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

DMITRY DOLGOPYAT, BASSAM FAYAD AND SIXU LIU

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 whose 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 $\{\Omega_{k,n}\}_{k=1}^n$ of independent events such that $p_n = \mathbb{P}(\Omega_{k,n})$ 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. In this case the classical Borel Cantelli Lemma says that infinitely many A_ns occur iff $\sum_n \mathbb{P}(\Omega_n) = \infty$.

A vast literature is devoted to extending the above classical results to the case independence is replaced by weak dependence. In particular, there are convenient moment conditions which imply similar results in for (weakly) dependent events. One important distinction between the Poisson Law and the other two results is that the Poisson regime additional geometric conditions on close-by events are required to extend the statement to dependent case. Without such conditions one can have clusters of rare events there the number of clusters has Poisson distribution but there may be several events inside each cluster. We refer the reader to [4] for the comprehensive discussion of Poisson clustering.

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: $\Omega_n(\rho_1) \subset \Omega_n(\rho_2)$ for $\rho_1 < \rho_2$ and such that $\sigma(\rho) = \mathbb{P}(\Omega_\rho)$ does not depend on n. Let $N_n(\rho)$ be a number of $\Omega_k(\rho)$, $k \leqslant n$ 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

$$N_n(\rho_n) = r$$

occurs. Even if the events Ω_ρ^n are independent for different ρ, the variables $N_{n_1}(\rho_{n_1})$ and $N_{n_2}(\rho_{n_2})$ 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 $N_{n_2}(\rho_{n_2}) \neq 0$ it is very likely that all the events $\Omega_k(\rho_{n_2})$ occur for $k > n_1$. Using this, one can show under appropriate monotonicity assumptions (see [114]) that $N_n(\rho_n) = r$ infinitely often iff

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

Under the condition $n\rho_n \to 0$ it follows that in the independent case

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

Therefore, under independence, infinitely many $N_n(\rho_n) = r$ occur iff

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

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 the 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 $d(f^k y, x)$ for $0 \leqslant k < n$. 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 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 dynamical Borel Cantelli Lemma which implies in particular that, if μ is smooth then for all x and almost all y we have

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

In Section 4 we extend this result to $r > 1$. For example, if f is an expanding map of the circle, we shall show that for Lebesgue almost all x and y we have

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

Both the smoothness assumption on the invariant measure, Lebesgue typicality assumption on x and hyperbolicity assumption on f are essential. Namely if μ is an invariant Gibbs measure which is **not** conformal, then we show in Section 6 that that there is a constant $c = c(\mu) \neq 0$ such that for μ almost all x and y and for all $r \in \mathbb{N}$

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

We shall also show that there is G_δ–dense set $\mathcal{B}$ such that for all $x \in \mathcal{B}$ and Lebesgue almost all y 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 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 Q of dimension $d+1$ which is not compact but has finite volume. Such manifold admits a thick-thin decomposition. Namely 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

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

be the maximal heights achieved during the excursions which occur before time T placed in the decreasing order. *Sullivan Logarithm Law* is equivalent to saying that for almost every geodesic

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

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

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

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

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

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 ([91, 69, 126]) says that for almost all $\alpha \in \mathbb{R}^d$ there are are infinitely many solutions to

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

iff

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

In particular the inequality

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

has infinitely many solutions for almost every α iff $s \leqslant 1$. We now replace the above inequality by

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

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 $k_1 > 0$) such that[1] $\gcd(k_1, \dots k_d, m) = 1$. (Our interest in smallness of

$$||k||^d|\langle k, \alpha\rangle + m| \tag{1.5}$$

is motivated by [44] where the discrepancy of Kronecker sequences with respect to convex sets is studied. Indeed ks where (1.5) is small are small denominators of the discrepancy and they determine its growth rate.) We show in Section 8 that almost every $\alpha \in \mathbb{R}^d$ is (r, s) approximable iff $s \leqslant \frac{1}{r}$.

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 for in the case the rare events in question involve visits to a sublevel set of a Lipshitz function by an orbit of a smooth exponentially mixing

[1]The last conditions is needed to disregard trivial solutions (lk, lm) where $1 < l$ is not to big while

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

dynamical systems. The results of Section 3 are used to obtain MultiLog Laws in various settings. Namely, Section 4 studies hitting and return times in the phase space, while Section 9 treats similar problems in the configuration space when the system at hand is a geodesic flow on compact negatively curved manifold. Geodesic excursion are discussed in Section 7, and Diophantine approximations are treated in Section 8. 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.

Some auxiliary 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 mixing or independence conditions between the rare events.

Namely, consider a probability space $(\mathbf{\Omega}, \mathbb{P})$ and a family of events Ω^n_ρ such that

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

Let N^n_ρ be the number of times $k \leqslant n$ such that Ω^k_ρ occurs. For $r \in \mathbb{N}$ and $\{\rho_n\}$ a decreasing sequence, we would like to give a criterion that allows to tell when $N^n_{\rho_n} \geqslant r$ will almost surely hold for infinitely many n.

We introduce several conditions quantifying asymptotic independence between the events Ω^n_ρ. The conditions require 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 $\varepsilon n \leqslant \hat{\mathfrak{s}}(n) < n(1-q)/(2(r-1))$ for some $0 < q < 1$, and some $0 < \varepsilon < (1-q)/(2(r-1))$,

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*

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

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

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

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

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

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

then

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

Fix $\bar{r} \in \mathbb{N}$.

Definition 2.1. *Ω_ρ^n are $\bar{r}$–*almost independent at a fixed scale *if $(M1)_r$ and $(M2)_r$ are satisfied for $r \leqslant \bar{r}$. Ω_ρ^n are $\bar{r}$–*almost independent at all scales *if $(M1)_r$, $(M2)_r$ and $(M3)_r$ are satisfied for $r \leqslant \bar{r}$.*

Theorem 2.2. *Fix $r \in \mathbb{N}^*$. Define*

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

(a) If $\mathbf{S}_r < \infty$, and Ω_n^r are $2r$–almost independent at a fixed scale then with probability 1, we have that for large n $N_{\rho_n}^n < r$.

(b) If $\mathbf{S}_r = \infty$, and Ω_n^r are $2r$–almost independent at all scales then with probability 1, there are infinitely many n such that $N_{\rho_n}^n \geqslant r$.

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

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

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

Remark 2.3. *An analogous statement has been obtained in* [1] *under different mixing conditions* [2]

[2]We refer the readers to the notes at the end of each section for a detailed discussion of the related literature.

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

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

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

The goal of this section is to prove the following estimates from which it will be easy to derive Theorem 2.2.

Proposition 2.4. *There exists constants $c, C > 0$ and a sequence $\theta_m \to 0$ such that if*

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

then for $m' > m + 1$ we have[3]

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

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

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

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

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

With these notations

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

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

Lemma 2.5. *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*

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

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

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

Proof. For $m \leqslant r$, denote

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

[3]Note that Proposition 2.4 does not provide a bound for $\mathbb{P}(\mathcal{D}_m \cap \mathcal{D}_{m+1})$.

Note that S_r includes $\frac{((a_2-a_1)n)^r}{r!}(1+\delta'_n)$ terms for some sequence $\delta'_n \to 0$, hence $(M1)_r$ yields

$$\sum_{\substack{a_1 n<k_1<k_2<\cdots<k_r\leqslant a_2 n\\ \mathrm{Sep}_n(k_1,\dots k_r)=r}} \mathbb{P}(A^{k_1,\dots,k_r}_{\rho_n}) = \frac{((a_2-a_1)n\sigma(\rho_n))^r}{r!}(1+\delta_n). \tag{2.10}$$

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

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

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

for some sequence $\eta_n \to 0$. Combining (2.10) with (2.11) we obtain (2.8). The proof of (2.9) is similar to that of (2.10), except that the number of terms is not anymore equivalent to $\frac{1}{r!}((a_2-a_1)n)^r(1+\delta'_n)$ but just larger than $\frac{1}{r!}((a_2-a_1)n-(r-1)\hat{\mathfrak{s}}(n))^r(1+\delta''_n)$ for some sequence $\delta''_n \to 0$, which is larger than $\frac{q^r}{r!}((a_2-a_1)n)^r(1+\delta''_n)$. □

Proof of Proposition 2.4. First, (2.3) follows directly from (2.6) and (2.8). Next, define

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

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

From (2.7) and Bonferroni inequalities we get that

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

Now, (2.9) implies that

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

On the other hand, since

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

we get that

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

and (2.8) then implies that

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

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

Finally, observe that

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

But $(M3)_r$ implies that

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

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

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

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

2.3. **Convergent case. Proof of Theorem 2.2 (a).** Suppose that $\mathbf{S}_r<\infty$. Then $n\sigma(\rho_n)\to 0$. By (2.3) of Proposition 2.4 we have that $\sum\mathbb{P}(\mathcal{A}_m)<\infty$. By Borel-Cantelli Lemma, $\mathcal{A}_m$ happen only finitely many times with probability 1. Observe that $\{N^n_{\rho_n}\geqslant r\}\subset\mathcal{A}_m$ for $2^m<n\leqslant 2^{m+1}$.

Hence with probability one $N^n_{\rho_n}\geqslant r$ happen only finitely many times. □

2.4. **Divergent case. Proof of Theorem 2.2 (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 2.6. *Let $Z_n=\sum_{m=1}^n 1_{\mathcal{D}_m}$. There exists a subsequence $\{Z_{n_k}\}$ such that*

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

Since $\mathbb{E}(Z_n)\to\infty$, the fact that $Z_n\to\infty$ almost surely immediately follows. 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 2.2 (b) in case (2.2) holds.

Proof of Claim 2.6. We first prove that (2.4) and (2.5) imply that

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

or equivalently that

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

Note that

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

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

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

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

$$\delta\sum_{i,j}\mathbb{P}(\mathcal{D}_j)\mathbb{P}(\mathcal{D}_j)\leqslant\delta(\mathbb{E}(Z_n))^2+\mathbb{E}(Z_n)+1$$

(b) The terms where $i\leqslant m$ or $j-i\leqslant m(\delta)$ (including $i=j$) 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

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

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

It remains to consider the case where (2.2) fails. That is, there exists $\delta_0 > 0$ and a sequence j_l such that $\sigma(\rho_{2^{j_l}})2^{j_l} \geqslant \delta_0$. In this case we consider the sets $\tilde{\Omega}^k_{\rho_{2^l}}$ defined as follows. If $\Omega^k_{\rho_{2^l}}$ does not occur then $\tilde{\Omega}^k_{\rho_{2^l}}$ does not occur. Conditionally on $\Omega^k_{\rho_{2^{j_l}}}$ occurring, $\tilde{\Omega}^k_{\rho_{2^l}}$ occurs with probability

$$\gamma_l = \min\left(1, \frac{1}{2^{j_l}\sigma(\rho_{2^{j_l}})\ln l}\right)$$

independently of the other events. In other words, each time $\Omega^k_{\rho_{2^{j_l}}}$ occurs, we keep it with probability γ_l and discard it with probability $1-\gamma_l$. Then the sets $\{\tilde{\Omega}^k_{\rho_{j_l}}\}_{k\leqslant 2^{j_l}, l\in\mathbb{N}}$ satisfy $(M1)_r$, $(M2)_r$, and $(M3)_r$ and

$$\sum_{l=1}^{\infty}\left(\mathbb{P}(\tilde{\Omega}_{\rho_{j_l}})2^{j_l}\right)^r \geqslant \sum_{l=l_0}^{\infty}\left(\frac{\delta_0}{\ln l}\right)^r = \infty$$

where l_0 is the number such that $\gamma_l = \dfrac{1}{2^{j_l}\sigma(\rho_{2^{j_l}})\ln l}$ for $l \geqslant l_0$. Hence more than r events among the events $\{\tilde{\Omega}^k_{\rho_{j_l}}\}_{k\leqslant 2^{j_l}}$ occurs for infinitely many l with probability 1. Since $\tilde{\Omega}^k_{\rho_{j_l}} \subset \Omega^k_{\rho_{j_l}}$ it follows that more than r events among the events $\{\Omega^k_{\rho_{j_l}}\}_{k\leqslant 2^{j_l}}$ occurs for infinitely many l with probability 1. The proof of Theorem 2.2 (b) is thus completed. □

2.5. **Prescribing some details.** In the remaining part of Section 2 we describe some extensions of Theorem 2.2(b).

Namely, we assume that $\Omega^n_\rho = \cup_{i=1}^p \Omega^{n,i}_\rho$ 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 $(k_1, k_2, \ldots, k_r)$ with $\mathrm{Sep}_n(k_1, k_2, \ldots, k_r) = r$ and each $(i_1, i_2, \ldots, i_r) \in \{1, \ldots, p\}^r$

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

and the following extension of $(M3)_r$: for each $i_1, i_2 \dots i_r$, $j_1, j_2 \dots j_r$

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

Theorem 2.7. *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 $i_1, i_2 \dots i_r$ 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 2.7 is similar to the proof of Theorem 2.2(b). Without the loss of generality we may assume that I_j does not contain 0. Then we consider the following modification of $\mathcal{D}_m$

$$\tilde{\mathcal{D}}_m := \{\exists 2^m < k_1 < \dots < k_r \leqslant 2^{m+1} \text{ such that } \frac{k_\alpha}{2^m} \in I_\alpha, \Omega_{\rho_{2^{m+1}}}^{k_\alpha, i_\alpha}$$
$$\text{happens and } k_{\alpha+1} - k_\alpha \geqslant \hat{\mathfrak{s}}(2^{m+1}),\ 0 \leqslant \alpha \leqslant r-1\}.$$

Arguing as in Proposition 2.4 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 2.2(b).

2.6. **Poisson regime.**

Theorem 2.8. *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 $\begin{pmatrix} m \\ r \end{pmatrix}$ denotes the binomial coefficient $\dfrac{m!}{r!(m-r)!}$. Since

$$\mathbb{E}\left(\begin{pmatrix} N_{\rho_n}^n \\ r \end{pmatrix}\right) = \sum_{k_1 < k_2 < \dots < k_r \leqslant n} \mathbb{P}(A_{\rho_n}^{k_1, \dots, k_r}), \tag{2.18}$$

Lemma 2.5 implies for each r

$$\lim_{n\to\infty} \mathbb{E}\left(\begin{pmatrix} N_{\rho_n}^n \\ r \end{pmatrix}\right) = \frac{\lambda^r}{r!} = \mathbb{E}\left(\begin{pmatrix} \mathcal{X} \\ r \end{pmatrix}\right). \tag{2.19}$$

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. □

Similarly to Borel-Cantelli Lemma, we also have the following extension of Theorem 2.8 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 $N^{n,i} := N_{[0,1]}^{n,i}$.

Theorem 2.9. *Suppose that* $\widetilde{(M1)}_r$ *and* $(M2)_r$ *hold for all* r *and that*

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

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

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

Proof. It suffices to prove the second statement. The proof is similar to the proof of Theorem 2.8. Namely, similarly to (2.19) we show that for each set $r_{ij}\in\mathbb{N}$ we have

$$\lim_{n\to\infty}\mathbb{E}\left(\prod_{i,j}\begin{pmatrix} N^{n,i}_{I_j} \\ r_{ij}\end{pmatrix}\right) = \prod_{ij}\frac{(\lambda_i|I_j|)^{r_{ij}}}{(r_{ij})!} = \prod_{ij}\mathbb{E}\begin{pmatrix}\mathcal{X}_{ij}\\ r_{ij}\end{pmatrix} \tag{2.20}$$

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

2.7. **Notes.** The usual Borel Cantelli Lemma is a classical subject in probability. There are many extensions to weakly dependent random variables, see e.g. [134, §12.15], [129, §1]. The connection between Borel-Cantelli Lemma and Poisson Limit Theorem is discussed in [49, 55] There is also a vast literature on Borel-Cantelli Lemmas for dynamical systems starting with [118]. Some representative examples dealing with hyperbolic systems are [5, 34, 54, 68, 71, 79, 73, 74, 87, 100] while [28, 29, 87, 93, 94, 95, 104, 130] deal with systems of zero entropy. The later cases is more complicated as counterexamples in [53, 64] show. Survey [6] reviews the results obtained up to 2009 and contains many applications, some of which parallel the results obtained in Sections 4–8 of the present paper. The multiple Borel Cantelli Lemma for independent events is proven in [114]. [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.

3. Multiple Borel Cantelli Lemma for exponentially mixing dynamical systems.

3.1. **Good maps, good targets.** Let f be a map of metric space X preserving a measure μ. Given a family of sets $\Omega_\rho\subset X$, we will, in a slight abuse of notations, sometimes call Ω_ρ the event 1_{Ω_ρ} and Ω^n_ρ the event $1_{\Omega_\rho}\circ f^n$. Recall the notation $\sigma(\rho) = \mu(\Omega_\rho)$.

To deal with recurrence, we need to consider slightly more complicated events.

