ISRAEL JOURNAL OF MATHEMATICS **247** (2022), 21–73
DOI: 10.1007/s11856-022-2289-3

MIXING PROPERTIES OF GENERALIZED T, T^{-1} TRANSFORMATIONS

BY

Dmitry Dolgopyat, Changguang Dong and Adam Kanigowski

Department of Mathematics, University of Maryland
College Park, MD 20742, USA
e-mail: dimadolgopyat@gmail.com, dongchangguang@gmail.com,
adkanigowski@gmail.com

AND

Peter Nándori

Department of Mathematical Sciences, Yeshiva University
New York, NY 10016, USA
e-mail: peter.nandori@yu.edu

Received April 25, 2020 and in revised form November 16, 2020

ABSTRACT

We study mixing properties of generalized T, T^{-1} transformations. We discuss two mixing mechanisms. In the case the fiber dynamics is mixing, it is sufficient that the driving cocycle is small with small probability. In the case the fiber dynamics is only assumed to be ergodic, one needs to use the shearing properties of the cocycle. Applications include the central limit theorem for sufficiently fast mixing systems and the estimates on deviations of ergodic averages.

Contents

1. Introduction

An important discovery made in the last century is that deterministic systems can exhibit chaotic behavior. Currently there are many examples of systems exhibiting a full array of chaotic properties including the Bernoulli property, exponential decay of correlations and the central limit theorem (see, e.g., [8, 9, 12, 61]). Systems which satisfy only some of the above properties are less understood. In fact, it is desirable to have more examples of such systems in order to understand the full range of possible behaviors of partially chaotic systems.

Generalized T, T^{-1} transformations are a rich source of examples in probability and ergodic theory. In fact, they were used to exhibit examples of systems with unusual limit laws [46, 14], a central limit theorem with non-standard normalization [7], K but non-Bernoulli systems in abstract [42] and smooth setting in various dimensions [44, 57, 43], very weak Bernoulli but not weak Bernoulli partitions [16], slowly mixing systems [17, 48], systems with multiple Gibbs measures [29, 50].

A comprehensive survey of a probabilistic version of T, T^{-1} transformations, which is a random walk in random scenery, is contained in [18]. On the other hand, there are no works addressing how statistical properties of T, T^{-1} transformations depend on the properties of the base and the fiber dynamics. Our paper provides a first step in this direction by investigating mixing properties of T, T^{-1} transformations.

Let us explain what we mean by smooth T, T^{-1} transformations. Let X, Y be compact manifolds, $f : X \to X$ be a smooth map preserving a measure μ and $G_t : Y \to Y$ be a d parameter flow on Y preserving a measure ν. Let $\tau : X \to \mathbb{R}^d$ be a smooth map. We study the following map $F : (X \times Y) \to (X \times Y)$:

$$F(x, y) = (f(x), G_{\tau(x)}y).$$

Note that F preserves the measure $\zeta = \mu \times \nu$ and that

$$F^N(x, y) = (f^N x, G_{\tau_N(x)}y) \quad \text{where } \tau_N(x) = \sum_{n=0}^{N-1} \tau(f^n x).$$

Clearly both mixing of f and ergodicity of G are necessary for F to be mixing. Under these assumptions there are two mechanisms for F to be mixing:

(1) If G itself is mixing, then it is enough to ensure that τ_N does not take small values with large probability (cf. [17, 48]).

(2) On the other hand, if we only assume that G is ergodic, then we need to rely on shearing properties of τ to ensure that τ_N is uniformly distributed in boxes of size 1. This can be done by assuming various extensions of the central limit theorem (cf. [10, 25]).

Abstract results detailing sufficient conditions for each of the two mechanisms described above are presented in Section 2. Estimates on the rates of mixing of F under the assumption that G is mixing are given in Section 4. In Section 5, we prove the central limit theorem in case F mixes sufficiently quickly. Section 6 contains mixing estimates in case G is only assumed to be ergodic (however, we need much stronger assumptions on the base map f). The results presented in Sections 4–6 rely on preliminary facts contained in Section 3. In Section 7, we discuss several examples which require a combination of ideas from Sections 4 and 6. Section 8 presents application of our mixing results to deviations of ergodic averages and also contains a survey of examples of systems satisfying various assumptions required in our results. We will have some strong assumptions that are sometimes non-trivial to check. In the appendix, we check one of our assumptions for an important example, namely the anticoncentration large deviation bounds for subshifts of finite type. This result may be interesting outside of the scope of the present work.

