and error terms for the number of visits to balls in nonuniformly hyperbolic dynamical systems. *Discrete Contin. Dyn. Syst.* **36** (2016), no. 5, 2585–2611.

[80] Hill, R., Velani, S.: The ergodic theory of shrinking targets. *Invent. Math.* **119** (1995), no. 1, 175–198.

[81] Hirata, M.: Poisson law for Axiom A diffeomorphisms. *Ergodic Theory Dynam. Systems* **13** (1993), no. 3, 533–556.

[82] Hirata, M.: Poisson law for the dynamical systems with self-mixing conditions. In *Dynamical systems and chaos, Vol. 1 (Hachioji, 1994)*, 87–96, World Sci. Publ., River Edge, NJ (1995).

[83] Hirata, M., Saussol, B. Vaienti, S.: Statistics of return times: a general framework and new applications. *Comm. Math. Phys.* **206** (1999), no. 1, 33–55.

[84] Hofbauer, F., Keller, G.: Ergodic Properties of Invariant Measures for Piecewise Monotonic Transformations, In: Hunt B.R., Li TY., Kennedy J.A., Nusse H.E. (eds) *The Theory of Chaotic Attractors*. Springer, New York (1982).

[85] Ibragimov, I. A.: Some limit theorems for stationary processes. (Russian) *Teor. Verojatnost. i Primenen.* **7** (1962), 361–392.

[86] Jaerisch, J., Kessebohmer, M., Stratmann, B. O.: A Fréchet law and an Erdös-Philipp law for maximal cuspidal windings, *Ergodic Theory Dynam. Systems* **33** (2013), no. 4, 1008–1028.

[87] Keller, G.: Rare events, exponential hitting times and extremal indices via spectral perturbation. *Dyn. Syst.* **27** (2012), no. 1, 11–27.

[88] Kelmer, D.: Shrinking targets for discrete time flows on hyperbolic manifolds. *Geom. Funct. Anal.* **27** (2017), no. 5, 1257–1287.

[89] Kelmer D., Oh H. Exponential mixing and shrinking targets for geodesic flow on geometrically finite hyperbolic manifolds. arXiv:1812.05251

[90] Kelmer, D., Yu, S.: Shrinking target problems for flows on homogeneous spaces. *Trans. Amer. Math. Soc.* **372** (2019), no. 9, 6283–6314.

[91] Kesseböhmer, M., Schindler, T.: Strong laws of large numbers for intermediately trimmed Birkhoff sums of observables with infinite mean. *Stochastic Process. Appl.* **129** (2019), no. 10, 4163–4207.

[92] Kesten H., Maller R. A.: Ratios of trimmed sums and order statistics. *Ann. Probab.* **20** (1992), no. 4, 1805–1842.

[93] Khintchine, A. Y.: Einige Sätze über Kettenbruche, mit Anwendungen auf die Theorie der Diophantischen Approximationen. (German) *Math. Ann.* **92** (1924), no. 1-2, 115–125.

[94] Kirsebom, M., Kunde, P., Persson, T.: Shrinking targets and eventually always hitting points for interval maps. ArXiv: 1903.06977.

[95] Kim, D. H.: The shrinking target property of irrational rotations. *Nonlinearity* **20** (2007), no. 7, 1637–1643.

[96] Kim, D. H.: The dynamical Borel-Cantelli lemma for interval maps. *Discrete Contin. Dyn. Syst.* **17** (2007), no. 4, 891–900.

[97] Kim, D. H.: Refined shrinking target property of rotations. *Nonlinearity* **27** (2014), no. 9, 1985–1997.

[98] Kleinbock, D.: Some applications of homogeneous dynamics to number theory. *Smooth ergodic theory and its applications (Seattle, WA, 1999),* 639–660, Proc. Sympos. Pure Math., **69** (2001).

[99] Kleinbock, D., Konstantoulas, I., Richter, F. K.: Zero-one laws for eventually always hitting points in mixing systems. ArXiv: 1904.08584.