Given a family of events $\bar\Omega_\rho$ in $X\times X$ let $\bar\Omega^n_\rho\subset X$ be the event

$$\bar\Omega^n_\rho = \{x : (x, f^n x)\in\bar\Omega_\rho\}.$$

Let

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

where $\mathbf{v}_j = \sigma(\rho_{2^j})$ for targets Ω^n_ρ and $\mathbf{v}_j = \bar\sigma(\rho_{2^j}) = \mu\times\mu(\bar A_\rho)$ for the targets $\bar\Omega^n_\rho$.

Let N_ρ^n be the number of times $k \leqslant n$ such that Ω_ρ^k (or $\bar{\Omega}_\rho^k$) occurs. We want to give conditions on the system (f, X, μ) and on a sequence of targets $\Omega_{\rho_n}^k$ (or $\bar{\Omega}_{\rho_n}^k$), that imply the validity of the dichotomy of Theorem 2.2 for the number of hits $N_{\rho_n}^n$.

Definition 3.1 (r-fold exponentially mixing systems)**.** *Let $\mathbb{B}$ be a space of real valued functions defined over X^{r+1}, with a norm $\|\cdot\|_{\mathbb{B}}$. We say that $(f, X, \mu, \mathbb{B})$ is r-fold exponentially mixing, if there exists 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) $\|A \circ (f^{k_1}, \dots, f^{k_r})\|_{\mathbb{B}} \leqslant C L^{\sum_{i=1}^r k_i}\|A\|_{\mathbb{B}}$,

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

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

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 collection of sets in X there are positive constants $\bar{C}, \eta, \tau$ 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* $\|A_\rho^\pm\|_{\mathbb{B}} \leqslant \bar{C}\rho^{-\tau}$;

(ii) $A_\rho^- \leqslant 1_{\Omega_\rho} \leqslant A_\rho^+$;

(iii) $\mu(A_\rho^+) - \mu(A_\rho^-) \leqslant \bar{C}\sigma(\rho)^{1+\eta}$.

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 if there exists $u > 0$ such that for some $u > 0$

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

and

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

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

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

Remark 3.4. A typical situation where one wants to verify these properties is the following. Suppose that f is Lipshitz and $\mathbb{B}$ is the space of Lipshitz functions. Then (Prod) is clear and (Gr) holds with L being the Lipshitz constant of f. Moreover in case

$$\Omega_\rho = \{\Phi(x) \in [a_1(\rho), a_2(\rho)]\}, \text{ and } \mu(\Omega_\rho) \geqslant \rho^{-a}, \text{ for some } a > 0,$$

and Φ is some Lipschitz function, one can often verify (Appr) by taking $A_\rho^+ = \phi^+(\Phi)$, $A_\rho^- = \phi^-(\Phi)$ where $\phi^\pm : \mathbb{R} \to \mathbb{R}$ are Lipshitz functions such that

$$1_{[a_1+\rho^s, a_2-\rho^s]} \leqslant \phi^- \leqslant 1_{[a_1,a_2]} \leqslant \phi^+ \leqslant 1_{[a_1-\rho^s, a_2+\rho^s]}$$

for some $s > 0$.

To deal with recurrence, the following definition is useful.

Definition 3.5 (Composite admissible targets)**.** *Let $\bar\Omega_\rho$ and be a collection of sets in $X\times X$ satisfying the following conditions for some positive constants $\bar C,\eta,\tau$.*

For any $\rho>0$

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

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

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

(iii) *There is a function $\bar\sigma(\rho)$ such that for any fixed x,*

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

(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

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

and there is a constant a such that

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

and

$$\bar\Omega^{k_1}_\rho\cap\bar\Omega^{k_2}_\rho\subset f^{-k_1}\bar\Omega^{k_2-k_1}_{2\rho}\ \text{for } k_1<k_2. \tag{$\overline{\text{Sub}}$}$$

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

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

3.2. **Multiple Borel-Cantelli Lemma for admissible targets.** The goal of this section is to establish the following.

Theorem 3.6. *Assume (f,X,μ) is $(2r)$-fold exponentially mixing. Then*
a) If Ω_{ρ_n} are as in Definition 3.2 then $(M1)_r$, $(M2)_r$ and $(M3)_r$ hold for the sequence Ω_{ρ_n} and the probability measure μ.
b) If $\bar\Omega_{\rho_n}$ are as in Definition 3.5 then $(M1)_r$, $(M2)_r$ and $(M3)_r$ hold for the sequence $\bar\Omega_{\rho_n}$ and the probability measure μ.

Hence, Theorem 2.2 implies

Corollary 3.7. *If (f,X,μ) is $(2r)$-fold exponentially mixing, and if Ω_{ρ_n} (or $\bar\Omega_{\rho_n}$) are as in Definition 3.2 (or Definition 3.5), then*

(a) *If $\mathbf{S}_r<\infty$, then with probability 1, we have that for large n $N^n_{\rho_n}<r$.*

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

In fact Theorem 3.6 is a direct consequence of the following Proposition.

Proposition 3.8. *Given a dynamical system (f,X,μ) and a sequence of decreasing sets Ω_{ρ_n} such that (Poly) holds, then:*

(i) If (Prod), $(\text{EM})_r$ *and* (Appr) *hold 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 $r,\theta,C,\bar C,\sigma$ and η.*

(ii) If (Prod), (Gr), $(\text{EM})_r$, (Appr) *and* (Mov) *hold, then $(M2)_r$ is satisfied.*

(iii) If (Prod), (Gr), $(\mathrm{EM})_{2r}$, *and* (Appr) *hold, then for arbitrary* $\varepsilon > 0$ $(M3)_r$ *is satisfied with* $\hat{\mathfrak{s}}(n) = \varepsilon n$.

Similarly, given a dynamical system (f, X, μ) *and a sequence of decreasing sets* $\bar{\Omega}_{\rho_n}$ *such that* $(\overline{\mathrm{Poly}})$ *holds, then:*

(i) If $(\overline{\mathrm{Prod}})$, $(\mathrm{EM})_r$ *and* $(\overline{\mathrm{Appr}})$ *hold 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* $r, \theta, C, \bar{C}, \sigma$ *and* η.

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

(iii) If $(\overline{\mathrm{Prod}})$, (Gr), $(\mathrm{EM})_{2r}$, *and* $(\overline{\mathrm{Appr}})$ *hold, then for arbitrary* $\varepsilon > 0$ $(M3)_r$ *is satisfied with* $\hat{\mathfrak{s}}(n) = \varepsilon n$.

Proof of Proposition 3.8. 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.

(i) For Ω_{ρ_n}, we prove (M1) in case $k_{i+1} - k_i \geqslant \sqrt{R} \ln n$ where R is a sufficently large constant.

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

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

Likewise

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

This proves (M1), since $\rho_n \geqslant n^{-u}, \mu(\Omega_{\rho_n}) \geqslant n^{-u}$, and R is sufficiently large.

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

$$\mu\left(\bigcap_{i=1}^{r} \bar{\Omega}^{k_i}_{\rho_n}\right) \leqslant \mu\left(B^+_{\rho_n}(x_0, f^{k_1}x_0, \cdots, f^{k_r}x_0)\right) \leqslant \mu\left(B^+_{\rho_n}(x_0, \cdots, x_r)\right) + C{\rho_n}^{-r\sigma} n^{\sqrt{R}\ln\theta}$$

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

Likewise,

$$\mu\left(\bigcap_{i=1}^{r} \bar{\Omega}^{k_i}_{\rho_n}\right) \geqslant \left(\bar{\sigma}(\rho_n) - C\bar{\sigma}(\rho_n)^{1+\eta}\right)^r - C{\rho_n}^{-r\sigma} n^{\sqrt{R}\ln\theta}$$

This proves $(M1)_r$ for the composite admissible sequence of targets $\bar{\Omega}_{\rho_n}$ by taking R large.

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

So we assume that $0 < k_j - k_{j-1} < R\ln n$ and $k_i - k_{i-1} \geqslant R\ln n$ 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 $(\mathrm{EM})_r$ and the fact that $\rho_n \geqslant n^{-u}, \mu(\Omega_{\rho_n}) \geqslant n^{-u}$,

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

$$\leqslant \mu\left(\prod_{i\neq j-1,j} A^+_{\rho_n}\left(f^{k_i}x\right)\right)\mu(B) + C{\rho_n}^{-\sigma r}L^{\sqrt{R}\ln n}\theta^{R\ln n} \leqslant C\mu(\Omega_{\rho_n})^{r-1}(\ln n)^{-1000r}$$

proving $(M2)_r$.

For $\bar\Omega_{\rho_n}$, as the proof of (i), we approximate $1_{\bar\Omega_{\rho_n}}$ by $A^+_{\rho_n}$. Consider

$$\tilde B_r(x_0,\cdots,x_{j-1},x_{j+1},\cdots,x_r)$$
$$= 1_{\bar\Omega_{\rho_n}}(x_0,x_1)\cdots 1_{\bar\Omega_{\rho_n}}(x_0,x_{j-1})1_{\bar\Omega^{k_j-k_{j-1}}_{2\rho_n}}(x_{j-1})1_{\bar\Omega_{\rho_n}}(x_0,x_{j+1})\cdots 1_{\bar\Omega_{\rho_n}}(x_0,x_r).$$

$$\hat B_r(x_0,\cdots,x_{j-1},x_{j+1},\cdots,x_r)$$
$$= A^+_{\rho_n}(x_0,x_1)\cdots A^+_{\rho_n}(x_0,x_{j-1})A^+_{2\rho_n}(x_{j-1},f^{k_j-k_{j-1}}x_{j-1})A^+_{\rho_n}(x_0,x_{j+1})\cdots A^+_{\rho_n}(x_0,x_r),$$

Since $(\overline{\mathrm{Appr}})$, $(\overline{\mathrm{Mov}})$ and $(\overline{\mathrm{Sub}})$ hold, we obtain

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

Integrating with respect to all all variables except for x_{j-1} we get

$$\mu^{r+1}(\hat B_r) \leqslant \left(\bar\sigma(\rho_n) + \bar\sigma(\rho_n)^{1+\eta}\right)^r \bar C\mu(\bar\Omega^{k_j-k_{j-1}}_{2\rho_n}). \tag{3.1}$$

Therefore, $(M2)_r$ follows when $\overline{(\mathrm{Mov})}$ holds.

(iii) Choose a large constant b and consider two cases.

Case 1 : $j \leqslant i+b$**.** then the proof of $(M3)_r$ is the same as the proof of $(M1)_r$ except that we need to use $(EM)_{2r}$ instead of $(EM)_r$.

Case 2 : $j > i+b$**.** Consider first simple targets Ω_{ρ_n}, Denoting $B(x) = \prod_{\alpha=1}^r A^+_{\rho_{2^i}}(f^{k_\alpha}x)$ we obtain from (Prod), (Gr), (Appr) and (Poly) so that

$$\|B\|_{\mathbb{B}} \leqslant CL^{2^{i+1}}.$$

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

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

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

Applying already established $(M1)_r$ to estimate $\mu(B)$, and observing that the second term is smaller than $C(L^{1/2^{b-1}})^{2^j}2^{\sigma rj}\theta^{2^j\varepsilon}$, which is thus much smaller than the first, we see that if b is sufficiently large $(M3)$ follows.

Next, we analyze Case 2 for $\bar{\Omega}_{\rho_n}$. Consider

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

By $(\overline{\text{Appr}})$ and $\overline{(\text{EM})}_{\text{r}}$, we get

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

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

Using $(M1)_r$ we conclude that

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

This completes the proof of $(M3)_r$. □

3.3. **Notes.** We refer the reader to Appendix A for more background on multiple exponential mixing as we for some examples. 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 [20, 32, 127]. [61] obtains Logarithm Law for hitting times under an assumption of superpolynomial mixing which is weaker than our exponentially mixing assumption.

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.

4.1. **Results.** Let f be a map of a compact $d-$dimensional Riemannian manifold M preserving a smooth measure μ. 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

$$\text{For almost all } x, \quad \limsup_{n\to\infty}\frac{|\ln d_n^{(1)}(x,x)|-\frac{1}{d}\ln n}{\ln\ln n}=\frac{1}{d}, \tag{4.1}$$

$$\text{For all } x \text{ and almost all } y, \quad \limsup_{n\to\infty}\frac{|\ln d_n^{(1)}(x,y)|-\frac{1}{d}\ln n}{\ln\ln n}=\frac{1}{d}. \tag{4.2}$$

In particular, the following result is a special case of [61, Corollary 9].

Theorem 4.1. *Suppose that f is smooth, preserves a smooth measure μ and has superpolynomial decay of correlations. That is,*

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

Then (4.1) and (4.2) hold.

MultiLog Law for recurrence and for hitting times. The goal of this section is to obtain an analogue of (4.1) (4.2) for $d_n^{(r)}$ with $r>1$, for exponentially mixing systems as in Definition 3.1.

Given a 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 4.2. *Suppose that $(f,M,\mu,\mathbb{B})$ is an r-fold exponentially mixing system. 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). It turns out that for larger r the answer is often negative. Define

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

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

Theorem 4.3. *Suppose that $(f,M,\mu,\mathbb{B})$ is a 1-fold exponentially mixing system and that the periodic points of f are dense. Then for any $r\in\mathbb{N}$, we have*
a) $\mathcal{H}$ contains a G_δ dense set.
b) $\bar{\mathcal{H}}$ 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 4.3 emphasizes the necessity of a restriction on x in Theorem 4.2.

In a similar spirit, we show that strong 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 4.4. *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 $r \geqslant 2$;
b) $\bar{\mathcal{E}}_r = M$ for all $r \geqslant 1$.

The proof requires different techniques, so it will be given in Section 8.

The case of flows. Here we describe the analogue results of Theorems 4.2 and 4.3 for flows. Let ϕ be a piecewise 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 ball. Namely, for some fixed $\rho > 0$ let $[t_i^-, t_i^+]$ be consecutive connected components such that $\phi^t x \in B(x, rho)$ for $t \in [t_i^-, t_i^+]$. Let t_i be the argmin of $d(\phi^t y, x)$ for $t \in [t_t^-, t_i^+]$. Let $d_n^{(r)}(x,y)$ be the $r-th$ minimum of

$$d(x, \phi^{t_1}(y)), \dots, d(x, \phi^{t_k}(y)), \quad t_k \leqslant n < t_{k+1}. \tag{4.3}$$

Theorem 4.5. *Suppose that for each $a > 0$, $(\phi^a, M, \mu, \mathbb{B})$ is an r-fold exponentially mixing system. 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) $\mathcal{H}$ contains a G_δ dense set;
d) $\bar{\mathcal{H}}$ contains a G_δ dense set.

4.2. **Weak non-recurrence and proof of Theorem 4.2.** In the following statement $r \geqslant 2$ is fixed, $\mathbb{B}$ is supposed to be equal to the space of C^1 functions on M^{r+1}. Since μ is a smooth measure there is a smooth function $\gamma(x)$ such that

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

Given $x \in M$, $c \in \mathbb{R}_+$ let

$$\Omega_{\rho,x} = \{y : d(x,y) \leqslant \rho\} \tag{4.5}$$

and

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

Proposition 4.6. *For each x, the targets $\Omega_{\rho,x}$ satisfy* (Appr) *and the targets $\bar{\Omega}_\rho$ satisfy* ($\overline{\text{Appr}}$)

Proof. For targets $\Omega_{\rho,x}$ the statement follows from Remark 3.4.

To handle the composite targets Denote $\Phi(x,y) = \frac{d(x,y)\gamma(x)^{1/d}}{c}$. For $s > 1$ let $\hat{\phi}^\pm(x,y)$ be Lipschitz functions such that

$$1_{B(0,\rho-\rho^s)} \leqslant \hat{\phi}^- \leqslant 1_{B(0,\hat{\rho})} \leqslant \hat{\phi}^+ \leqslant 1_{B(0,\rho+\rho^s)}.$$

Take $A^\pm_\rho = \hat{\phi}^\pm(\Phi)$. Then

$$(\rho-\rho^s)^d \leqslant \int 1_{B(0,\rho-\rho^s)}(\Phi(x,y))d\mu(y) \leqslant \int A^-_\rho(x,y)d\mu(y)$$
$$\leqslant \int A^+_\rho(x,y)d\mu(y) \leqslant \int 1_{B(0,\hat{\rho}+\hat{\rho}^s)}(\Phi(x,y))d\mu(y) \leqslant (\rho+\rho^s)^d.$$

Hence ($\overline{\text{Appr}}$)(i−iii) follow.

Denote $C_0 = \sup_{x\in M}\gamma(x)/\gamma(y)$, we get

$$\int A^+_\rho(x,y)d\mu(x) \leqslant \int 1_{B(0,\rho+\rho^s)}(\Phi(x,y))d\mu(x) \leqslant C_0\int 1_{B(0,\rho+\rho^s)}(\Phi(x,y))d\mu(y) \leqslant C_0(\rho+\rho^s)^d$$

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

To check (Mov) and ($\overline{\text{Mov}}$) we need the following definitions.