We also mention that in a followup paper [23] we provide a description of further statistical properties of the generalized T, T^{-1} transformation, using the mixing bounds obtained in the present paper.

Acknowledgements. D. D. was partially supported by the NSF grant DMS-1956049, A. K. was partially supported by the NSF grant DMS-1956310, P. N. was partially supported by the NSF grants DMS-1800811 and DMS-1952876.

2. The local limit theorem and mixing

For a function $A \in L^1(X, \mu)$ we denote $\mu(A(\cdot)) := \int_X A(x)\, d\mu$.

Definition 2.1: We say that τ satisfies the **mixing LLT** if there exist sequences $(L_n)_{n\in\mathbb{N}} \subset \mathbb{R}$, $(D_n)_{n\in\mathbb{N}} \subset \mathbb{R}^d$ and a bounded probability density $\mathfrak{p}$ on $\mathbb{R}^d$ such that for any sequence $(\delta_n)_{n\in\mathbb{N}} \subset \mathbb{R}$, with $\lim_{n\to\infty} \delta_n = 0$, $(z_n)_{n\in\mathbb{N}} \subset \mathbb{R}^d$ such that $|\frac{z_n}{L_n} - z| < \delta_n$ for any cube $\mathcal{C} \subset \mathbb{R}^d$ and any continuous functions $A_0, A_1 : X \to \mathbb{R}$,

$$\lim_{n\to\infty} L_n^d \mu(A_0(\cdot)A_1(f^n\cdot)\mathbb{1}_{\mathcal{C}}(\tau_n - D_n - z_n)) = \mathfrak{p}(z)\mu(A_0)\mu(A_1)\operatorname{Vol}(\mathcal{C}),$$

and the convergence is uniform once $(\delta_n)_{n\in\mathbb{N}}$ is fixed and A_0, A_1, z range over compact subsets of $C(X), C(X)$ and $\mathbb{R}^d$ respectively.

Definition 2.2: We say that τ satisfies the **mixing multiple LLT** if for each $m \in \mathbb{N}$, any sequence $(\delta_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ with $\lim_{n\to\infty} \delta_n = 0$, and any family of sequences $(z_n^{(1)}, \ldots, z_n^{(m)})_{n\in\mathbb{N}}$ with $|\frac{z_n^{(j)}}{L_n} - z^{(j)}| < \delta_n$, any cubes $\{\mathcal{C}_j\}_{j\le m} \subset \mathbb{R}^d$ and continuous functions $A_0, \ldots, A_m : X \to \mathbb{R}$, for any sequences $n_k^{(1)}, \ldots, n_k^{(m)} \in \mathbb{N}$

such that $n_k^{(j)} - n_k^{(j-1)} \geq \delta_k^{-1}$ (with $n_k^{(0)} = 0$),

$$\lim_{k\to\infty} \left(\prod_{j=1}^{m} L^d_{n_k^{(j)} - n_k^{(j-1)}} \right) \mu \left(\prod_{j=0}^{m} A_j(f^{n_k^{(j)}} \cdot) \prod_{j=1}^{m} \mathbb{1}_{\mathcal{C}_j}(\tau_{n_k^{(j)}} - D_{n_k^{(j)}} - z^{(j)}_{n_k^{(j)}}) \right)$$
$$= \prod_{j=0}^{m} \mu(A_j) \prod_{j=1}^{m} \mathfrak{p}(z^{(j)} - z^{(j-1)}) \prod_{j=1}^{m} \mathrm{Vol}(\mathcal{C}_j)$$

where $z^{(0)} = 0$. Moreover, the convergence is uniform once $(\delta_n)_{n\in\mathbb{N}}$ is fixed, $A_0, \dots, A_m$ range over compact subsets of $C(X)$ and $z^{(j)}$ range over a compact subset of $\mathbb{R}^d$ for every $j \leq m$.