[100] Kleinbock, D. Y., Margulis, G. A.: Bounded orbits of nonquasiunipotent flows on homogeneous spaces. *Amer. Math. Soc. Transl. Ser. 2,* **171** (1996), 141–172.

[101] Kleinbock, D. Y., Margulis, G. A.: Flows on homogeneous spaces and Diophantine approximation on manifolds. *Ann. of Math. (2)* **148** (1998), no. 1, 339–360.

[102] Kleinbock, D. Y., Margulis, G. A.: Logarithm laws for flows on homogeneous spaces. *Invent. Math.* **138** (1999), no. 3, 451–494.

[103] Kleinbock, D. Y., Margulis, G. A.: On effective equidistribution of expanding translates of certain orbits in the space of lattices. *Number theory, analysis and geometry,* 385–396, Springer, New York, (2012).

[104] Kleinbock, D., Shah, N., Starkov, A.: 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.

[105] Kleinbock, D., Shi, R., Weiss, B.: Pointwise equidistribution with an error rate and with respect to unbounded functions. *Math. Ann.* **36** (2017), no. 1-2, 857–879.

[106] Kleinbock, D. Y. and Zhao, X.: An application of lattice points counting to shrinking target problems. *Discrete Contin. Dyn. Syst.* **38** (2018), no. 1, 155–168.

[107] Knieper, G.: Hyperbolic dynamics and Riemannian geometry. *Handbook of dynamical systems, Vol. 1A,* 453, North-Holland, Amsterdam, (2002).

[108] Kurzweil, J.: On the metric theory of inhomogeneous Diophantine approximations. *Studia Math.* **15** (1955), 84–112.

[109] Leadbetter, M. R., Lindgren, G., Rootzen, H.: *Extremes and related properties of random sequences and processes,* Springer, New York–Berlin (1983) xii+336 pp.

[110] Ledrappier, F., Young, L.-S.: The metric entropy of diffeomorphisms. Part II. Relations between entropy, exponents and dimension. *Ann. of Math. (2)* **122** (1985), no. 3, 540–574.

[111] Leplaideur, R., Saussol, B.: Central limit theorem for dimension of Gibbs measures in hyperbolic dynamics. *Stoch. Dyn.* **12** (2012), no. 2, paper 1150019, 24 pp.

[112] Liverani, C.: On contact Anosov flows. *Ann. of Math. (2)* **159** (2004), no. 3, 1275–1312.

[113] Lucarini V., Faranda D., de Freitas A., de Freitas J., Holland M., Kuna T., Nicol M., Todd M., Vaienti S.: *Extremes and recurrence in dynamical systems.* John Wiley & Sons, Inc., Hoboken, NJ (2016), xi+295 pp.

[114] Marklof, J.: The *n*-point correlations between values of a linear form. *Ergodic Theory Dynam. Systems* **20** (2000), no. 4, 1127–1172.

[115] Marklof, J.: Distribution modulo one and Ratner's theorem. *Equidistribution in number theory, an introduction,* NATO Sci. Ser. Math. Phys. Chem. **237** (2007), 217–244.

[116] Marklof, J.: Entry and return times for semi-flows. *Nonlinearity* **30** (2017), no. 2, 810–824.

[117] Maucourant, F.: Dynamical Borel-Cantelli lemma for hyperbolic spaces. *Israel J. Math.* **152** (2006), 143–155.

[118] Mori, T.: The strong law of large numbers when extreme terms are excluded from sums. *Z. Wahrsch. Verw. Gebiete* **36** (1976), no. 3, 189–194.

[119] Paccaut, F.: Statistics of return times for weighted maps of the interval. *Ann. Inst. H. Poincare Probab. Statist.* **36** (2000), no. 3, 339–366.

[120] Pacifico, M. J., Yang, F.: Hitting times distribution and extreme value laws for semi-flows. *Discrete Contin. Dyn. Syst.* **37** (2017), no. 11, 5861–5881.