Definition 4.7 (Weakly non-recurrent points)**.** *Call x* weakly non-recurrent *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*

$$\mu(B(x,\rho)\cap f^{-n}B(x,\rho)) \leqslant \mu(B(x,\rho))|\ln\rho|^{-A}. \tag{4.7}$$

Definition 4.8 (Weakly non-recurrent system)**.** *Call the system $(f, M, \mu, \mathbb{B})$* weakly non-recurrent *if for each $A > 0$ $\exists\rho_0$ such that for all $\rho < \rho_0$ for all $n \in \mathbb{N}^*$ we have*

$$\mu\left(\{x : d(x, f^n x) < \rho\}\right) \leqslant |\ln\rho|^{-A}. \tag{4.8}$$

Proposition 4.9. *Suppose that $(f, M, \mu, \mathbb{B})$ is an r-fold exponentially mixing system. Then*

i) $(f, M, \mu, \mathbb{B})$ is weakly non-recurrent.
ii) Almost every point is weakly non-recurrent.

Proof. Take $B := A^2$. If $k \geqslant B\ln|\ln\rho|$, take $\hat{\rho} = |\ln\rho|^{-A}$.

By exponential mixing, we get

$$\mu(x : d(x, f^k x) \leqslant \rho) \leqslant \mu(x : d(x, f^k x) \leqslant \hat{\rho}) \tag{4.9}$$
$$\leqslant \mu\left(\bar{A}^+_{\hat{\rho}}(x, f^k x)\right) \leqslant C\left(\hat{\rho}^d + \hat{\rho}^{d+d\eta} + \frac{1}{\hat{\rho}^\sigma}\bar{\theta}^k\right) \leqslant |\ln\rho|^{-2A},$$

provided ρ is sufficiently small.

Now fix any $k \leqslant B \ln|\ln\rho|$. 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 (4.9) 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

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

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 $ii)$. □

Proof of Theorem 4.2. a). If we take $\rho_n = n^{-\frac{1}{d}}\ln^{-s} n$, we see that for almost every x, $\{\Omega_{\rho_n,x}\}$ is a simple admissible sequence of targets as in Definition 3.5 ((Appr) follows from Proposition 4.6(a) and (Mov) follows from Proposition 4.9ii)). Moreover, with the notation $\mathbf{v}_j = \sigma(\rho_{2^j})$ and $\mathbf{S}_r = \sum_{j=1}^{\infty}(2^j\mathbf{v}_j)^r$, we see that $\mathbf{S}_r = \infty$ iff $s \geqslant rd$. Hence, Theorem 4.2 a) follows from Corollary 3.7.

b). Take $\rho_n = n^{-\frac{1}{d}}\ln^{-s} n$. Note that when $x \in \bar{\Omega}_\rho^{k_1} \cap \bar{\Omega}_\rho^{k_2}$, we have

$$d(f^{k_1}x, f^{k_2}x) \leqslant d(x, f^{k_1}x) + d(x, f^{k_2}x) \leqslant \frac{2c\rho}{(\gamma(x))^{1/d}} \leqslant \frac{c_1\rho}{(\gamma(f^{k_1}x))^{1/d}}.$$

Hence ($\overline{\text{Sub}}$) is satisfied. Combining this with Propositions 4.6(b) and 4.9(i), we know that $\{\bar{\Omega}_{\rho_n}\}$ defined by (4.6) is a composite admissible sequence of targets. Thus we get b) by a similar argument to a). □

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

Proof. To prove part a), we first prove that periodic points belong to $\mathcal{H}_r$. We know from Theorem 4.1 that for any $x \in M$ and almost every y,

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

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

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

To prove the opposite inequality let

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

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 $\mathcal{H}_{l,m}$ is an open set. Hence we finish if we show that for any fixed r and l, $\bigcup_m \mathcal{H}_{m,l,r}$ contains the dense set of periodic points.

Let $\bar{x}$ be a periodic point of period p. Take U to be some small neighbourhood of $\bar{x}$ and denote by Λ the Lipschitz constant of f^p in U.

By (4.10), there exists $n \geqslant \exp\circ\exp(\Lambda + l)$ 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 $n \geqslant \exp\circ\exp(\Lambda + r)$. Hence $\bar{x} \in \mathcal{H}_{m,l}$ and the proof of (a) is finished.

We now turn to the proof of (b). Define

$$\mathcal{A}_{m,l} = \left\{ x : |\ln d_m^{(l)}(x,x)| > e^{2m} \right\}.$$

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

4.4. **The case of flows. Proof of Theorem 4.5.** The proof of Theorem 4.5 for flows proceed in the same way as for maps with minimal modifications. Namely, let

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

and

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

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

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

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

Fix $A >> 1$ arbitrarily large.

Definition 4.10 (Weakly non-recurrent points for flow)**.** *Call x* weakly non-recurrent *if for each $A, K > 0$, there exists $\rho_0 > 0$ such that for all $\rho < \rho_0$ for all $n \in \mathbb{N}$ and $0 < n < K|\ln\rho|$,*

$$\mu\left(\Omega_{\rho,x} \cap \Omega^n_{\rho,x}\right) \leqslant \mu(\Omega_{\rho,x})|\ln\rho|^{-A}. \tag{4.12}$$

Definition 4.11 (Weakly non-recurrent flow)**.** *Call the system $(\phi, M, \mu, \mathbb{B})$* weakly non-recurrent *if for each $A > 0$ $\exists \rho_0 > 0$ such that for all $\rho < \rho_0$ and for all $n \in \mathbb{N}^*$ we have*

$$\mu\left(\bar{\Omega}^n_\rho\right) \leqslant |\ln\rho|^{-A}. \tag{4.13}$$

We proceed analogously to the proof of flow version of Proposition 4.9.

Proposition 4.12. *Suppose that for any $a > 0$, $(\phi^a, M, \mu, \mathbb{B})$ is an exponentially mixing system. Then*

i) $(\phi^a, M, \mu, \mathbb{B})$ is weakly non-recurrent.
ii) Almost every point is weakly non-recurrent.

Proof of Proposition 4.12. Take $B = A^2$. If $k \geqslant B\ln|\ln\rho|$, the same argument as in the proof of Proposition 4.9 gives that

$$\mu\left(\bar{\Omega}^k_\rho\right) \leqslant |\ln\rho|^{-A}. \tag{4.14}$$

Now fix any $k < B\ln|\ln\rho|$. Denote $\|\phi\|_1 = \max_{t\in[1,2]}\|D\phi^t\|$. Assume that x satisfies $d(x, \phi^{k+t}x) \leqslant \rho$ for some $0 \leqslant t \leqslant 1$, then for any l we have that

$$d(\phi^{(l-1)(k+t)}(x), \phi^{l(k+t)}x) \leqslant \|\phi\|_1^{(l-1)k}\rho$$

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

$$d(x, \phi^{L(k+t)}x) \leqslant \sum_{l\leqslant L-1} \|\phi\|_1^{lk}\rho \leqslant \sqrt{\rho},$$

provided ρ is sufficiently small. But $kL \geqslant B\ln|\ln\rho|$, hence (4.14) applies and we get

$$\mu\left(\bar{\Omega}^k_\rho\right) \leqslant \sum_{m=0}^{L-1} \mu\left(\bar{\Omega}^{Lk+m}_{\sqrt{\rho}}\right) \leqslant 4B(\ln|\ln\rho|)|\ln\sqrt{\rho}|^{-A} \leqslant |\ln\rho|^{-A/2},$$

proving $i)$.

We proceed now to prove *ii*). Denote $M_0 = \max_{t\in[-1,1]} \|D\phi_t\|$,

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

where $\gamma(x) = \lim_{\rho\to 0} \frac{\mu(\Omega_\rho)}{\rho^d}$ and $\tilde{\Omega}^n_\rho = \{x : (x,\phi^n(x)) \in \tilde{\Omega}_\rho\}$. Define for $j,k \in \mathbb{N}^*$,

$$H_{j,k}(x) = \mu\left(\Omega_{1/2^j,x} \cap \Omega^k_{1/2^j,x}\right).$$

Note that

$$\begin{aligned}
\int H_{j,k}(x)d\mu(x) &= \int 1_{\bar{\Omega}_{\gamma(x)/2^j}}(x,y) 1_{\bar{\Omega}_{\gamma(x)/2^j}}(x,f^k y) d\mu(x)d\mu(y) \\
&\leqslant M_0 C_1 \int 1_{\bar{\Omega}_{\gamma(x)/2^j}}(x,y) 1_{\tilde{\Omega}^k_{2/2^j}}(y) d\mu(x)d\mu(y) \\
&\leqslant M_0 C_2 \int \mu\left(\Omega_{\gamma(x)/2^j,y}\right) 1_{\tilde{\Omega}^k_{2/2^j}}(y) d\mu(y) \\
&\leqslant M_0 C_3 \mu(\Omega_{1/2^j,x}) \mu\left(\tilde{\Omega}^k_{2/2^j}\right)
\end{aligned}$$

where we used that $C^{-1} \leqslant \gamma(x) \leqslant C$ for any $x \in M$ and $\mu(B(y,\rho)) \leqslant C\mu(B(x,\rho))$ for any $x,y \in M$.

Part *i*) then implies that for sufficiently large j it holds that

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

For such j we get from Markov inequality

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

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

Proof of Theorem 4.5. If the point x is weakly non-recurrent, we see that $\Omega_{\rho_n,x}$ is an simple admissible sequence as in Definition 3.2 by taking $\rho_n = n^{-\frac{1}{d}} \ln^{-s} n$. Moreover, with the notation $\mathbf{v}_j = \sigma(\rho_{2^j})$ and $\mathbf{S}_r = \sum_{j=1}^\infty \left(2^j \mathbf{v}_j\right)^r$, we see that $\mathbf{S}_r = \infty$ iff $s \geqslant rd$. Hence, $x \in \mathcal{G}_r$.

The proof of part b) is similar when we replace the targets $\Omega_{\rho_n,x}$ by $\bar{\Omega}_{\rho_n}$ and apply the same argument. It remains to verify ($\overline{\text{Sub}}$). Note that when $x \in \bar{\Omega}^{t_1}_\rho \cap \bar{\Omega}^{t_2}_\rho$ for $t_1 < t_2$, we have some s_1, $s_2 \in [0,1]$ such that

$$d(x,\phi^{t_1+s_1}x) \leqslant \frac{c\rho}{(\gamma(x))^{1/d}}, \quad d(x,\phi^{t_2+s_2}x) \leqslant \frac{c\rho}{(\gamma(x))^{1/d}}.$$

Hence

$$d(\phi^{t_1}x,\phi^{t_2+s_2-s_1}x) \leqslant M_0 d(\phi^{t_1+s_1}x,\phi^{t_2+s_2}x) \leqslant \frac{2M_0 c\rho}{(\gamma(x))^{1/d}} \leqslant \frac{c_1\rho}{(\gamma(\phi^{t_1}x))^{1/d}}.$$

It follows that $\bar{\Omega}^{t_1}_\rho \cap \bar{\Omega}^{t_2}_\rho \in \phi^{-t_1}\tilde{\Omega}^{t_2-t_1}_\rho$, which gives $(M2)_r$ by a similar argument of Proposition 3.8 (ii).

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

4.5. **Notes.** Many authors obtain Logarithm Law (4.2) for hitting times as a consequence of dynamical Borel-Cantelli Lemmas. See [34, 42, 60, 79] and references wherein. [61] studies also return times. We note that [61] obtains (4.1) and (4.2) under much weaker conditions than those imposed in the present paper, however, his results are valid only for $r = 1$ (the first visit).

[87, 92, 97] study the recurrence problem in the case in the case lim sup in (4.2) is replaced by lim inf. In particular, [97] proves that for several expanding maps the

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

exists for almost all y.

5. Poisson Law for near returns.

5.1. **Introduction.** In this section we suppose that μ is a smooth measure and that (f, M, μ, C^1) is an r-fold exponentially mixing system for all r. In the previous section we verified properties $(M1)_r$ and $(M2)_r$ for targets $\Omega_{\rho,x}$ given by (4.5), for almost every x, and for targets $\bar{\Omega}_\rho$ given by (4.6). Accordingly Theorem 2.8 gives

Theorem 5.1. *(a) For almost all x the following holds. Let y be uniformly distributed with respect to μ the number of visits to $B(x,\rho)$ up to time $\tau\rho^{-d}$ converges to a Poisson distribution with parameter $\tau\gamma(x)$ as $\rho \to 0$. Moreover letting $n = \tau\rho^{-d}$ we have he sequence*

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

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

(b) Let x be chosen uniformly with respect to μ. Then the number of visits to $B\left(x, \dfrac{\rho}{\gamma^{1/d}(x)}\right)$ up to time $\tau\rho^{-d}$ 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 2.8. To prove the Poisson limit for (5.1) we need to check that for each $r_1^- < r_1^+ < r_2^- < r_2^+ < \dots < r_s^- < r_s^+$ the number of times where $d(f^k y, x) \in \left[r_j^-\rho, r_j^+\rho\right]$ are independent Poisson random variables with parameters

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

but this follows from Theorem 2.9. This theorem can be applied since $\widetilde{(M1)}_r$ follows from condition (Appr) for targets

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

and the (Appr) holds due to Remark 3.4. □

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 law for hits hold. 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 to $B(x,\rho)$ up to time $\tau\rho^{-d}$ converges to a mixture of Poisson distributions. Namely, for each l*

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

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

Corollary 5.4. *If f preserves a smooth measure and is r fold exponentially mixing on C^1 for all r 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). □

5.2. **Poisson Law for Returns.**

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 $\hat\Omega^k_\rho = \{x : (x, f^k x)\in\hat\Omega_\rho\}$. Note that $(M2)_r$ for $\bar\Omega^k_\rho$ implies $(M2)_r$ for $\hat\Omega^k_\rho$. However, $(M1)_r$ is false for targets $\hat\Omega^k_\rho$. We now argue similarly to the proof of Theorem 3.6 to obtain that for separated tuples $k_1, k_2, \dots, k_r$,

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

Namely, note that

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

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

Thus approximating $1_{\hat{\Omega}_\rho}$ by $\hat{A}^{\pm}_\rho$ satisfying $(\overline{\text{Appr}})$, and applying $\overline{(\text{EM})}_{\text{r}}$ 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

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

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

and, likewise,

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

Taking R large we obtain (5.3).

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

$$\lim_{\rho\to 0}\int_M \begin{pmatrix} N^r_{\rho,\tau,x} \\ r \end{pmatrix} d\mu(x) = \int_M \frac{\lambda^r(z)}{r!}d\mu(z) \tag{5.4}$$

where

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

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

5.3. **Notes.** Early works on Poisson Limit Theorems for dynamical systems include [35, 40, 80, 81, 82, 119]. [30, 75, 78, 120] 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. [42] obtains Poisson Limit Theorem for partially hyperbolic systems. Some of those papers, including [27, 42, 75] 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 [112, 116]. Convergence of 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 [42, 56, 57]. A mixed exponential distribution for a return time for dynamical systems similar to Corollary 5.4 has been obtained in [37] in a symbolic setting. For more discussion of the distribution of the entry times to small measure sets we refer the readers to [36, 86, 123, 137] and references wherein.

6. Gibbs measures on the circle: Law of iterated logarithm for 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.

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

This means (see [125] 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

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

then

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

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

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

and by $\mathbf{d}$ the dimension of the point x,

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

We know from [106] that the limit exists for μ-a.e. x and $\mathbf{d} = \frac{h_\mu(f)}{\lambda} = -\frac{\mu(g)}{\mu(f_u)}$ with the notation $f_u = \ln |f'|$.

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

$$\frac{1}{K} \leqslant \frac{\mu(B(x, r))}{r^{\mathbf{d}}} \leqslant K. \tag{6.1}$$

It is known (see e.g. [117]) that μ is conformal iff g can be represented in the form

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

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

Theorem 6.1. *(a) If μ is conformal then Theorems 4.2 and 4.3 remain valid with d replaced by $\mathbf{d}$.*

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

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x, x)| - \frac{1}{\mathbf{d}} \ln n}{\sqrt{2(\ln n)(\ln\ln\ln n)}} = \frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}, \tag{6.2}$$

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x, y)| - \frac{1}{\mathbf{d}} \ln n}{\sqrt{2(\ln n)(\ln\ln\ln n)}} = \frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}. \tag{6.3}$$

6.2. **Good targets for Gibbs measures.** Here we prepare for the proof of Theorem 6.1 by showing that certain targets are good for expanding maps equipped with a Gibbs measure. We caution the reader that the targets $\bar{\Omega}_\rho$ from Section 4 are **not** good targets in the non-conformal case so we need to use a roundabout approach to proving (6.2). We collect several useful facts.

Proposition 6.2. *For each (f, μ) enjoys r fold exponential mixing with respect to Lipschitz functions.*

The proof is given in §A.2.

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

$$r^a \leqslant \mu(B(x, r)) \leqslant r^b. \tag{6.4}$$

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

$$\mu(B(x, 4\rho) \leqslant R\mu(x, \rho). \tag{6.5}$$

Proposition 6.3. *For any Gibbs measure μ for μ almost all x the targets $B(x, \rho_n)$ are admissible provided that $\rho_n > n^{-u}$ for some u.*

Proof. (Poly) follows from (6.4).