Remark 2.3: We note that τ is bounded and consequently τ_n/n is bounded, too. Thus if the mixing LLT holds, then $L_n < Cn$. We assume that $D_n = n\mu(\tau)$. In case $\mu(\tau) = 0$, we say that τ has zero drift.

Remark 2.4: By the Portmanteau theorem on vague convergence, the mixing LLT is equivalent to saying that for all continuous functions $A_0, A_1 : X \to \mathbb{R}$ for any compactly supported almost everywhere continuous function $\phi : \mathbb{R}^d \to \mathbb{R}$ for any sequence z_N such that $|\frac{z_N}{L_N} - z| < \delta_n$, we have

$$\lim_{n\to\infty} L_n^d \mu(A_0(\cdot)A_1(f^n\cdot)\phi(\tau_n - D_n - z_n)) = \mathfrak{p}(z)\mu(A_0)\mu(A_1) \int_{\mathbb{R}^d} \phi(t)dt \tag{2.1}$$

and the convergence is uniform if A_0, A_1 range over compact subsets of $C(X)$ and z ranges over a compact subset of $\mathbb{R}^d$. A similar remark applies to the multiple mixing LLT.

Theorem 2.5: *Suppose that (G_t) is ergodic.*

(a) *If τ satisfies the mixing LLT then F is mixing.*
(b) *If τ satisfies the mixing multiple LLT then F is multiple mixing.*

Proof. (a) For $i = 1, 2$, let $\Phi_i(x, y) = A_i(x)B_i(y)$ be a continuous function on $X \times Y$. Since linear combinations of products as above are dense in $L^2(\mu\times\nu)$, it suffices to show that for every $\epsilon > 0$ there exists $N_0 \in \mathbb{N}$ such that for every $N \geq N_0$, we have

$$\left| \int_{X\times Y} \Phi_1(x, y)\Phi_2(F^N(x, y))d(\mu \times \nu) - \mu(A_1)\mu(A_2)\nu(B_1)\nu(B_2) \right| < \epsilon. \tag{2.2}$$

Let $\rho(t) := \int_Y B_1(y)B_2(G_t y)d\nu(y)$. Note that

$$\int_{X\times Y}\Phi_1(x,y)\Phi_2(F^N(x,y))d(\mu\times\nu)=\int_X A_1(x)A_2(f^N(x))\cdot\rho(\tau_N(x))d\mu(x). \tag{2.3}$$

Let $\delta = \delta(\epsilon) > 0$ be small with respect to ϵ, and $I_0 \subset \mathbb{R}^d$ be a cube of volume δ^d, centered at 0. Consider a (disjoint) cover of $\mathbb{R}^d$ by a union of small cubes $\{I_j\}$, where I_j is a translation of I_0, and let t_j denote the center of I_j. Now let $\mathbf{B}_\ell \subset \mathbb{R}^d$ be a ball centered at 0 with radius ℓ, and denote

$$S_\ell := \{j : I_j \cap \mathbf{B}_\ell \neq \emptyset\}.$$

By the mixing LLT (with $A_0 = A_1 = 1$) it follows that there exists $K = K(\epsilon)$ and $N_0' \in \mathbb{N}$ such that for every $N \geq N_0'$,

$$\mu(\{x \in X : |\tau_N - D_N| > KL_N/2\}) < \epsilon/2.$$

Let $\hat{S}_1 := S_{KL_N}$. Therefore (see (2.2) and (2.3)) it is enough to show that

$$\begin{aligned}\Bigg|\sum_{j\in\hat{S}_1}\int A_1(x)A_2(f^N(x))\cdot\rho(\tau_N(x))\mathbb{1}_{I_j+D_N}(\tau_N(x))d\mu(x)\\ -\mu(A_1)\mu(A_2)\nu(B_1)\nu(B_2)\Bigg| < \epsilon/2.\end{aligned} \tag{2.4}$$

If δ is small enough (using continuity of (G_t)), the above sum is, up to an error less than $\epsilon/16$, equal to

$$\sum_{j\in\hat{S}_1}\rho(D_N+t_j)\mu(A_1(\cdot)A_2(f^N(\cdot))\mathbb{1}_{I_j}(\tau_N(\cdot)-D_N)). \tag{2.5}$$