[121] Parry, W., Pollicott, M.: Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque* **187-188** (1990), 268 pp.

[122] Paulin, F., Pollicott, M.: Logarithm laws for equilibrium states in negative curvature. *Comm. Math. Phys.* **346** (2016), no. 1, 1–34.

[123] Philipp, W.: Some metric theorems in number theory. *Pacific J. Math* **20** (1967), 109–127.

[124] Pitskel, B.: Poisson limit law for Markov chains. *Ergodic Theory Dynam. Systems* **11** (1991), no. 3, 501–513.

[125] Pene, F., Saussol, B.: Poisson law for some non-uniformly hyperbolic dynamical systems with polynomial rate of mixing. *Ergodic Theory Dynam. Systems* **36** (2016), no. 8, 2602–2626.

[126] Ross, S.: *A first course in probability. 8th edition.* Macmillan Co., New York; Collier Macmillan Ltd., London (2010), xviii+530 pp.

[127] Saussol, B.: On fluctuations and the exponential statistics of return times. *Nonlinearity* **14** (2001), no. 1, 179–191.

[128] Saussol, B.: Recurrence rate in rapidly mixing dynamical systems. *Discrete Contin. Dyn. Syst.* **15** (2006), no. 1, 259–267.

[129] Saussol, B.: An introduction to quantitative Poincaré recurrence in dynamical systems. *Rev. Math. Phys.* **21** (2009), no. 8, 949–979.

[130] Sevast'janov, B. A.: A Possion limit law in a scheme of sums of dependent random variables. (Russian) *Teor. Verojatnost. i Primenen.* **17** (1972), 733–738.

[131] Sinaĭ, Ja. G.: Gibbs measures in ergodic theory. (Russian) *Uspehi Mat. Nauk* **27** (1972), no. 4(166), 21–64.

[132] Sprindžuk, V. G. *Metric theory of Diophantine approximations.* V. H. Winston & Sons, Washington, D.C. (1979).

[133] Stenlund, M.: A strong pair correlation bound implies the CLT for Sinai billiards, *J. Stat. Phys.* **140** (2010) 154–169.

[134] Stratmann, B., Velani, S. L.: The Patterson measure for geometrically finite groups with parabolic elements, new and old. *Proc. London Math. Soc. (3)* **71** (1995), no. 1, 197–220.

[135] Sullivan, D.: Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics. *Acta Math.* **149** (1982), no. 3-4, 215–237.

[136] Tseng, J.: On circle rotations and the shrinking target properties. *Discrete Contin. Dyn. Syst.* **20** (2008), no. 4, 1111–1122.

[137] Varandas, P.: Correlation decay and recurrence asymptotics for some robust nonuniformly hyperbolic maps. *J. Stat. Phys.* **133** (2008), no. 5, 813–839.

[138] Viana, M.: *Stochastic dynamics of deterministic systems.* Rio de Janeiro: IMPA(1997).

[139] Vinogradov, I.: Limiting distribution of visits of several rotations to shrinking intervals, preprint.

[140] Williams, D.: *Probability with martingales.* Cambridge Math.Textbooks. Cambridge University Press, Cambridge, xvi+251 pp (1991).

[141] Yang F.: Rare event process and entry times distribution for arbitrary null sets on compact manifolds. ArXiv: 1905.09956.

[142] Young L.-S.: Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.* **147** (1998) 585–650.

[143] Zhang X.: Note on limit distribution of normalized return times and escape rate. *Stoch. Dyn.* **16** (2016), no. 3, paper 1660014, 15 pp.

(D. Dolgopyat) Department of Mathematics, University of Maryland, 4417 Mathematics Bldg, College Park, MD 20742, USA
E-mail address: dolgop@umd.edu

(B. Fayad) Institut de Mathématiques de Jussieu-Paris Rive Gauche, Paris, France
E-mail address: bassam.fayad@imj-prj.fr

(S. Liu) Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, China
E-mail address: liusixu@mail.tsinghua.edu.cn