Next take a large s and consider $A^\pm_\rho$ such that

$$1_{B(x,\rho-\rho^s)} \leqslant A^-_\rho \leqslant A^+_\rho \leqslant 1_{B(x,\rho+\rho^s)}.$$

Then $A^+_\rho - A^-_\rho$ is supported on two segments of length $2\rho^s$ showing that

$$\mu(A^+_\rho - A^-_\rho) \leqslant \rho^{bs} \leqslant \rho^a \leqslant \mu(B(x, \rho))$$

if s is large proving (Appr).

It remains to show that (Mov) holds for a. e. x. Namely we show that for a.e. x and all k

$$\mu(B(x, \rho) \cap f^{-k}B(x, \rho)) \leqslant \mu(B(x, \rho))^{1+\eta}. \tag{6.6}$$

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 $A^+_\rho = 0$ on $B(x, \rho)$, $\mu(A^+_\rho) - \mu(B(x, \rho)) \leqslant 2\mu(B(x, \rho))$ 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

$$\mu(B(x, \rho) \cap f^{-k}B(x, \rho)) \leqslant \mu(A^+_{\hat{\rho}}(x)(A^+_\rho \circ f^k))$$
$$\leqslant 4\mu(B(x, \rho))\mu(B(x, \hat{\rho})) + C\theta^k\hat{\rho}^{-\tau}\mu(B(x, \rho)) = C\mu(B(x, \rho))\left[\mu(\rho^{\sigma b} + \rho^\varepsilon\rho^{-\tau\sigma}\right].$$

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

(II) $k > \varepsilon|\ln \rho|$. In this case for a.e. x the intersection $B(x, \rho) \cup f^{-k}B(x, \rho)$ is empty for small ρ due to the Proposition 6.4 below. □

Proposition 6.4. *([14, 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 $n_0(x)$ such that for all $n \geqslant n_0(x)$, and all $0 < k \leqslant \frac{1}{2L}|\ln \rho|$, we have $T^{-k}B(x,\rho) \cap B(x,\rho) = \varnothing$.*

Proposition 6.5. *If μ is conformal, that is $B(x,\rho) = \gamma(x)\rho^{\mathbf{d}}(1 + o_{\rho\to 0}(1))$ then the targets $\bar{\Omega}_{\rho_n}$ defined by (4.6) are admissible provided that $\rho_n > n^{-u}$ for some u.*

The proof of this fact is the same as in the smooth measure case, so we leave it to the reader.

Proof of Theorem 6.1(a). Given Propositions 6.2, 6.3 and 6.5 the proof is exactly the same as in Section 4.

Consider for example, the MultiLog Law for hitting times. Taking $\rho_n = n^{-\frac{1}{d}} \ln^{-s} n$, we see that for almost every x, $\Omega_{\rho_n,x} = \{y : d(x,y) \leqslant \rho_n\}$ is a simple admissible sequence of targets as in Definition 3.2. Moreover, with the notation $\mathbf{v}_j = \sigma(\rho_{2^j})$ and $\mathbf{S}_r = \sum_{j=1}^{\infty} (2^j \mathbf{v}_j)^r$, we see from (6.1), that $\mathbf{S}_r = \infty$ iff $s \geqslant r\mathbf{d}$. Hence the MultiLog Law for hitting times follows from Corollary 3.7.

The proof of other results are similar. □

6.3. **Law of iterated logarithm for hitting times.** The proof of part (b) of Theorem 6.1 relies on the following lemma on the fluctuations of local dimension of Gibbs measures of expanding circle maps. We leave the proof of the lemma to Appendix B.

Denote $\psi(x) = g(x) + \mathbf{d} f_u(x)$, then we have $\int \psi d\mu = 0$ under the assumption $P(g) = 0$.

Define $\sigma = \sigma(\mu)$ by the relation

$$\sigma^2 = \int \psi^2 d\mu + 2\sum_{n=1}^{\infty} \int f_u(x) f_u(f^n x) d\mu(x).$$

Lemma 6.6.
(a) $\sigma(\mu) = 0$ iff μ is conformal.
(b) For μ almost every x

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

We proceed to the proof of part (b) of Theorem 6.1, that uses only the lim inf of part (b) of Lemma 6.6.

Take an arbitrary ε and let

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

Take

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

Denote $\tilde{\vartheta}^{\pm}(n) = \vartheta^{\pm}(\rho_n)$, then

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

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

The lim inf in Lemma 6.6, has the following straightforward consequences, for any $\varepsilon > 0$ and for μ almost every x:

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

$$\mu\left(B(x, \rho_n)\right) \leqslant \tilde{\vartheta}^+(n)$$

(ii) For a subsequence $n_l \to \infty$ we have

$$\mu\left(B(x, \rho_{n_l})\right) \geqslant \tilde{\vartheta}^-(n_l)$$

From there it follows that for any $r \geqslant 1$, $S_r = \sum_{k=1}^{\infty} (2^k \mu B(x, \rho_{2^k}))^r$ is finite if $c > (1+\varepsilon)\frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}$ and is infinite if $c < (1-\varepsilon)\frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}$. Thus the result follows from Theorems 2.2 and 3.6 since the assumptions of Section 3 have been verified in §6.2.

6.4. **Law of iterated logarithm for return times. Lower bound.** Here we prove that

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

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 $P_n(x) = \{P_n \in \mathcal{P}_n : x \in P_n\}$ and a sequence $n_j(x)$, $j \in \mathbb{N}$ such that $n_0(x) = 0$,

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

where

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

Let

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

and denote $\mathcal{F}_j = \mathcal{B}\left(\mathcal{P}_{k_1}, \cdots, \mathcal{P}_{k_j}\right)$. Then

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

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

We will use Lévy's extension of the Borel-Cantelli Lemma.

Theorem 6.7. *([134, §12.15]) If*

$$\sum_j \mathbb{P}\left(\mathcal{A}_{j+1}|\mathcal{F}_j\right) = \infty \quad a.s.$$

then $\mathcal{A}_j$ happen infinitely many times almost surely.

Hence (6.7) 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

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

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 Proposition 6.3 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 2.8 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. □

6.5. **Law of iterated logarithm for return times. Upper bound.** Now we turn to the proof of

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(x,x)| - \frac{1}{\mathbf{d}}\ln n}{\sqrt{2(\ln n)(\ln\ln\ln n)}} \leqslant \frac{\sigma}{\mathbf{d}\sqrt{\mathbf{d}\lambda}}. \tag{6.8}$$

Since $d_n^{(r)}(x,x) \geqslant d_n^{(1)}(x,x)$, we only need to show (6.8) for $r = 1$.

Denote

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

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

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

Proposition 6.4 allows us to further restrict the range of m by assuming $m \geqslant \bar{\varepsilon}\ln N_k$ there $\bar{\varepsilon}$ 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

$$\mathcal{A}_k = \{x : x \text{ is } n-\text{good for } n \geqslant N_{k_0} \text{ but } d(f^m x, x) \leqslant r_{N_k} \text{ for } m = \bar{\varepsilon}\ln N_k, \dots N_k.\}$$

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 then for $m \leqslant K\ln N_k$ we use (6.6) telling us that

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

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

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

Summing those estimate for we obtain

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

for some $\kappa = \kappa(\varepsilon) > 0$. Since $B(x_{j,k}, r_{N_k}/2)$ are disjoint for different j we conclude using (6.5) that

$$\sum_j \mu(B(x, 2r_{N_k}) \leqslant R \sum_j \mu(B(x, r_{N_k}/2) \leqslant R.$$

It follows that

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

Now the result follows from Borel–Cantelli Lemma.

6.6. **Notes.** The fact that return times for the non-conformal Gibbs measures are dominated by fluctuations of measures of the balls has been explored in various settings [23, 24, 31, 37, 77, 84, 115, 121, 131]. In particular, [72] obtains a result similar to our Lemma 6.6 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 [13], while fluctuations are determined only for a limited class of systems [107]. 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 manifold of curvature -1 and dimension $d+1$. For $q \in \mathcal{Q}, v \in S\mathcal{Q}$ let $\gamma(t, q, v)$ be the geodesic such that $\gamma(0) = q$ $\dot{\gamma}(0) = v$. We call g_t the corresponding flow, that is $g_t(\gamma(0), \dot{\gamma}(0)) = (\gamma(t), \dot{\gamma}(t))$. Fix a reference point $O \in \mathcal{Q}$ and let $D(q, v, t) = \operatorname{dist}(\gamma(t), O)$, According to Sullivan's Logarithm Law for excursions [129] for almost all q and v

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

In fact, the argument used to prove (7.1) also shows using the Borel–Cantelli Lemma of [129] that

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

Here we present a multiple version of (7.1). Recall ([16, Proposition D.3.12]) that $\mathcal{Q}$ admits a decomposition $\mathcal{Q} = \mathcal{K} \cup \left(\cup_{j=1}^{p} \mathcal{C}_j\right)$ 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

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

where V_i is a compact flat manifold and dx is the Euclidean metric on V_j. Cusp are disjoint, so that a geodesic can not pass between different cusps without visiting the thick part $\mathcal{K}$ in between. We note that for each $(x_0, y_0) \in \mathcal{C}_i$ there is a unique geodesic $(\{x = x_0\})$ which remains in the cusp for all positive time. We will call this geodesic *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 $\gamma(q, v, t) \in \mathcal{C}_i$ and has coordinates $(x(t), y(t))$ where. 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 $\gamma(t)$ belong to some cusp $\mathcal{C}_i$ for all $t \in I$. $h(I) = \max_{t \in I} h(q, v, t)$ is called *the height of the excursion* I. Let

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

be the heights of excursions finished up to time T ordered in a decreasing order. By the foregoing discussion (7.2) can be restated as

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

We have the following extension of this result

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

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

We also have the following byproduct of our analysis

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* $S\mathcal{Q}$*. Then the number of excursions in the cusp* $\mathcal{C}_i$ *which finished before time* T *and reached the height* $\frac{\ln T}{d} + \mathfrak{h}$ *is asymptotically Poisson with parameter* $a_i e^{-d\mathfrak{h}}$.

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

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

In particular, taking $r = 1$ in (7.5) we obtain

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

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

7.2. **Multilog law for geodesic excursions. Proof of Theorem 7.1.** To prove this result we 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 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 $\sin\theta \leqslant \frac{y(n^*)}{R}$. Thus the probability that a geodesic which enters the cusp at time n^* will reach height H is sandwiched between

$$\frac{e^{-dH}}{C} \quad \text{and} \quad Ce^{-dH}$$

for some constant C.

Thus given H we consider the set $\mathcal{A}_H$ which consists of points $(q, v) \in M$ such that
(i) The first backward time such that $\gamma(-t, q, v)$ exits the cusp is between 0 and 1;
(ii) The angle v makes with the exiting geodesic is less than e^{-H}.
By the foregoing discussion

$$\frac{e^{-dH}}{C} \leqslant \mu(\mathcal{A}_H) \leqslant Ce^{-dH}. \tag{7.6}$$

In fact, it is not difficult to see that for each cusp $\mathcal{C}_i$ there is a constant a_i such that for each $n \in \mathbb{Z}$

$$\mu(x \text{ enters } \mathcal{C}_i \text{ at time } n \text{ and reaches height } H) = a_i e^{-dH}(1 + o(1)) \tag{7.7}$$

but this will be needed in the proof of Corollary 7.2 and not in the proof of Theorem 7.1.

To prove Theorem 7.1 we verify the conditions of Corollary 3.7. Let $\mathbb{B} = C^1(S\mathcal{Q})$. Then the system (g_1, M, μ) is exponentially mixing in the sense of Definition 3.1. Indeed, $(EM)_r$ follows from [100] and (Prod) and (Gr) are clear.

Next, for $\rho > 0$, we consider the following functions defined in a neighborhood of $\mathcal{A}_{-\ln\rho}$: $t(x)$-time since entering the cusp, and $\psi(x)$ the angle with escaping geodesics. We then define the following sequence of targets.

$$\Omega_\rho = \mathcal{A}_{-\ln\rho} = \left\{ x \in M : \ t(x) \in [0, 1], \ \ \psi(x) \in \left[0, \frac{1}{\rho}\right] \right\}. \tag{7.8}$$

We need to check the conditions of Definition 3.2 for these targets.

Since both t and ψ are Lipshitz (Appr) follows as in Remark 3.4.

As for (Mov), it is a direct consequence of the following.

Lemma 7.4. *There is a constant K such that for each n_1, n_2*

$$\mu\left(\mathcal{A}_H(n_1)\mathcal{A}_H(n_2)\right) \leqslant K\mu\left(\mathcal{A}_H(n_1)\right)\mu\left(\mathcal{A}_H(n_2)\right).$$

Before we prove Lemma 7.4, we finish the proof of Theorem 7.1.

We let

$$\rho_n := \frac{1}{n^{\frac{1}{d}} \ln^s n}. \tag{7.9}$$

By (7.6), we have that

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

hence (Poly) is satisfied.

Finally, we have that $(2^j \mu(\Omega_{\rho_{2^j}}))^r = \frac{1}{j^{sdr}}$, hence by Corollary 3.7 we get that if $s > \frac{1}{dr}$, then with probability one, we have that for large n, $N_n(\rho_n) < r$. This implies that for almost every $(q, v) \in M$, it holds that for n sufficiently large

$$H^{(r)}(q, v, n) < -\ln \rho_n = \frac{\ln n}{d} + s \ln \ln n.$$

Conversely, if $s \leqslant \frac{1}{dr}$, then with probability one, we have that for large n, $N_n(\rho_n) \geqslant r$. This implies that for almost every $(q, v) \in M$, it holds that for n sufficiently large

$$H^{(r)}(q, v, n) \geqslant -\ln \rho_n = \frac{\ln n}{d} + s \ln \ln n.$$

Thus (7.4) is proved.

To finish we still need to give the

Proof of Lemma 7.4. Let $\tilde{\mathcal{A}}_H = I\mathcal{A}_H$ where I denotes the involution $I(q, v) = (q, -v)$. Given n_1, n_2 define

$$\mathcal{B}_{H,n_1,n_2} = \{x : g^{n_1} x \in \mathcal{A}_H,\ g^{n_2} x \in \tilde{\mathcal{A}}_H,\ g^n x \notin \mathcal{K} \text{ for } n_1 < n < n_2\}.$$

Thus $\mathcal{B}_{H,n_1,n_2}$ consists of points which enter a cusp at time n_1, reach the height H, and then exit the cusp at time n_2. Fix a small δ and let $\tilde{\mathcal{B}}_{H,n_1,n_2} = \cup_{x \in \mathcal{B}_{H,n_1,n_2}} W^u(x, \delta e^{-n_2})$. Note that if $y \in \tilde{\mathcal{B}}_{H,n_1,n_2}$ 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 $\bar{n}_2 \in [n_2 - 1, n_2 + 1]$. In particular for each n_1 the sets $\{\tilde{\mathcal{B}}_{H,n_1,n_2}\}_{n_2 \geqslant n_1}$ have at least 3 intersection multiplicity and hence

$$\sum_{n_2} \mu(\tilde{\mathcal{B}}_{H,n_1,n_2}) \leqslant 3\mu(\mathcal{A}_{H-1}) \leqslant Ce^{-dH}. \tag{7.10}$$

Also for $\hat{n} > n_1$

$$\mu(\mathcal{A}_H(n_1)\mathcal{A}_H(\hat{n})) \leqslant \sum_{n_2} \mu(\mathcal{B}_{H,n_1,n_2}\mathcal{A}_H(\hat{n})).$$

We claim that for each $n_1, n_2, \hat{n}$

$$\mu(\mathcal{A}_H(\hat{n})\mathcal{B}_{H,n_1,n_2}) \leqslant Ce^{-dH} \mu(\tilde{\mathcal{B}}_{H,n_1,n_2}) \tag{7.11}$$

(7.11) implies that

$$\mu(\mathcal{A}_H(n_1)\mathcal{A}_H(\hat{n})) \leqslant Ce^{-dH} \sum_{n_2} \mu(\tilde{\mathcal{B}}_{H,n_1,n_2}) \leqslant \hat{C}e^{-dH} \mathbb{P}(\mathcal{A}_H)$$

where the last step relies on (7.10).