By the definition of the mixing LLT (with $A_1, A_2, \mathcal{C} = I_0$ and $z = t_j$), and since the number of j's such that $j \in \hat{S}_1$ is bounded above by $C(\delta,\epsilon)L_N^d$, there exists $N_1 = N_1(\epsilon,\delta) \in \mathbb{N}$ such that for every $N \geq N_1$, the above expression is, up to an error less than $\epsilon/16$, equal to

$$\sum_{j\in\hat{S}_1}\frac{1}{L_N^d}\operatorname{Vol}(I_0)\mathfrak{p}\Big(\frac{t_j}{L_N}\Big)\mu(A_1)\mu(A_2)\rho(D_N+t_j). \tag{2.6}$$

Enlarging K and N, if necessary, we can guarantee that

$$\Bigg|\sum_{j\in\hat{S}_1}\frac{1}{L_N^d}\operatorname{Vol}(I_j)\mathfrak{p}\Big(\frac{t_j}{L_N}\Big)-1\Bigg| < \frac{\epsilon}{16}. \tag{2.7}$$

Now, fix $R>0$ and for $c\in\mathbf{B}_R$, let

$$\alpha(c) := \sum_{j\in\hat{S}_1} \frac{1}{L_N^d}\operatorname{Vol}(I_0)\mathfrak{p}\Big(\frac{t_j}{L_N}\Big)\rho(D_N+t_j+c).$$

We claim that there exists $N_2=N_2(R)$ such that for $N\geq N_2$, we have

$$|\alpha(c)-\alpha(0)|<\epsilon/16.$$

Indeed, let k be such that $c\in I_k$, then $|t_k|\leq R+1$ and $|t_k-c|\leq\delta$, by choosing $\delta\ll\epsilon$ small enough, and N_2 large so that $\frac{R+1}{L_{N_2}}\leq\delta$, we have

$$\begin{aligned}
|\alpha(c)-\alpha(0)| \leq& |\alpha(c)-\alpha(t_k)|+|\alpha(t_k)-\alpha(0)|\\
\leq& \frac{\operatorname{Vol}(I_0)}{L_N^d}\sum_{j\in\hat{S}_1}\mathfrak{p}\Big(\frac{t_j}{L_N}\Big)|\rho(D_N+t_j+c)-\rho(D_N+t_j+t_k)|\\
&+\frac{\operatorname{Vol}(I_0)}{L_N^d}\sum_{j\in\hat{S}_1}\Big|\mathfrak{p}\Big(\frac{t_j}{L_N}\Big)-\mathfrak{p}\Big(\frac{t_j-t_k}{L_N}\Big)\Big||\rho(D_N+t_j)|\\
&+\frac{\operatorname{Vol}(I_0)}{L_N^d}\sum_{j\in\hat{S}_1:|t_j-t_k|\geq KL_N}\mathfrak{p}\Big(\frac{t_j}{L_N}\Big)|\rho(D_N+t_j+t_k)|\\
(\star)\qquad \leq& C_1(\mathfrak{p},\rho)|t_k-c|+C_2(\mathfrak{p},\rho,K)R/L_N+K^dC(\rho)R/L_N\\
\leq& \epsilon/64+\epsilon/64+\epsilon/64<\epsilon/16,
\end{aligned}$$

where for the inequality $(\star)$, the first term is due to the fact that ρ is continuous on t and (2.7), the second term is due to continuity of $\mathfrak{p}$ and the choice of N_2 (that is, $\frac{R+1}{L_N}\leq\delta$), and the last term contains a sum of $K^dRL_N^{d-1}$ many terms and hence $\leq K^dC(\rho)R/L_N$.

Therefore

$$\Big|\alpha(0)-\frac{1}{\operatorname{Vol}(\mathbf{B}_R)}\int_{c\in\mathbf{B}_R}\alpha(c)dc\Big|<\epsilon/16. \tag{2.8}$$

Now by the ergodicity of G and the mean ergodic theorem for the G-action, there exist a subset $Y_0\subset Y$ with $\nu(Y_0)\geq 1-\frac{\epsilon}{32C_3^2}$ and $R_0>0$, such that for any $y\in Y_0$ and $R\geq R_0$,

$$\Big|\frac{1}{\operatorname{Vol}(\mathbf{B}_R)}\int_{t\in\mathbf{B}_R}B_2(G_ty)dt-\nu(B_2)\Big|<\frac{\epsilon}{32C_3}.$$

Here the constant

$$C_3 := 10\max_{y\in Y}\{|B_1(y)|,|B_2(y)|\}.$$

Hence for any t, if $R \geq R_0$,

$$
\begin{aligned}
&\left|\frac{1}{\mathrm{Vol}(\mathbf{B}_R)}\int_{c\in\mathbf{B}_R}\rho(t+c)dc-\nu(B_1)\nu(B_2)\right| \\
&\quad\leq\left|\int_{G_{-t}(Y_0)}B_1(y)\left(\frac{1}{\mathrm{Vol}(\mathbf{B}_R)}\int_{c\in\mathbf{B}_R}B_2(G_{t+c}y)dc-\nu(B_2)\right)d\nu(y)\right| \\
&\qquad+\int_{Y\setminus G_{-t}(Y_0)}|B_1(y)|\left|\frac{1}{\mathrm{Vol}(\mathbf{B}_R)}\int_{c\in\mathbf{B}_R}B_2(G_{t+c}y)dc-\nu(B_2)\right|d\nu(y) \\
&\quad\leq\max\{|B_1|\}\frac{\epsilon}{32C_3}+\max\{|B_1|\}\max\{|B_2|\}2(1-\nu(Y_0))\leq\frac{\epsilon}{16}.
\end{aligned}
\tag{2.9}
$$

Note that (2.6) is equal to $\mu(A_1)\mu(A_2)\alpha(0)$. By (2.8) and (2.9), up to an error less than $\epsilon/8$, $\mu(A_1)\mu(A_2)\alpha(0)$ is equal to

$$\mu(A_1)\mu(A_2)\nu(B_1)\nu(B_2)\left[\sum_{j\in\hat{S}_1}\frac{1}{L_N^d}\mathrm{Vol}(I_j)\mathfrak{p}\left(\frac{t_j}{L_N}\right)\right].$$

Combining the estimates (2.7), (2.5) and (2.6) we obtain (2.4) (and consequently (2.2)), completing the proof.

(b) The proof is essentially the same as that for (a), therefore we leave it to the reader. ∎

3. Background

Definition 3.1: We say that G is **mixing with rate $\psi(t)$ on a space $\mathbb{B}$** if

$$\left|\int B_1(y)B_2(G_ty)d\nu(y)-\nu(B_1)\nu(B_2)\right|\leq C\psi(t)\|B_1\|_{\mathbb{B}}\|B_2\|_{\mathbb{B}}. \tag{3.1}$$

We call G **exponentially mixing** if (3.1) holds with $\mathbb{B}=C^r$ for some $r>0$ and $\psi(t)=e^{-\delta\|t\|}$ for some $\delta>0$.

We call G **polynomially mixing** if (3.1) holds with $\mathbb{B}=C^r$ for some $r>0$ and $\psi(t)=\|t\|^{-\delta}$ for some $\delta>0$.

We call G **rapidly mixing** if for each m there exists r such that (3.1) holds with $\mathbb{B}=C^r$ and $\psi(t)=\|t\|^{-m}$.

These definitions extend to maps (such as to f and F) in the natural way.

Definition 3.2: We say that τ satisfies **exponential large deviation bounds**, if for each $\varepsilon > 0$ there exist C and $\delta > 0$ such that for any $N \in \mathbb{N}$,

$$\mu\left(\left\|\frac{\tau_N}{N} - \mu(\tau)\right\| \geq \varepsilon\right) \leq Ce^{-\delta N}. \tag{3.2}$$

We say that τ satisfies **polynomial large deviation bounds**, if for each $\varepsilon > 0$ there exist C and $\delta > 0$ such that for any $N \in \mathbb{N}$,

$$\mu\left(\left\|\frac{\tau_N}{N} - \mu(\tau)\right\| \geq \varepsilon\right) \leq CN^{-\delta}.$$

We say that τ satisfies **superpolynomial large deviation bounds**, if for each $w > 0$, $\varepsilon > 0$ there exist $C = C(\varepsilon, w)$ such that for any $N \in \mathbb{N}$,

$$\mu\left(\left\|\frac{\tau_N}{N} - \mu(\tau)\right\| \geq \varepsilon\right) \leq CN^{-w}.$$

We will often use the following standard fact.