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. Let

$$\hat{\mathcal{B}}_{H,n_1,n_2} = \bigcup_{x\in B_{H,n_1,n_2}} \mathcal{W}^u_\delta(g^{n_2}x)$$

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

$$\mathcal{B}_{H,n_1,n_2} \subset g^{-n_2}\hat{\mathcal{B}}_{H,n_1,n_2} \subset \tilde{\mathcal{B}}_{H,n_1,n_2}.$$

Thus

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

where $n^* = \hat{n} - n_2$. We claim that

$$\mu(\hat{\mathcal{B}}_{H,n_1,n_2}\mathcal{A}_H(n^*)) \leqslant C\mu(g^{n_2}\tilde{\mathcal{B}}_{H,n_1,n_2})e^{-dH}. \tag{7.12}$$

Since $\mu(g^{n_2}\tilde{\mathcal{B}}_{H,n_1,n_2}) = \mu(\tilde{\mathcal{B}}_{H,n_1,n_2})$ (7.12) implies (7.11). By construction $\hat{\mathcal{B}}_{H,n_1,n_2}$ is partitioned into unstable cubes of size δ, it suffices to show that for any such cube $\mathcal{W}$ which is contained in $\mathcal{U}$ we have

$$\mu(\tilde{A}_H(n^*)|\mathcal{W}) \leqslant Ce^{-dH} \tag{7.13}$$

where $\mu(\cdot|\cdot)$ denotes the conditional expectation. Let $Q = \cup_{x\in\mathcal{W}} \cup_{|t|<\delta} W^s_\delta(g^t x)$. 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 $y \in W^s_\delta(g^t x)$. 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

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

This establishes (7.13) and, hence (7.12) completing the proof of Lemma 7.4. □

With the proof of Lemma 7.4, we finished the proof of Theorem 7.1. □

7.3. **Poisson Law for excursions. Proof of Corollary 7.2.** We fix $\mathfrak{h} \in \mathbb{R}$ and fix a cusp index i. We change the targets Ω_{ρ_n} where Ω_ρ is defined by (7.8) and ρ_n is defined by (7.9) to

$$\Omega^{(i)}_\rho = \mathcal{A}_{-\ln\rho} = \left\{x \in M : \ t_i(x) \in [0,1], \ \ \psi_i(x) \in [0, \frac{1}{\mathfrak{h}\rho}]\right\}. \tag{7.14}$$

and

$$\rho_n := n^{-\frac{1}{d}}. \tag{7.15}$$

where $t_i(x)$ is the time since entering the cusp $\mathcal{C}_i$, and $\psi_i(x)$ the angle with escaping geodesics.

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

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

Therefore Corollary 7.2 follows from Theorem 2.8. □

7.4. **Notes.** The logarithm law for the highest excursion was proven in [129]. The extensions for infinite volume hyperbolic manifolds is studied in [128]. Corollary 7.3 for surfaces is obtained in [85] where the authors also consider infinite volume surfaces. Papers [11, 50] obtain stable laws for geodesic windings on hyperbolic manifolds. Indeed 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. [46, 48]. 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, [100] studies partially hyperbolic flows on homogenous spaces and presents applications to metric number theory, cf. Section 8 of the present paper. Logarithm Law for unipotent flows is considered in [8, 9, 65, 87]. In the next section we obtain MultiLog Law for certain diagonal flows on the space of lattices.

8. Multiple Khintchine-Groshev Theorem.

8.1. Statements.

Homogenous approximations.

Definition 8.1 ((r,s)-approximable vectors)**.** *Given $\alpha = (\alpha_1, \dots \alpha_d) \in \mathbb{R}^d$, $s \geqslant 0$, $c > 0$, let $D_N(\alpha, s, c)$ be the set of $k = (k_1, \dots, k_d) \in \mathbb{Z}^d$ such that*

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

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

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

Theorem 8.2. *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 8.3. *Observe that an equivalent statement of Theorem 8.2 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 8.4 ((r,s)-approximable couples)**.** *Given $\alpha = (\alpha_1, \dots \alpha_d) \in \mathbb{R}^d$ and $z \in \mathbb{R}$, $s \geqslant 0$ and $c > 0$, let $D_N(\alpha, z, s, c)$ be the set of $k = (k_1, \dots, k_d) \in \mathbb{Z}^d$ such that*

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

*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 8.5. *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 8.6 ((r,s)-simultaneously approximable vectors)**.** *Given $\alpha = (\alpha_1, \dots \alpha_d) \in \mathbb{R}^d$, $s \geqslant 0$, $c > 0$, let $D_N(\alpha, s, c)$ be the set of $k \in \mathbb{Z}^*$ such that*

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

*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 8.7. *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 of Theorem 8.7 since it is obtained by routine modification of the proof of Theorem 8.2.

8.2. **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 $SL_{d+1}(\mathbb{R})/SL_{d+1}(\mathbb{Z})$.

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

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

$$g_t = \begin{pmatrix} 2^{-t} & & & \\ & \ddots & & \\ & & 2^{-t} & \\ & & & 2^{dt} \end{pmatrix} \tag{8.1}$$

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

$$\mathcal{S}(f)(\mathcal{L}) = \sum_{e \in \mathcal{L},\ e \text{ prime}} f(e). \tag{8.2}$$

For $a > 0$, let ϕ_a be the indicator of the set[4]

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

[4]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

$$\nu := \frac{c}{M \ln M^s}, \quad \Phi_\nu := \mathcal{S}(\phi_\nu). \tag{8.3}$$

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

$$.\exists m, \quad gcd(k_1, \ldots k_d, m) = 1, \quad |k|^d |\langle k, \alpha \rangle + m| \leqslant \frac{c}{M \ln^s M}. \tag{8.4}$$

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 8.2 we are counting all solutions we have to deal with this issue.

The following proposition shows that almost surely on α, multiple solutions do not occur. Its proof is based on Rogers identity for the second moment of the Siegel transforms.

Proposition 8.8. *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 8.2 is equivalent to the following.

Theorem 8.9. *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

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

8.3. **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 $\Lambda_\alpha : \alpha \in \mathbb{T}^d$.

Hence, we define

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

where we recall that

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

and ϕ_ν is the indicator of the set $E_\nu = \big\{(x,y) \in \mathbb{R}^d \times \mathbb{R} \mid x_1 > 0, |x| \in [1,2], |x|^d|y| \in [0,\nu]\big\}$.

Our goal becomes to prove the following.

Proposition 8.10. *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 8.11. *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 $\tilde{\mu}$-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 8.10 and Theorem 8.11 imply Proposition 8.8 and Theorem 8.9.

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 (8.7) 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

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

The idea now is that if $0 < \varepsilon \ll \eta$ is fixed, and if B is distributed according to a smooth measure with respect to Haar measure of $SL_d(\mathbb{R})$ in an ε neighborhood of the Identity, and if 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 $\mathbb{T}^d$, we have that the lattice $\tilde{\Lambda}_{\alpha,\beta,B}$ is distributed according to a smooth distribution in $\mathcal{M}$ according to the Haar measure μ.

Moreover, because Λ_β^- forms the stable direction of g_t and because U_B forms the centralizer of g_t, we have that if M is sufficiently large, then

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

hence, the implication of Proposition 8.8 and Theorem 8.9 from Proposition 8.10 and Theorem 8.11. □

8.4. **Rogers identities.** The following identities (see [110, 133]) 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 8.12. *We have*

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

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

8.5. **Proof of Proposition 8.10.** Recall that $\nu = \frac{c}{M \ln M^s}$

Lemma 8.13. *There exists a constant $C > 0$, such that for every M, for every $t \in \mathbb{R}$, it holds that*

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

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

$$\mathbb{P}_{\text{Haar}}\left(\exists t \leqslant 2^K, \Phi_{\nu_{4c}}(g_t \mathcal{L}) > 1\right) \leqslant 16 C c^2 2^{-K} \ln^{-2s}(2^K).$$

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

To finish the proof of Proposition 8.10 we give

Proof of Lemma 8.13. 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

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

It then follows from Rogers identity (b) of Lemma 8.12 that

$$\mathbb{P}_{\text{Haar}}\left(\Phi_{\nu_c}(\mathcal{L}) > 1\right) \leqslant \mathbb{E}\left(\Phi_{\nu_c}^2(\mathcal{L}) - \Phi_{\nu_c}(\mathcal{L})\right) \leqslant \mathbf{c}_2 \left(\int_{\mathbb{R}^{d+1}} \phi_{\nu_c}(u) du\right)^2 \leqslant C c^2 M^{-2} \ln^{-2s} M.$$

□

8.6. **Proof of Theorem 8.11.**

We want to apply Corollary 3.7. 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 $\Omega_\rho = \{\mathcal{L} : \Phi_\rho(\mathcal{L}) \geqslant 1\}$. 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 $\rho_M := \frac{c}{M \ln^s M}$. The conclusions of Theorem 8.11 will then follow from the conclusion of Corollary 3.7 applied to $N_M(\rho_M)$, where $N_n(\rho)$ is the number of times $t \leqslant n$ such that $\Omega_t(\rho)$ 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 $\mathbf{S}_r = \infty$ iff $rs < 1$.

This being said, to be able to apply Corollary 3.7 and finish, we still need to check the conditions of Theorem 3.6 for the system $(g_1, \mathcal{M}, \mu)$ and for the family of targets given by Ω_ρ and the sequence $\rho_M = \frac{c}{M \ln^s M}$.

Fix $s \in \mathbb{N}$, $s \geqslant 2$. We define $C_c^s(\mathcal{M})$ the space of smooth functions defined over $\mathcal{M}$ that are constant outside a compact set.

Next we let $\mathbb{B} = C_c^s(\mathcal{M})$. For this norm, it is known that the system $(g_1, \mathcal{M}, \mu)$ is exponentially mixing as in Definition 3.1 [18].

We need to check the conditions of Definition 3.2. The approximation condition (Appr) can be checked as follows.

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

(i) $\|A_\rho^\pm\|_\infty \leqslant 2$ and $\|A_\rho^\pm\|_{\mathbb{B}} \leqslant \rho^{-\sigma}$;
(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 $\mu(\Omega_\rho) = \mathcal{O}(\rho)$. The proof of the claim is similar to the corresponding argument in [47].

Next we show now how Rogers identity of Lemma 8.12(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

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

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 8.12 (b) we get that

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

which is stronger than the required (Mov).

Finally, the condition (Poly) clearly holds for the sequence $\rho_M = \frac{c}{M \ln^s M}$ due to Lemma 8.12 (a). □

8.7. **The argument in the inhomogeneous case.** The proof of Theorem 8.5 is very similar to that of Theorem 8.2, and below we only outline the main differences.

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

$$\Lambda_{\alpha,z} = (\Lambda_\alpha, (0,\dots,0,z)) \tag{8.5}$$

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

$$\tilde{\mathcal{S}}(f)(\tilde{\mathcal{L}}) = \sum_{e \in \tilde{\mathcal{L}}} f(e). \tag{8.6}$$

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[5]

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

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

$$\nu := \frac{c}{M \ln^s M}, \quad \tilde{\Phi}_\nu := \tilde{\mathcal{S}}(\phi_\nu). \tag{8.7}$$

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 $k = (k_1,\dots,k_d)$ such that $2^t < |k| \leqslant 2^{t+1}$ and that

$$\exists m, \quad |k|^d|z + \langle k,\alpha\rangle + m| \leqslant \frac{c}{M \ln^s M}. \tag{8.8}$$

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 8.2 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.

[5]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 8.2 do not appear in the inhomogeneous Diophantine approximation problem of Theorem 8.5.

Proposition 8.14. *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 8.5 is equivalent to the following.

Theorem 8.15. *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

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

8.8. **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 8.14 and Theorem 8.15 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 8.16. *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 8.17. *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

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

8.9. **Proofs of Proposition 8.16 and Theorem 8.17.** Again, the proofs of Proposition 8.16 and Theorem 8.17 are very similar to the proofs of their counterpart in the homogeneous case, Proposition 8.10 and Theorem 8.11.

Similarly to the homogeneous case, we want to apply Corollary 3.7. 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 $\Omega_\rho = \{\tilde{\mathcal{L}} : \tilde{\Phi}_\rho(\mathcal{L}) \geqslant 1\}$. 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 8.16 and Theorem 8.17 compared to that of Proposition 8.10 and Theorem 8.11, is in the application of Rogers identities to prove Proposition 8.16 as well as in the proof of (Mov) that is part of the proof of Theorem 8.17.

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 (8.6). Rogers identities for affine lattices read ([110])

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

(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 decomposes into pairs of independent vectors and pairs of equal vectors.)

Now for the proof Proposition 8.16, we have that

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

and Proposition 8.16 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

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

which is stronger than the required (Mov). □

8.10. **Multiple recurrence for toral translations.**

Proof of Theorem 4.4. 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 $\tilde{d}_n^{(r)}(z,\alpha)$ 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

$$\limsup_{n\to\infty} \frac{|\ln\tilde{d}^{(1)}(z,\alpha)|-\frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{d}, \tag{8.9}$$

$$\limsup_{n\to\infty} \frac{|\ln \tilde{d}^{(r)}(z,\alpha)| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{2d}, \text{ for } r \geqslant 2. \tag{8.10}$$

Next we claim that it suffices to prove (8.10) only for $r = 2$. Indeed, since $\tilde{d}_n^{(r)}$ is non decreasing in r, (8.10) with $r = 2$ implies that for $r > 2$

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

To get the upper bound, suppose that $d_n^{(2)}(z,\alpha) \leqslant \varepsilon$. 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

$$k_2 + (r-2)\alpha \in B(z, 1 + 2(r-2)\varepsilon).$$

Thus $d_{(r-1)n}^{(r)}(z,\alpha) \leqslant (2r-1)d_n^{(2)}(z,\alpha)$. Taking lim sup we obtain that if (8.10) holds for $r = 2$ then it holds for arbitrary r. In summary, we only need to show (8.9) and

$$\limsup_{n\to\infty} \frac{|\ln \tilde{d}^{(2)}(z,\alpha)| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{2d}. \tag{8.11}$$

The proofs of (8.9) and (8.11) is similar to but easier than the proof of Theorem 8.5 so we only explain the changes. First, it is suffices to take lim sup for n of the form 2^M since for $2^{M-1} \leqslant n \leqslant 2^M$ we have

$$d_{2^M}^{(r)} \leqslant d_n^{(r)} \leqslant d_{2^{M-1}}^{(r)}.$$

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

$$E_\nu = \{e = (e', e'') \in \mathbb{R}^d \times \mathbb{R} : ||e'|| \leqslant \nu, e'' \in [0,1)\}. \tag{8.12}$$

Then a direct inspection shows that

$$d_{2^M}^{(r)} \geqslant r \Leftrightarrow \tilde{S}(\mathbf{1}_{E_{\nu_M}})(g_M\Lambda_{\alpha,z}\mathbb{Z}^{d+1})) \geqslant r$$

where g_M is defined by (8.1) and $\Lambda_{\alpha,z}$ is defined by (8.5). As in the proof of Theorem 8.5 one can show that $\tilde{S}(\mathbf{1}_{E_{\nu_M}})(g_M\Lambda_{\alpha,z}\mathbb{Z}^{d+1}) \geqslant r$ infinitely often for almost every (z,α) iff $\tilde{S}(\mathbf{1}_{E_{\nu_M}})(g_M\tilde{\mathcal{L}})) \geqslant r$ infinitely often for almost every $\tilde{\mathcal{L}} \in \widetilde{\mathcal{M}}$. Thus we need to show that for almost every $\tilde{\mathcal{L}} \in \widetilde{\mathcal{M}}$

$$\tilde{S}(\mathbf{1}_{E_{\nu_M}})(g_M\tilde{\mathcal{L}}) \geqslant 1 \text{ infinitely often if } s < \frac{1}{d}, \tag{8.13}$$

$$\tilde{S}(\mathbf{1}_{E_{\nu_M}})(g_M\tilde{\mathcal{L}}) \geqslant 1 \text{ finitely often if } s > \frac{1}{d}, \tag{8.14}$$

$$\tilde{S}(\mathbf{1}_{E_{\nu_M}})(g_M\tilde{\mathcal{L}}) \geqslant 2 \text{ infinitely often if } s < \frac{1}{2d}, \tag{8.15}$$

$$\tilde{S}(\mathbf{1}_{E_{\nu_M}})(g_M\tilde{\mathcal{L}}) \geqslant 2 \text{ finitely often if } s > \frac{1}{2d}. \tag{8.16}$$

To prove (8.13)-(8.16), we need the following fact.

Lemma 8.18. *(a)* $\mathbb{P}(\tilde{S}(1_{E_\nu}) = 1) = c_d\nu^d(1 + o(1))$,
(b) $c'\nu^{2d} \leqslant \mathbb{P}(\tilde{S}(1_{E_\nu}) \geqslant 2) \leqslant c''\nu^{2d}$.

Before we give the proof of the lemma, we see how it allows to obtain (8.13)–(8.16) and thus finish the proof of part (a) of Theorem 4.4.

Indeed, Lemma 8.18 shows that

$$\sum_M \mathbb{P}(\tilde{S}(1_{E_{\nu_M}}) = 1) = \infty \text{ iff } s < \frac{1}{d}, \quad \sum_M \mathbb{P}(\tilde{S}(1_{E_{\nu_M}}) \geqslant 2) = \infty \text{ iff } s < \frac{1}{2d}.$$

From there, (8.13)–(8.16) follow from the the classical Borel Cantelli Lemma, that is, from the case $r = 1$ in our Theorem 2.2. We note that in case $r = 1$ Theorem 2.2 is a minor variation of standard dynamical Borel Cantelli Lemmas such as e.g., the Borel Cantelli Lemma of [100].

It remains to verify the conditions of Theorem 2.2 with $r = 1$. Thus we need to verify conditions $(M1)_1, (M2)_2$ and $(M3)_1$. This verification is very similar to the proof of Theorem 8.5 so we leave it to the readers.