LEMMA 3.3: *For each r, there is $w = w(r)$ such that functions $\Phi \in C^w(X \times Y)$ admit a decomposition $\Phi(x, y) = \sum_{k=1}^{\infty} A_k(x)B_k(y)$, where $A_k \in C^r(X)$, $B_k \in C^r(Y)$ and*

$$\sum_k \|A_k\|_{C^r(X)}\|B_k\|_{C^r(Y)} \leq C(r, w)\|\Phi\|_{C^w(X\times Y)}. \tag{3.3}$$

COROLLARY 3.4: *Suppose that there are positive constants K and r, such that*

$$\begin{aligned}\left|\iint A'(x)B'(y)A''(f^n x)B''(G_{\tau_n(x)}y)d\mu(x)d\nu(y) - \mu(A')\nu(B')\mu(A'')\nu(B'')\right| \\ \leq K\|A'\|_{C^r(X)}\|B'\|_{C^r(Y)}\|A''\|_{C^r(X)}\|B''\|_{C^r(Y)}\psi(n).\end{aligned} \tag{3.4}$$

Then F is mixing with rate ψ.

Proof. Let

$$\bar\rho_n(\Phi', \Phi'') := \zeta(\Phi'(\Phi'' \circ F^n)) - \zeta(\Phi')\zeta(\Phi'').$$

Decomposing $\Phi', \Phi'' \in C^w$ as in (3.3), we get

$$\begin{aligned}|\bar\rho_n(\Phi', \Phi'')| = &\left|\sum_{j,k} \bar\rho_n(A'_jB'_j, A''_kB''_k)\right| \leq K\psi(n)\sum_{j,k}(\|A'_j\|_r\|B'_j\|_r\|A''_k\|_r\|B''_k\|_r) \\ \leq &K\psi(n)\sum_j(\|A'_j\|_r\|B'_j\|_r)\sum_k(\|A''_k\|_r\|B''_k\|_r) \\ \leq &K\psi(n)C^2(r, w)\|\Phi'\|_w\|\Phi''\|_w. \qquad \blacksquare\end{aligned}$$

4. Mixing rates for mixing fibers

4.1. Double mixing.

Theorem 4.1: *Suppose that $\mu(\tau) \neq 0$.*

(a) *If τ satisfies exponential large deviation bounds and f and G are exponentially mixing, then F is exponentially mixing.*
(b) *If τ satisfies polynomial large deviation bounds and f and G are polynomially mixing, then F is polynomially mixing.*
(c) *If τ satisfies superpolynomial large deviation bounds and f and G are rapidly mixing, then F is rapidly mixing.*

Proof. (a) For $i = 1, 2$, let $\Phi_i(x, y) = A_i(x)B_i(y)$ be a C^r function on $X \times Y$. Let

$$\rho(t) := \int_Y B_1(y)B_2(G_t y)d\nu(y).$$

Since G is exponentially mixing, there exist constants $C_1 > 0$ and $\kappa > 0$ such that

$$|\rho(t) - \nu(B_1)\nu(B_2)| \leq C_1\|B_1\|_{C^r}\|B_2\|_{C^r}e^{-\kappa\|t\|}. \tag{4.1}$$

Taking $\varepsilon = \|\mu(\tau)\|/2$ in the definition of exponential large deviation bounds, we find that there exist $C_0 > 0$ and $\delta > 0$ such that $\mu(T_N) \leq C_0e^{-\delta N}$, where

$$T_N := \{x \in X : \|\tau_N(x) - N\mu(\tau)\| \geq N\|\mu(\tau)\|/2\}.$$

Now note that

$$\int_{X\times Y}\Phi_1(x, y)\Phi_2(F^N(x, y))d(\mu\times\nu) = \int_X A_1(x)A_2(f^N(x))(\rho(\tau_N(x)))d\mu(x). \tag{4.2}$$