Proof of Lemma 8.18. Letting $\Phi = \tilde{S}(1_{E_n u})$ we get by Rogers

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

It follows that

$$\mathbb{P}(\Phi \geqslant 2) \leqslant \frac{\mathbb{E}(\Phi^2 - \Phi)}{2} \leqslant C\nu^{2d}$$

proving the upper bound of part (b).

In addition

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

so that

$$\mathbb{P}(\Phi = 1) = \mathbb{E}(\Phi) - \mathbb{E}(\Phi \mathbf{1}_{\Phi \geqslant 2}) = c_d\nu^d + \mathcal{O}\left(\nu^{2d}\right). \tag{8.17}$$

This proves part (a).

To prove the lower bound in part (b) we need the following estimate. Let

$$\begin{aligned}
\hat{E}_1 &= \{(e', e'') \in \mathbb{R}^d \times \mathbb{R} : |e'| < \frac{\nu}{10}, e'' \in [0.1, 0.2]\},\\
\hat{E}_2 &= \{(e', e'') \in \mathbb{R}^d \times \mathbb{R} : |e'| < \frac{\nu}{10}, |e''| \leqslant 0.2\},\\
\mathcal{A} &= \{\mathcal{L} \in \mathcal{M} : \mathrm{Card}(\mathcal{L}_{prime} \cap \hat{E}_1(\nu)) = \mathrm{Card}(\mathcal{L}_{prime} \cap \hat{E}_2(\nu)) = 1\}.
\end{aligned}$$

Claim. *We have*

$$\mathbb{P}(\mathcal{A}) = c\nu^d. \tag{8.18}$$

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

$$\hat{E}_3(\nu) = \{(e', e'') \in \mathbb{R}^d \times \mathbb{R} : |e'| \leqslant 0.01\nu, |e''| \leqslant 0.01.\}$$

We thus have

$$\mathbb{P}((\mathcal{L}+z) : \mathrm{Card}((\mathcal{L}+z)\cap E_\nu) \geqslant 2)) \geqslant \mathbb{P}(\mathcal{A})\mathbb{P}(\mathrm{Card}((\mathcal{L}+z)\cap E_\nu) \geqslant 2|\mathcal{A})$$

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

This gives the lower bound in part (b) of Lemma 8.18. 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 $\Psi_j = \tilde{S}(\mathbf{1}_{\hat{E}_j})$. By Rogers

$$\mathbb{E}\left(\Psi_1\right) = 0.1c_d\nu^d, \quad \mathbb{E}\left(\Psi_1^2 - \Psi_1\right) = \left(0.1c_d\nu^d\right)^2.$$

Thus arguing as in the proof of (8.17) we conclude that

$$\mathbb{P}(\Psi_1 = 1) = 0.1c_d\nu^d + \mathcal{O}\left(\nu^{2d}\right). \tag{8.19}$$

Rogers identities also give

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

Hence

$$\mathbb{P}(\mathrm{Card}(\mathcal{L}_{prime}\cap \hat{E}_1) \geqslant 1 \text{ and } \mathrm{Card}\left(\mathcal{L}_{prime}\cap\left(\hat{E}_2\backslash\hat{E}_1\right)\right) \geqslant 1) = \mathcal{O}\left(\nu^{2d}\right). \tag{8.20}$$

Combining (8.19) and (8.20) we obtain (8.18) 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 $\mathrm{Card}(\mathcal{L}_{prime}\cap \hat{E}_2(\nu)) \leqslant 1$ since $\mathcal{L}$ is unimodular. Thus

$\mathbb{E}(\Psi_1) = \mathbb{P}(\Psi_1 = 1) = \mathbb{P}(\Psi_1 = 1 \text{ and } \Psi_2 - \Psi_1 = 0) = c\nu.$ □

This completes the proof of Lemma 8.18 and thus of part (a) of Theorem 4.4. □

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 (8.10) from (8.11), 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 $\mathcal{L} \in \mathcal{M}$

$$\mathcal{S}(\mathbf{1}_{E_{\nu_n}})(g_n\mathcal{L}) \geqslant 1 \text{ infinitely often if } s < \frac{1}{d}, \tag{8.21}$$

$$\mathcal{S}(\mathbf{1}_{E_{\nu_n}})(g_n\mathcal{L}) \geqslant 1 \text{ finitely often if } s > \frac{1}{d}, \tag{8.22}$$

where E_{ν_n} is as in (8.12), and $\mathcal{S}$ designates the Siegel transform as in (8.2). By Rogers identity, Lemma 8.12(a), we have that $\mathbb{E}\left(\mathcal{S}(\mathbf{1}_{E_{\nu_n}})(g_n\mathcal{L})\right) = cn^{-sd}$, hence (8.21) and (8.22) follow by classical Borel Cantelli Lemma (see for example the Borel Cantelli Lemma of [100]) or by the case $r = 1$ of our Theorem 2.2 .

This completes the proof of Theorem 4.4. □

8.11. **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. [12, 17]). The applications of dynamics to Diophantine approximation is based on Dani correspondence [38]. In particular, [99] 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 [110]. Surveys on applications of dynamics to metric Diophantine approximations include [15, 19, 45, 51, 52, 66, 96, 101, 111]. Limit Theorems for Siegel transforms are discussed in [7, 10, 21, 46, 47]

9. Recurrence in configuration space.

9.1. **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 manifold of curvature 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 $S\mathcal{Q}$, and μ the Liouville measure on $S\mathcal{Q}$.

Fix a small number $\bar{\rho}$. 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 9.1. *(a) For each $a \in \mathcal{Q}$ and almost every $(q,v) \in S\mathcal{Q}$,*

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(a,(q,v))| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{rd}; \tag{9.1}$$

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

$$\limsup_{n\to\infty} \frac{|\ln d_n^{(r)}(q,(q,v))| - \frac{1}{d}\ln n}{\ln\ln n} = \frac{1}{rd}. \tag{9.2}$$

Note that in contrast with Section 4 there are no exceptional points for hitting. We also obtain Poisson limit theorem. Denote $B(a,\rho) = \{q \in \mathcal{Q} : d(a,q) < \rho\}$,

$$\hat{B}(a,\rho) = \{(q,v) \in S\mathcal{Q} : d(a,q) < \rho, v \in S_q\mathcal{Q}\},$$

and

$$\hat{A}(a,\rho) = \cup_{t\in[0,\varepsilon]}\phi^t \hat{B}(a,\rho),$$

for fixed small ε.

Let

$$\gamma = \frac{1}{\varepsilon}\lim_{\rho\to 0}\frac{\mu(\hat{A}(a,\rho))}{\rho^d}. \tag{9.3}$$

We will show in §9.3 that this limit exists and does not depend on $a \in \mathcal{Q}$.

Theorem 9.2. *For each a and almost every (q,v) if $N_{a,\rho,\tau}(q,v)$ is the maximal r such that $d^{(r)}_{\tau\rho^{-d}}(a,(q,v)) < \rho$ then*

$$\lim_{\rho\to 0}\mu\left((q,v) \in S\mathcal{Q} : N_{a,\rho,\tau}(q,v) = r\right) = e^{-\tau\gamma}\,\frac{(\gamma\tau)^r}{r!} \tag{9.4}$$

and

$$\lim_{\rho\to 0}\mu\left((q,v)\in S\mathcal{Q}: N_{q,\rho,\tau}(q,v)=r\right)=e^{-\tau\gamma}\,\frac{(\gamma\tau)^r}{r!} \tag{9.5}$$

9.2. **MultiLog Law.**

Proof of Theorem 9.1. We shall check the conditions Corollary 3.7 for $f=\phi^\varepsilon$, $\mathbb{B}$ the space of Lipshitz functions, and the sets $\hat{A}(q,r)$. Then (Prod) and (Gr) are clear. $(EM)_2$ is proven in [108] and $(EM)_r$ for $r>2$ follows from $(EM)_2$ due to [42]. Note that $\hat{A}(a,\rho)$ is a sublevel set of a Lipshitz function

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

so (Appr) follows as in Remark 3.4. It remains to verify (Mov). That is, we need to prove the following Proposition.

Proposition 9.3. *There exists $\eta>0$ and $t_0>0$ such that for any $a\in Q$ and ρ sufficiently small, t*

$$\mu(\hat{A}(a,\rho)\cap\phi^t\hat{A}(a,\rho))\leqslant\mu(\hat{A}(a,\rho))^{1+\eta}, \tag{9.6}$$

for all $t>t_0$.

Denote $S(q,t)=\phi^t S_q\mathcal{Q}$ and let $\mathcal{A}(q,\varepsilon)=\bigcup_{t\in[0,\varepsilon]} S(q,t)$. We shall show that for each $\bar{q}\in\mathcal{Q}$

$$\mathrm{mes}(\mathcal{A}(q,\varepsilon)\cap\phi^{-t}\hat{A}(a,\rho))\leqslant C\rho^\eta \tag{9.7}$$

which implies (9.6) by integration over $q\in B(a,\rho)$.

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

$$(J^i)''(t)+\sum_{j=0}^{n-1}K^i_j(t)J^j(t)=0, i=1,\dots,n-1,$$

where the matrix K^i_j is symmetric and its spectrum lies between $-K_1^2$ and $-K_2^2$ when $\mathcal{Q}$ has negative curvature.

Note that $J(0)=0$, $\|J'(0)\|\leqslant C_0$ for some $C_0>0$, we have the following Proposition.

Proposition 9.4. *There are constants $C>0$, $t_0>0$ such that $\|J'(t)\|\leqslant C\|J(t)\|$ for $t>t_0$.*

Proof. Suppose $J'(t)=V(t)J(t)$, $S(t)=\langle J(t),J'(t)\rangle$ and $n(t)=\|J'(t)\|^2$.

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

$$\geqslant C_1\left(\|J'(t)\|^2+\|J(t)\|^2\right)\geqslant\frac{C_1}{2}S(t)$$

for some $C_1 > 0$. It follows that $S(t) > 0$, $t > 0$, and $S(t) > S(t_0)e^{\frac{C_1}{2}(t-t_0)}$, $t > t_0$.

Besides,

$$\frac{d}{dt}n(t) = 2|\langle J''(t), J'(t)\rangle| = 2|\langle -K(t)J(t), J'(t)\rangle| \leqslant C_2\frac{d}{dt}S(t),$$

for some $C_2 > 0$, which gives that $n(t) - C \leqslant C_2 S(t)$ and $S(t) \geqslant \frac{2}{C_2}n(t), t > t_0$ for some $t_0 > 0$.

Therefore,

$$\begin{aligned} \|J'(t)\|\|J(t)\| &\geqslant \langle\, J(t), J'(t)\rangle \geqslant \frac{2}{C_2}\|J'(t)\|^2, \\ \|J'(t)\| &\leqslant C\|J(t)\| \end{aligned}$$

for some $C > 0$. □

Let us now introduce

$$\tilde{\mathcal{A}}(q,\varepsilon) = \cup_{t\in[-\varepsilon,\varepsilon]}\phi^t S_q\mathcal{Q}$$

By elementary geometry

$$\mu(\mathcal{A}(q,\varepsilon)\cap\phi^t\hat{A}(a,\rho)) \leqslant \frac{C}{\rho}\mu(\phi^t\tilde{\mathcal{A}}(q,\varepsilon)\cap B(a,2\rho)).$$

To estimate the measure of the last intersection we note that $\Sigma(t,q,\varepsilon) := \phi^t\tilde{\mathcal{A}}(q,\varepsilon)$ is an embedded submanifold on $S\mathcal{Q}$ of dimension d. By Proposition 9.4, $\pi : \Sigma(t,q,\varepsilon) \to \mathcal{Q}$ is a local diffeomorphism and moreover $||d\pi^{-1}||$ is uniformly bounded as a map $T\mathcal{Q} \to T\Sigma(t,q,\varepsilon)$. Since $\Sigma(t,q,\varepsilon)$ approaches the weak unstable manifold of ϕ as $t \to \infty$, it has uniformly bounded curvature. Hence, there exist uniform constants $\varepsilon_1, \varepsilon_2$ independent of t such that $\Sigma(t,q,\varepsilon)$ can be cut into disjoint piece $\Sigma_j(t)$ satisfying that for each j, $\pi\Sigma_j(t)$ is contained in a ball of radius ε_2 and contains a ball of radius ε_1. Decreasing ε_2 if necessary we obtain that the intersection $\pi\Sigma_j(t)\cap B(a,\rho)$ has only one component and hence

$$\mu(\pi\Sigma_j(t)\cap B(a,\rho)) \leqslant C(\varepsilon_1)\rho^d.$$

Since $d\pi^{-1}$ is bounded we also have

$$\mu(\Sigma_j(t)\cap\hat{B}(a,\rho)) \leqslant C(\varepsilon_1)\rho^d.$$

By bounded distortion

$$\mu(\phi^{-t}(\Sigma_j(t)\cap\hat{B}(a,\rho))) \leqslant C\rho^d\mu(\phi^{-t}\Sigma_j(t)).$$

Summing over j we obtain (9.7) completing the proof of Theorem 9.1(a). □

In order to complete the proof of Theorem 9.1(b) it remains to verify the conditions $\overline{\text{(Appr)}}$, $\overline{\text{(Mov)}}$ and $(\overline{\text{Sub}})$.

Denote

$$\hat{\Phi}(a,(q,v)) = \min_{s\in[0,\varepsilon]} d\left(a,\pi(\phi^s(q,v))\right)\gamma(a)^{1/d}/c,$$

We regard $\hat{\Phi}$ is a function of $x = (a,u)$ and $y = (q,v)$ on $M\times M$ even though $\hat{\Phi}$ not depend on u. For $b > 1$ let $\hat{\phi}^{\pm}(x,y)$ be Lipschitz functions such that

$$1_{B(0,\hat{\rho}-\hat{\rho}^b)} \leqslant \hat{\phi}^- \leqslant 1_{B(0,\hat{\rho})} \leqslant \hat{\phi}^+ \leqslant 1_{B(0,\hat{\rho}+\hat{\rho}^b)}.$$

Take $A^{\pm}_{\hat\rho} = \hat\phi^{\pm}(\hat\Phi)$, then

$$(\hat\rho - \hat\rho^b)^d \leqslant \int 1_{B(0,\hat\rho-\hat\rho^b)}(\hat\Phi(x,y))d\mu(y) \leqslant \int A^-_{\hat\rho}(x,y)d\mu(y)$$

$$\leqslant \int A^+_{\hat\rho}(x,y)d\mu(y) \leqslant \int 1_{B(0,\hat\rho+\hat\rho^b)}(\hat\Phi(x,y))d\mu(y) \leqslant (\hat\rho+\hat\rho^b)^d.$$

This proves $\overline{\text{(Appr)}}$(i−iii).

Denoting $C_0 = \sup_{x\in M} \gamma(x)/\gamma(y)$, we get

$$\int A^+_{\hat\rho}(x,y)d\mu(x) \leqslant \int 1_{B(0,\hat\rho+\hat\rho^b)}(\Phi(x,y))d\mu(x) \leqslant C_0 \int 1_{B(0,\hat\rho+\hat\rho^b)}(\Phi(x,y))d\mu(y) \leqslant C_0(\hat\rho+\hat\rho^b)^d,$$

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

Next, we verify $\overline{\text{(Mov)}}$. Denote

$$\bar\Omega_\rho = \left\{((a,u),(q,v)) \in S\mathcal{Q}\times S\mathcal{Q} : \exists\, s\in[0,\varepsilon], d\left(a,\pi(\phi^s(q,v))\right) < \frac{c\rho}{\gamma(a)}\right\},$$

and

$$\bar\Omega^t_\rho = \left\{(q,v)\in S\mathcal{Q} : \exists\, s\in[0,\varepsilon], d\left(q,\pi(\phi^{s+t}(q,v))\right) < \frac{c\rho}{\gamma(q)}\right\}.$$

Take $x_i\in\mathcal{Q}$, $1\leqslant i\leqslant k$ such that $\bigcup_{i=1}^k B(x_i,\rho)$ covers $\mathcal{Q}$. By (9.6),

$$\begin{aligned}
\mu(\bar\Omega^t_\rho) &\leqslant \sum_i \left\{(q,v)\in S\mathcal{Q} : \exists\, s\in[0,\varepsilon], d\left(q,\pi(\phi^{s+t}(q,v))\right) < \frac{c\rho}{\gamma(q)}, q\in B_i\right\}\\
&\leqslant \sum_i \left\{(q,v)\in S\mathcal{Q} : \exists\, s\in[0,\varepsilon], d\left(x_i,\pi(\phi^{s+t}(q,v))\right) < c_0\rho, q\in B_i\right\}\\
&\leqslant \sum_i C\rho^{d(1+\eta)} \leqslant C\rho^\eta.
\end{aligned}$$

This verifies $\overline{\text{(Mov)}}$.

It remains to verify $(\overline{\text{Sub}})$. Denote $M_0 = \max_{t\in[-\varepsilon,\varepsilon]}\|D\phi_t\|$,

$$\tilde\Omega_\rho = \left\{((a,u),(q,v)) \in S\mathcal{Q}\times S\mathcal{Q} : \exists\, s\in[-\varepsilon,\varepsilon], d\left(a,\pi(\phi^s(q,v))\right) < \frac{c\rho}{\gamma(a)}\right\},$$

and

$$\tilde\Omega^t_\rho = \left\{(q,v)\in S\mathcal{Q} : \exists\, s\in[-\varepsilon,\varepsilon], d\left(q,\pi(\phi^{t+s}(q,v))\right) < \frac{c\rho}{\gamma(q)}\right\}.$$

Note that when $(q,v)\in\bar\Omega^{t_1}_\rho\cap\bar\Omega^{t_2}_\rho$ for $t_1<t_2$, we have some s_1, $s_2\in[0,1]$ such that

$$d\left(q,\pi(\phi^{t_1+s_1}(q,v))\right) \leqslant \frac{c\rho}{(\gamma(q))^{1/d}}, \quad d\left(q,\pi(\phi^{t_2+s_2}(q,v))\right) \leqslant \frac{c\rho}{(\gamma(q))^{1/d}}.$$

Hence

$$\begin{aligned} d\left(\pi(\phi^{t_1}(q,v)), \pi(\phi^{t_2+s_2-s_1}(q,v))\right) &\leqslant M_0 d\left(\pi(\phi^{t_1+s_1}(q,v)), \pi(\phi^{t_2+s_2}(q,v))\right) \\ &\leqslant \frac{2M_0 c\rho}{(\gamma(q))^{1/d}} \leqslant \frac{c_1\rho}{(\gamma(\pi(\phi^{t_1}(q,v))))^{1/d}}. \end{aligned}$$

It follows that $\bar{\Omega}_\rho^{t_1} \cap \bar{\Omega}_\rho^{t_2} \in \phi^{-t_1}\tilde{\Omega}_\rho^{t_2-t_1}$, which complete the proof of Theorem 9.1(b) by a similar argument to Proposition 3.8 (ii).

9.3. **Poisson regime.**

Proof of Theorem 9.2. (9.4) follows from Theorem 2.8, since conditions $(M1)_r$ and $(M2)_r$ are satisfied for all r, due to the results of §9.2.

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 (9.3) does not depend on a. It remains to verify that the limit in (9.3) indeed does not depend on a base point.

If $(q,v) \in \hat{B}(a,\rho)$, denote

$$L(q,v) = L^+(q,v) + L^-(q,v) \text{ where } L^\pm(q,v) = \sup\{t : \phi^{\pm s}(q,v) \in \hat{B}(a,\rho) \text{ for } 0 \leqslant s \leqslant t\}.$$

Then we have the following estimate

$$\mu\left(\hat{A}(a,\rho)\right) = \varepsilon\left(\int_{\hat{B}(a,\rho)} \frac{1}{L(q,v)} d\mu\right)(1 + O(\rho) \tag{9.8}$$

(see e.g. [33]). Note that μ is of the form $d\mu(q,v) = \frac{d\lambda(q)d\sigma(v)}{\lambda(\mathcal{Q})}$ where λ is the Riemann volume on $\mathcal{Q}$ and σ is normalized volume on the d dimensional sphere. If ρ is small then the integral in parenthesis equals to $\rho^d\gamma(1 + O(\rho))$ where

$$\gamma = \frac{1}{\lambda(\mathcal{Q})} \int_{\mathcal{B}\times\mathbb{S}^d} \frac{1}{\mathcal{L}(x,v)} dx d\sigma(v) \tag{9.9}$$

where $\mathcal{B}$ is the unit ball in $\mathbb{R}^{d+1}$ and $\mathcal{L}(\cdot)$ is defined similarly $L(\cdot)$ with geodesics in $\mathcal{Q}$ replaced by geodesics in $\mathbb{R}^{d+1}$. Specifically, an elementary plane geometry gives $\mathcal{L}(x,v) = \sqrt{1 - r_{min}^2}$ 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 (9.3) with γ given by (9.9). □

9.4. **Notes.** In [113], Maucourant proved that for all $p \in \mathcal{Q}$ and almost every $(q,v) \in S\mathcal{Q}$

$$\limsup_{t\to+\infty} \frac{|\ln d\left(p, \pi\left(\phi^t(q,v)\right)\right)|}{\log t} = \frac{1}{d}. \tag{9.10}$$

[103] 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, 87, 88].

Concerning Poisson Limits we note that visits to sets with complicated geometry naturally appears in Extreme Value Theory, see Section 10 for details. [135] provided

a general conditions for the number of visits to a small neighborhood of arbitrary submanifold to be asymptotically Poisson.

10. Extreme values.

10.1. **From hitting balls 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 the definitions of the sets $\mathcal{G}_r$ and $\mathcal{H}$ preceding the statement of Theorem 4.2, that satisfy by Theorems 4.2 and 4.3 : $\mu(\mathcal{G}_r) = 1$ and, if the periodic points are dense, $\mathcal{H}$ 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 $\{\phi(f^j y)\}_{j=0}^n$.

Theorem 10.1. *(a) Suppose f is $2r$-fold exponentially mixing. 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 $y \in M$

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

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

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

(b) If f is an expanding map of $\mathbb{T}$ and μ is a non-conformal Gibbs measure of dimension $\mathbf{d}$ then there is a set $\mathcal{G}_\mu$ with $\mu(\mathcal{G}) = 1$ and such that if ϕ is a function with a unique non degenerate minimum at $x \in \mathcal{G}$ then for μ–almost every $y \in M$

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

(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 $\mathcal{Q}$.

(d) For toral translations we have that for almost all α and almost all y we have

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

Proof. Since x is non degenerate we have that for

$$\frac{d^2(y,x)}{K} \leqslant \phi(y) - \phi(x) \leqslant K d^2(y,x) \tag{10.1}$$

so (a) holds for $x \in \mathcal{G}_r$ and (b) holds for $x \in \mathcal{H}$. Hence, the part (a) follows from Theorems 4.2 and 4.3, part (b) follows from Theorem 6.1, part (c) follows from Theorem 4.4, and part (d) follows from Theorem 9.1. □

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*

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

with $n=\tau\rho^{-d}$ converges as $\rho\to 0$ to the Poisson process on $\mathbb{R}^+$ with intensity $\gamma(\phi)\tau\frac{2t^{(d/2)-1}}{d}$.

Proof. (10.1) does not provide enough information to deduce the result from (5.1) of Theorem 5.1 however it is not difficult to verify directly conditions $\widetilde{(M1)}_r$ and $(M2)_r$ from §2.5 for targets

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

using the results of Section 3. Since (Mov) for targets (10.2) follows from (Mov) for balls, only (Appr) needs to be checked but it follows immediately from Remark 3.4. □

Next, we consider functions of the form

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

Theorem 10.3. *Let f be $2r$-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

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

converges as $\rho\to 0$ to the Poisson process on $\mathbb{R}^+$ with intensity

$$\frac{1}{\frac{d}{s}t^{(d/s)+1}}.$$

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 $r-$ 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 $t > 0$*

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

(a) (Weibull Law for unbounded functions) *If f is $r-$ fold exponentially mixing, ϕ is given by (10.3) with $x \in \mathcal{G}$ then there is $\sigma = \sigma(x)$ such that for each $t > 0$*

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

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 [105] 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 [56]. The book [109] 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, 25, 39, 89, 90, 114] and references wherein.

Appendix A. Multiple exponentially mixing

A.1. **Basic properties.** Let f be a smooth map of a compact manifold M preserving a smooth measure μ. In the dynamical system literature f is called r fold exponentially mixing if there are constant s, C and $\bar\theta < 1$ such that for any C^s functions $A_1, \dots A_r$ for any r tuple $k_1 < k_2 < \dots < k_r$

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

where $m = \min_j (k_j - k_{j-1})$ with $k_j = 0$.

In this paper we need to consider a larger class of functions, namely we need that there are constants s, C and $\theta < 1$ such that for any $B \in C^s(M^{r+1})$ we have

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

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

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

Then

$$\int B(x_0, f^{k_1}x_0, \cdots, f^{k_r}x_0)d\mu(x_0) = \int \bar{B}(x_0, f^{k_1}x_0, \cdots, f^{k_r}x_0)d\mu(x_0) \pm e^{-\varepsilon\alpha m}$$
$$= \int \bar{B}(x_0, x_1, \dots x_r)d\mu(x_0)d\mu(x_1)\dots d\mu(x_r) + O\left(e^{-\varepsilon\alpha m}\right) + O\left(\theta^m e^{\varepsilon s m}\right)$$
$$= \int B(x_0, x_1, \dots x_r)d\mu(x_0)d\mu(x_1)\dots d\mu(x_r) + O\left(e^{-\varepsilon\alpha m}\right) + O\left(\theta^m e^{\varepsilon s m}\right).$$

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 $H^s(M^{r+1})$. 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

$$\phi_\lambda(x_0, x_1, \dots, x_r) = \prod_{j=0}^{r+1} \psi_j(x_j)$$

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 $H^s(M) \subset C^{s-\frac{d}{2}-1-\varepsilon}(M)$. 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

$$\left|\int \phi_\lambda(x, f^{k_1}x, \dots, f^{k_r}x)d\mu(x)\right| \leqslant C\lambda^{u(r+1)}\theta^m.$$

Therefore

$$\left|\int B(x, f^{k_1}x, \dots, f^{k_r}x)d\mu(x) - \int B(x_0, \cdots, x_r)d\mu(x_0)\cdots d\mu(x_r)\right|$$
$$\leqslant C\theta^m \sum_\lambda b_\lambda \lambda^{u(r+1)} \leqslant C\theta^m ||B||_{H^{u(r+1)}(M^{r+1})}.$$

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

A.2. **Mixing for Gibbs measures.**

Proof of Proposition 6.2. Denote by $\mathrm{Lip}(\mathbb{T})$ the set of Lipschitz functions on $\mathbb{T}$.

The proof consists of three steps.

Step 1. By the same argument as in [132, Proposition 3.8], we have that for $\hat{\psi}_1 \in \mathrm{Lip}(\mathbb{T}, \mathrm{a})$, $\hat{\psi}_2 \in L^1(\mu)$,

$$\left| \int \hat{\psi}_1 (\hat{\psi}_2 \circ f^n) d\mu - \int \hat{\psi}_1 d\mu \int \hat{\psi}_2 d\mu \right| \leqslant C \|\hat{\psi}_1\|_{Lip} \|\hat{\psi}_2\|_{L^1} \bar{\theta}^n, \; n \geqslant 0. \tag{A.3}$$

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

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

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 $\hat{\psi}_1 = \psi_1$, $\hat{\psi}_2 = \prod_{j=2}^{r} \psi_j (f^{k_j - k_2} x)$ we get

$$\left| \int \left(\prod_{i=1}^{r} \psi_i (f^{k_i} x) \right) d\mu(x) - \left(\int \psi_1(x) d\mu(x) \right) \left[\int \left(\prod_{i=2}^{r} \psi_i (f^{k_i} x) \right) d\mu(x) \right] \right|$$
$$\leqslant C \bar{\theta}^m \|\psi_1\|_{Lip} \left\| \left(\prod_{i=2}^{r} \psi_i (f^{k_i} x) \right) \right\|_{L_1} \leqslant C \bar{\theta}^m \prod_{i=1}^{r} ||\psi_i||_{Lip}.$$

Applying inductive estimate to

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

we obtain (A.4).

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

A.3. **Examples of exponentially mixing systems.** There are many results about double (=twofold) 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. ([42, Theorem 2]) Suppose that for each subset D in a single unstable leave of bounded geometry[6] and any Hölder probability density ρ on D we have

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

then f is r fold exponentially mixing for all $r \in \mathbb{Z}$.

[6]We refer the reader to [42] for precise requirements on D since those requirements are not essential for the present discussion.

Examples of maps satisfying the conditions of Theorem A.3 include expanding maps, volume preserving Anosov diffeomorphsims [22, 117], time one maps of contact Anosov flows [108], mostly contracting systems [26, 41], partially hyperbolic translations on homogeneous spaces [98], and partially hyperbolic automorphisms of nilmanifolds [67].

We also not 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 [43] for more details.

Another source of exponential mixing is spectral gap for transfer operators (cf. §A.2 as well as [117, 132]). This allows to handle non-uniformly hyperbolic systems admitting Young tower with exponential tails [136] as well as piecewise expanding maps [132].

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. Fluctuations of local dimension. Proof of Lemma 6.6.

We start with the proof of (b). Denote $\mu_n = \mu(B_n(x,\varepsilon))$. Since $P(g) = 0$ we have

$$\ln \mu_n = \sum_{j=0}^{n-1} g(f^j x) + O(1). \tag{B.1}$$

Denote

$$r_n = \sup_{r>0}\{r \mid B(x,r) \subset B_n(x,\varepsilon)\}, \quad \bar{r}_n = \inf_{r>0}\{r \mid B(x,r) \supset B_n(x,\varepsilon)\}.$$

By bounded distortion property, there exists $C_0 > 0$ such that if $d(f^n y, f^n x) < \varepsilon$ then

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

Thus

$$e^{-\sum_{j=0}^{n-1}\varphi(f^j x)} (C_0 \exp \varepsilon^\alpha)^{-1} d(x,y) \leqslant d(f^n x, f^n y) \leqslant e^{-\sum_{j=0}^{n-1}\varphi(f^j x)} (C_0 \exp \varepsilon^\alpha) d(x,y).$$

Hence

$$\varepsilon_0 C_0^{-1} e^{-\sum_{j=0}^{n-1}\varphi(f^j x)-\varepsilon^\alpha} \leqslant r_n \leqslant \bar{r}_n \leqslant \varepsilon_0 C_0^{-1} e^{-\sum_{j=0}^{n-1}\varphi(f^j x)+\varepsilon^\alpha}.$$

It follows that

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

Accordingly $\ln \mu_n - \mathbf{d} \ln r_n = \ln \mu_n - \mathbf{d} \ln \bar{r}_n = \sum_{j=0}^{n-1} \psi(f^j x) + O(1)$. By Law of Iterated Logarithm [83],

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

Since $B(x,r_n)\subset B_n(x,\varepsilon)\subset B(x,\bar r_n)$

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

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

$$\sigma\leqslant\limsup_{\delta\to0}\frac{|\ln\mu B(x,r_n)|-\mathbf{d}|\ln r_n|}{\sqrt{2n(\delta)\ln\ln n(\delta)}}\leqslant\limsup_{\delta\to0}\frac{|\ln\mu B(x,\delta)|-\mathbf{d}|\ln\delta|}{\sqrt{2n(\delta)\ln\ln n(\delta)}}$$

$$\leqslant\limsup_{\delta\to0}\frac{|\ln\mu B(x,\bar r_n)|-\mathbf{d}|\ln\bar r_n|}{\sqrt{2n(\delta)\ln\ln n(\delta)}}\leqslant\sigma.$$

It follows that all inequalities above are in fact equalities. In particular,

$$\limsup_{\delta\to0}\frac{|\ln\mu B(x,\delta)|-\mathbf{d}|\ln\delta|}{\sqrt{2n(\delta)\ln\ln n(\delta)}}=\sigma.$$

On the other hand by (B.2) and the ergodic theorem we see that $\lim\limits_{n\to\infty}\dfrac{|\ln r_n|}{n}=\lambda$. Therefore $\lim\limits_{n\to\infty}\dfrac{|\ln r_n|(\ln\ln|\ln r_n|)}{n\ln\ln n}=\lambda$. Since $r_n/C\leqslant\delta\leqslant r_n$ we have

$$\lim_{\delta\to0}\sqrt{\frac{n(\delta)\ln\ln n(\delta)}{|\ln\delta|(\ln\ln|\ln\delta|)}}=\frac{1}{\sqrt{\lambda}}.$$

Multiplying the last two displays we obtain

$$\limsup_{\delta\to0}\frac{|\ln\mu B(x,\delta)|-\mathbf{d}|\ln\delta|}{\sqrt{2|\ln\delta|(|\ln\ln|\ln\delta|)}}=\frac{\sigma}{\sqrt{\lambda}}.$$

Likewise

$$\liminf_{\delta\to0}\frac{|\ln\mu B(x,\delta)|-\mathbf{d}|\ln\delta|}{\sqrt{2|\ln\delta|(|\ln\ln|\ln\delta|)}}=-\frac{\sigma}{\sqrt{\lambda}}.$$

This proves part (b).

To prove part (a) suppose that $\sigma^2=0$. Since we also have that $\int\psi d\mu=0$ [117, 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 $\sum\limits_{k=0}^{n-1}\psi(f^kx)=\eta(x)-\eta(f^nx)$ is uniformly bounded with respect to both n and x. Recalling the definition of ψ we see that in this case

$$\sum_{k=0}^{n-1}g(f^kx)=\left[\mathbf{d}\sum_{k=0}^{n-1}-f_u(f^kx)\right]+O(1).$$

Now (B.1) and (B.2) show that μ is conformal. □

References

[1] Aaronson, J., Nakada, H.: Trimmed sums for non-negative, mixing stationary processes. *Stochastic Process. Appl.* **104** (2003) 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** (1989) Springer, New York, 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 Poincare Probab. Stat.* **53** (2017) 1371–1401.