We rewrite the last integral as the sum of two integrals $\mathcal{I}_1 + \mathcal{I}_2$, where

$$\mathcal{I}_1 = \int_{T_N} A_1(x)A_2(f^N(x))\rho(\tau_N(x))d\mu(x)$$

and

$$\mathcal{I}_2 = \int_{X\setminus T_N} A_1(x)A_2(f^N(x))\rho(\tau_N(x))d\mu(x).$$

By exponential large deviation bounds, $|\mathcal{I}_1| \leq C_2\mu(T_N) \leq C_3e^{-\delta N}$. For $\mathcal{I}_2$, since f is exponentially mixing, it is enough to show that

$$\Delta := \left|\mathcal{I}_2 - \nu(B_1)\nu(B_2)\int_{X\setminus T_N} A_1(x)A_2(f^N(x))d\mu(x)\right|$$

is exponentially small. Indeed, by (4.1)

$$\begin{aligned}\Delta &\leq \left| \int_{X\setminus T_N} |A_1(x)||A_2(f^N(x))||\rho(\tau_N(x)) - \nu(B_1)\nu(B_2)|d\mu(x) \right| \\ &\leq C_4\|A_1\|_0\|A_2\|_0\|B_1\|_r\|B_2\|_r \cdot e^{-\kappa_1 N} \\ &\leq C_4\|A_1 \times B_1\|_r\|A_2 \times B_2\|_r \cdot e^{-\kappa_1 N}\end{aligned}$$

with $\kappa_1 = \kappa/2$. This finishes the proof. The proofs of parts (b) and (c) are analogous to part (a). We will omit them. ∎

Remark 4.2: In part (b) above, if τ satisfies polynomial large deviation bounds with rate $N^{-\delta_1}$, and f, G are polynomially mixing with rate $N^{-\delta_2}$ and $N^{-\delta_3}$ respectively, then F is polynomially mixing with rate $N^{-\min\{\delta_1,\delta_2,\delta_3\}}$.

Remark 4.3: Observe that the LLT was not needed in Theorem 4.1 and so the theorem remains valid if $\mathbb{R}^d$ is replaced by an arbitrary Lie group, in which case τ_N means the product

$$\tau_N(x) = \tau(f^{N-1}x)\cdots\tau(fx)\tau(x).$$

Definition 4.4: Assume that a cocycle τ is such that $\frac{\tau_n - D_n}{L_n}$ converges as $n \to \infty$ to a non-atomic distribution. We say that τ satisfies the **anticoncentration inequality** if for every unit cube $\mathcal{C} \subset \mathbb{R}^d$,

$$\mu(\{x \in X : \tau_N(x) \in \mathcal{C}\}) \leq CL_N^{-d},$$

for some global constant $C > 0$.

Remark 4.5: Note that by assumption there is a constant R such that

$$\mu(\|\tau_n\| \leq RL_n) \geq 0.5,$$

so the power of L_N in the anticoncentration inequality is optimal.

THEOREM 4.6: *Assume that for some $r \in \mathbb{N}$, f is mixing with rate $\psi_f(N) = L_N^{-\alpha}$, for some $\alpha > 0$ on C^r, τ satisfies the anticoncentration inequality and G is mixing with rate $\psi_G(\cdot)$ on C^r, where*

$$\int_{\mathbb{R}^d} \psi_G(t)dt < +\infty. \tag{4.3}$$

Then F is mixing with rate $\psi_F(N) := L_N^{-\min\{d,\alpha\}}$ on C^w for some $w = w(r) \in \mathbb{N}$.

THEOREM 4.7: *Assume that for some $r \in \mathbb{N}$, f is mixing with rate $\psi_f(N) = L_N^{-\alpha}$, for some $\alpha > 0$ on C^r, G is mixing with rate $\psi_G(\cdot)$ on C^r, τ satisfies the mixing LLT with zero drift.*

(a) *Suppose τ satisfies the anticoncentration inequality. If $\psi_G(\cdot)$ satisifies (4.3) and*

$$\int \Phi_1(x,y)d\nu(y) \equiv 0, \tag{4.4}$$

then

$$\begin{aligned}&\int \Phi_1(z)\Phi_2(F^N z)d\zeta(z)\\ &\quad = \mathfrak{p}(0)L_N^{-d} \iiiint \Phi_1(x,y)\Phi_2(\bar{x}, G_t y)d\mu(x)d\nu(y)d\mu(\bar{x})dt + o(L_N^{-d}).\end{aligned} \tag{4.5}$$

(b) *If $\psi_G(t) = \|t\|^{-\beta}$, for $\beta < d$, then F is mixing with rate*

$$\psi_F(N) := L_N^{-\min\{\beta,\alpha\}}$$

on C^w for some $w = w(r) \in \mathbb{N}$.