[6] Athreya, J. S.: Logarithm laws and shrinking target properties. *Proc. Indian Acad. Sci. Math. Sci.* **119** (2009) 541–557.

[7] 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.

[8] Athreya, J. S., Margulis, G. A.: Logarithm laws for unipotent flows-I. *J. Mod. Dyn.* **3** (2009) 359–378.

[9] Athreya, J. S., Margulis, G. A.: Logarithm laws for unipotent flows-II. *J. Mod. Dyn.* **11** (2017) 1–16.

[10] 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.

[11] Babillot, M., Peigne, M.: Asymptotic laws for geodesic homology on hyperbolic manifolds with cusps. *Bull. SMF* **134** (2006) 119–163.

[12] Badziahin D., Beresnevich V., Velani S. Inhomogeneous theory of dual Diophantine approximation on manifolds. *Adv. Math.* 232 (2013), 1–35.

[13] Barreira L., Pesin Y., Schmeling J.: Dimension and product structure of hyperbolic measures. *Ann. of Math.* **149** (1999) no. 3, 755–783.

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

[15] Bekka M. B., Mayer M. Ergodic theory and topological dynamics of group actions on homogeneous spaces. *LMS Lecture Notes* **269** (2000) x+200 pp.

[16] Benedetti, R., Petronio, C.: Lectures on hyperbolic geometry. Universitext. Springer, Berlin, 1992. xiv+330 pp.

[17] Bernik V. I., Dodson M. M.: Metric Diophantine approximation on manifolds. *Cambridge Tracts in Mathematics* **137** (1999) xii+172 pp.

[18] Björklund, M., Einsiedler M, Gorodnik, A.: Quantitative multiple mixing, arXiv:1701.00945v2.

[19] Bjourklund M., Gorodnik A.: Central limit theorems in the geometry of numbers. *Electron. Res. Announc.* **24** (2017) 110–122.

[20] Bjourklund M., Gorodnik A.: Central Limit Theorems for group actions which are exponentially mixing of all orders, to appear in *Journal d'Analyse Mathematiques.*

[21] Bjourklund M., Gorodnik A.: Central limit theorems for Diophantine approximants. *Math. Ann.* **374** (2019), no. 3-4 1371–1437.

[22] Bowen R. Equilibrium states and the ergodic theory of Anosov diffeomorphisms. 2d revised edition. *Springer Lecture Notes in Math.* **470** (2008) viii+75 pp.

[23] 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.

[24] Bruin H., Vaienti S. Return time statistics for unimodal maps. *Fund. Math.* **176** (2003) 77–94.

[25] 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.

[26] Castro, A.: Fast mixing for attractors with a mostly contracting central direction. *Ergodic Theory Dynam. Sys.* **24** (2004) 17–44.

[27] 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.

[28] Chaika, J.: Shrinking targets for IETs: extending a theorem of Kurzweil. *GAFA* **21** (2011) 1020–1042.

[29] Chaika, J., Constantine, D.: Quantitative shrinking target properties for rotations and interval exchanges. *Israel J. Math.* **230** (2019) 275–334.

[30] Chazottes, J.-R., Collet P.: Poisson approximation for the number of visits to balls in non-uniformly hyperbolic dynamical systems. *Ergodic Th. Dynam. Sys.* **33** (2013) no. 1, 49–80.

[31] 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.

[32] Chernov, N. I.: Limit theorems and Markov approximations for chaotic dynamical systems, *Probab. Theory Related Fields,* **101** (1995) 321–362.

[33] Chernov, N.: Entropy, Lyapunov exponents, and mean free path for billiards. *J. Statist. Phys.* **88** (1997) 1–29.

[34] Chernov, N., Kleinbock, D.: *Dynamical Borel-Cantelli lemmas for Gibbs measures,* Israel J. Math. **122** (2001) 1–27.

[35] Coelho Z.: Asymptotic laws for symbolic dynamical processes. *LMS Lecture Notes 279* **2000** 123–165.

[36] Collet P.: Statistics of closest return for some non-uniformly hyperbolic systems. *Ergodic Th. Dynam. Sys.* **21** (2001) 401–420.

[37] Collet P., Galves A., Schmitt B.: Repetition times for Gibbsian sources. *Nonlinearity* **12** (1999), no. 4, 1225–1237.

[38] Dani S. G.Divergent trajectories of flows on homogeneous spaces and Diophantine approximation. *J. Reine Angew. Math.* **359** (1985) 55–89.

[39] Diamond H. G., Vaaler J. D. Estimates for partial sums of continued fraction partial quotients. *Pacific J. Math.* **122** (1986) 73–82.

[40] Doeblin W.: Remarques sur la thorie mtrique des fractions continues. *Compositio Math.* **7** (1940) 353–371.

[41] Dolgopyat, D.: On dynamics of mostly contracting diffeomorphisms. *Comm. Math. Phys.* **213** (2000) 181–201.

[42] Dolgopyat, D.: Limit theorems for partially hyperbolic systems. *Transactions AMS* **356** (2004), 1637–1689.

[43] Dolgopyat D., Dong C., Kanigowski A., Nandori P.: *Mixing properties of generalized T, T^{-1} transformations,* in preparation.

[44] Dolgopyat D., Fayad B.: Deviations of ergodic sums for toral translations I: Convex bodies. *GAFA* **24** (2014) 85–115.

[45] Dolgopyat D., Fayad B.: Limit theorems for toral translations. *AMS Proc. Sympos. Pure Math.* **89** (2015) 227–277.

[46] Dolgopyat, D., Fayad, B.: Deviations of Ergodic sums for Toral Translations II. Squares. arXiv:1211.4323v1.

[47] Dolgopyat, D., Fayad, B., Vinogradov, I.: Central limit theorems for simultaneous Diophantine approximations. *Journal de l'École Polytechnique* **4** (2017) 1–35.

[48] Dolgopyat, D., Sarig, O.: Quenched and annealed temporal limit theorems for circle rotations, to appear in *Asterisque.*

[49] 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.

[50] 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) 377–401.

[51] 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.
[52] Eskin A.: Counting problems and semisimple groups. *Doc. Math.* Extra Vol. **II** (1998) 539–552.
[53] Fayad, B.: Mixing in the absence of the shrinking target property. *Bull. LMS* **38** (2006) 829–838.
[54] Fernandez, J. L., Melian, M. V., Pestana, D.: Expanding maps, shrinking targets and hitting times. *Nonlinearity* **25** (2012) 2443–2471.
[55] Freedman, D.: Another note on the Borel-Cantelli Lemma and the strong law with the Poisson approximation as a by-product. *Ann. Probab.* **1** (1973) 910–925.
[56] Freitas A., Freitas J., Magalhaes M.: Convergence of marked point processes of excesses for dynamical systems. *JEMS* **20** (2018), no. 9, 2131–2179.
[57] Freitas A., Freitas J., Magalhaes M.: Complete convergence and records for dynamically generated stochastic processes. *Trans. AMS* **373** (2020) no. 1 435–478.
[58] Freitas A., Freitas J., Todd M.: Hitting time statistics and extreme value theory. *Probab. Theory Related Fields* **147** (2010) no. 3-4 675–710.
[59] Freitas A., Freitas J., Todd M.: Extreme value laws in dynamical systems for non-smooth observations. *J. Stat. Phys.* **142** (2011) no. 1 108–126.
[60] Galatolo, S.: Dimension via Waiting Time and Recurrence. *Math. Res. Lett.* **12** (2005) 377–386.
[61] Galatolo, S.: Dimension and Hitting in Rapidly Mixing System. *Math. Res. Lett.* **14** (2007) 797–805.
[62] Galatolo, S.: Hitting time in regular sets and logarithm law for rapidly mixing dynamical systems. *Proc. AMS* **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) 185–204.
[65] Ghosh, A., Kelmer, D.: Shrinking targets for semisimple groups. *Bull. LMS* **49** (2017) 235–245.
[66] Gorodnik A., Nevo A. *Quantitative ergodic theorems and their number-theoretic applications,* Bull. AMS **52** (2015) no. 1 65–113.
[67] Gorodnik, A., Spatzier, R.: Exponential mixing of nilmanifold automorphisms. *J. Anal. Math.* **123** (2014) 355–396.
[68] Gouezel, S.: A Borel-Cantelli lemma for intermittent interval maps. *Nonlinearity* **20** (2007) 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. ENS* **26** (1993) 23–50, **29** (1996) 811–814.
[71] Gupta, C., Nicol, M., Ott, W.: A Borel-Cantelli lemma for nonuniformly expanding dynamical systems, *Nonlinearity* **23** (2010) 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) 475–498.
[74] Haydn, N., Nicol, M., Vaienti, S., Zhang, L.: Central limit theorems for the shrinking target problem. *J. Stat. Phys.* **153** (2013) 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. 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) 2585–2611.

[79] Hill, R., Velani, S.: *The ergodic theory of shrinking targets, Invent. Math.* **119** (1995) 175–198.

[80] Hirata M.: Poisson law for Axiom A diffeomorphisms. *Erg. Th. & Dyn. Sys.* **13** (1993) 533–556.

[81] Hirata M.: Poisson law for the dynamical systems with self-mixing conditions. In *Dynamical systems and chaos,* Ed. N. Aoki, K. Shiraiwa & Y. Takahashi, World Scientific, Singapore (1995) v. 1, 87–96.

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

[83] 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.

[84] Ibragimov I. A.: Some limit theorems for stationary processes. *Teor. Verojatnost. i Primenen.* **7** (1962) 361–392.

[85] Jaerisch, J., Kessebohmer, M., Stratmann, B. O.: A Frechet law and an Erdos-Philipp law for maximal cuspidal windings, *Ergodic Theory Dynam. Systems* **33** (2013) 1008–1028.

[86] Keller G.: Rare events, exponential hitting times and extremal indices via spectral perturbation. *Dyn. Syst.* **27** (2012) 11–27.

[87] Kelmer, D.: Shrinking targets for discrete time flows on hyperbolic manifolds. *GAFA* **27** (2017) 1257–1287.

[88] Kelmer D. ,Yu S.: Shrinking target problems for flows on homogeneous spaces. Trans. AMS **372** (2019) 6283–6314.

[89] Kessebhmer 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.

[90] Kesten H., Maller R. A. Ratios of trimmed sums and order statistics. *Ann. Probab.* **20** (1992) 1805–1842.

[91] Khintchine A. Y.: Einige Satze uber Kettenbruche, mit Anwendungen auf die Theorie der Diophantischen Approximationen. *Math. Ann.* **92** (1924) 115–125.

[92] Kirsebom M., Kunde P., Persson T.: Shrinking targets and eventually always hitting points for interval maps, arXiv: 1903.06977.

[93] Kim, D. H.: The shrinking target property of irrational rotations. *Nonlinearity 20* **7** (2007) 1637–1643.

[94] Kim, D. H.: The dynamical Borel-Cantelli lemma for interval maps. *Discrete Contin. Dyn. Syst.* **17** (2007) 891–900.

[95] Kim, D. H.: Refined shrinking target property of rotations. Nonlinearity **27** (2014) 1985–1997.

[96] Kleinbock D. Some applications of homogeneous dynamics to number theory. Smooth ergodic theory and its applications (Seattle, WA, 1999), 639?660, *AMS Proc. Sympos. Pure Math.* **69** (2001) 639–660.

[97] Kleinbock D., Konstantoulas I., Richter F. K.: Zero-one laws for eventually always hitting points in mixing systems, arXiv:1904.08584.

[98] Kleinbock, D. Y., Margulis, G. A.: Bounded orbits of nonquasiunipotent flows on homogeneous spaces. *AMS Transl.* **171** (1996) 141–172.

[99] Kleinbock, D. Y., Margulis, G. A.: Flows on homogeneous spaces and Diophantine approximation on manifolds. *Ann. Math.* 148 (1998) 339–360.

[100] Kleinbock, D. Y., Margulis, G. A.: Logarithm laws for flows on homogeneous spaces. *Invent. Math,* **138** (1999) 451–494.

[101] 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, 813–930 (2002) North-Holland, Amsterdam.

[102] 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.

[103] Kleinbock, D. Y. and Zhao, X.: An application of lattice points counting to shrinking target problems. *Discrete Contin. Dyn. Syst.* **38** (2018) 155–168.

[104] Kurzweil, J.: On the metric theory of inhomogeneous Diophantine approximations. *Studia Math.* **15** (1955) 84–112.

[105] Leadbetter M. R., Lindgren G., Rootzen H.: Extremes and related properties of random sequences and processes, Springer, New York–Berlin, 1983. xii+336 pp.

[106] Ledrappier, F., Young, L. S.: The metric entropy of diffeomorphisms: Part II: Relations between entropy, exponents and dimension. *Annals of Mathematics* **122** (1985) 540–574.

[107] 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.

[108] Liverani, C.: On contact Anosov flows. *Ann. Math.* **159** (2004) 1275–1312.

[109] 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, Hoboken, NJ, 2016. xi+295 pp.

[110] Marklof, J.: The n-point correlations between values of a linear form. *Erg. Th., Dynam. Sys.* **20** (2000) 1127–1172.

[111] Marklof J.: Distribution modulo one and Ratner's theorem. *NATO Sci. Ser. Math. Phys. Chem.* **237** (2007) 217–244.

[112] Marklof J.: Entry and return times for semi-flows. *Nonlinearity* **30** (2017) no. 2 810–824.

[113] Maucourant, F.: Dynamical Borel-Cantelli lemma for hyperbolic spaces. *Israel J. Math.* **152** (2006) 143–155.

[114] Mori, T.: The strong law of large numbers when extreme terms are excluded from sums. *Z. Wahrsch. Verw. Gebiete* **36** (1976) 189–194.

[115] Paccaut F.: Statistics of return times for weighted maps of the interval. Ann. Inst. H. Poincare Probab. Statist. **36** (2000) 339–366.

[116] 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.

[117] Parry, W., Pollicott, M.: Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Asterisque* **187-188** (1990) 268 pp.

[118] Philipp, W.: Some metric theorems in number theory. *Pacific J. Math* **20** (1967) 109–127.

[119] Pitskel, B.: Poisson limit law for Markov chains. *Ergodic Th. Dynam. Sys.* **11** (1991), no. 3, 501–513.

[120] Pene F., Saussol B. Poisson law for some non-uniformly hyperbolic dynamical systems with polynomial rate of mixing. *Ergodic Th. Dynam. Sys.* **36** (2016) no. 8 2602–2626.

[121] Saussol, B.: On fluctuations and the exponential statistics of return times. Nonlinearity **14** (2001) no. 1 179–191.

[122] Saussol, B.: Recurrence rate in rapidly mixing dynamical systems. *Discrete Contin. Dyn. Syst.* **15** (2006) 259–267.

[123] Saussol, B.: An introduction to quantitative Poincar recurrence in dynamical systems. *Rev. Math. Phys.* **21** (2009) 949–979.

[124] Saussol, B., Troubetzkoy, S., Vaienti, S.: Recurrence, dimensions, and Lyapunov exponents. *J. Stat. Phys.* **106** (2002) 623–634.

[125] Sinai, Ya. G.: Gibbs measures in ergodic theory. *Russian Mathematical Surveys*, 27, 21, 1972.

[126] Sprindzhuk V. G. Metric theory of Diophantine approximations. V. H. Winston & Sons, Washington, D.C., 1979.

[127] Stenlund M.: A strong pair correlation bound implies the CLT for Sinai billiards, *J. Stat. Phys.* **140** (2010) 154–169.

[128] Stratmann, B., Velani, S. L.: The Patterson measure for geometrically finite groups with parabolic elements, new and old. *Proc. London Math. Soc.* **71**(3)197–220 (1995)

[129] Sullivan, D.: Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics. *Acta Math.* **149**(3-4) 215–237 (1982)

[130] Tseng, J.: On circle rotations and the shrinking target properties. *Discrete Contin. Dyn. Syst.* **20**(4), 1111–1122 (2008)

[131] Varandas P.: Correlation decay and recurrence asymptotics for some robust nonuniformly hyperbolic maps. *J. Stat. Phys.* **133** (2008) no. 5, 813–839.

[132] Viana, M.: *Stochastic dynamics of deterministic systems.* Rio de Janeiro: IMPA(1997)

[133] Vinogradov, I.: Limiting distribution of visits of several rotations to shrinking intervals, preprint.

[134] Williams, D.: *Probability with martingales.* Cambridge Math.Textbooks. Cambridge University Press, Cambridge, xvi+251 pp (1991).

[135] Yang F.: Rare event process and entry times distribution for arbitrary null sets on compact manifolds. arXiv:1905.09956.

[136] Young L.–S.: Statistical properties of dynamical systems with some hyperbolicity. *Ann. Math.* **147** (1998) no. 3, 585–650.

[137] Zhang X.: Note on limit distribution of normalized return times and escape rate. *Stoch. Dyn.* **16** (2016) no. 3, paper 1660014, 15 pp.