(c) *If $\min\{\alpha, d\} > \beta$ and for zero mean functions we have*

$$\int B_1(y)B_2(G_t y)d\nu = q(B_1, B_2)\Psi(t) + o(\|t\|^{-\beta}),$$

where q is a bounded bilinear form on $C^r(Y)$ and Ψ is a homogeneous function of degree $-\beta$, then

$$\int \Phi_1(z)\Phi_2(F^N z)d\zeta(z) = L_N^{-\beta} Q(\Phi_1, \Phi_2) \int_{\mathbb{R}^d} \mathfrak{p}(t)\Psi(t)dt + o(L_N^{-\beta}) \tag{4.6}$$

where

$$Q(\Phi_1, \Phi_2) = \int q(\Phi(x_1, \cdot), \Phi_2(x_2, \cdot))d\mu(x_1)d\mu(x_2).$$

Remark 4.8: In the case $d = 1$, (4.5) is proven in [48] under a slightly more restrictive condition.

Remark 4.9: We note that the integral in (4.6) converges. In fact, convergence near 0 follows because $\mathfrak{p}$ is bounded and $d > \beta$, while convergence near infinity follows since Ψ is bounded outside of the unit sphere. We also observe that for $\Phi_j(x,y) = A_j(x)B_j(y)$

$$Q(\Phi_1, \Phi_2) = \mu(A_1)\mu(A_2)q(B_1, B_2). \tag{4.7}$$

Proof of Theorem 4.6. For $i = 1, 2$, let

$$\Phi_i(x, y) = A_i(x)\tilde{B}_i(y), \quad \text{where } A_i \in C^r(X) \text{ and } \tilde{B}_i \in C^r(Y).$$

Let $B_i = \tilde{B}_i - \nu(\tilde{B}_i)$. Let $\rho(t) := \int_Y B_1(y)B_2(G_t y)d\nu(y)$. Note that

$$\begin{aligned}\int_{X\times Y}\Phi_1(x,y)\Phi_2(F^N(x,y))d(\mu\times\nu) &= \int_X A_1(x)A_2(f^N(x))\cdot\rho(\tau_N(x))d\mu(x)\\ &\quad + \nu(\tilde{B}_1)\nu(\tilde{B}_2)\int_X A_1(x)A_2(f^N(x))d\mu(x).\end{aligned} \tag{4.8}$$

Since f is mixing with rate $L_N^{-\alpha}$ on C^r, the second summand is equal to $\mu(A_1)\mu(A_2)$ up to an error less than $C\|A_1\|_r\|A_2\|_r L_N^{-\alpha}$. It remains to estimate the first summand.

Let $\{\mathcal{C}_i\}_{i=1}^{\infty}$ be a countable disjoint family of unit cubes in $\mathbb{R}^d$ such that

$$\mathbb{R}^d = \bigcup_i \mathcal{C}_i.$$

Below we assume without loss of generality that the function ψ from (4.3) satisfies

$$\sup_{\mathcal{C}_i}\psi(t) \le K\inf_{\mathcal{C}_i}\psi(t). \tag{4.9}$$

Indeed, given $t, \bar{t} \in \mathcal{C}_i$ we have

$$\nu(B_1\cdot B_2\circ G_t) = \nu(B_1\cdot\hat{B}_2\circ G_{\bar{t}}),$$

where $\hat{B}_2 = B_2\circ G_{t-\bar{t}}$. The last integral is smaller in absolute value than

$$\psi(\bar{t})\|B_1\|_{C^r}\|\hat{B}_2\|_{C^r} \le K\psi(\bar{t})\|B_1\|_{C^r}\|B_2\|_{C^r}.$$

Thus decreasing ψ if necessary we may assume that (4.9) holds.

Note first that since τ is bounded, we have

$$\begin{aligned}&\int_X A_1(x)A_2(f^N(x))\cdot\rho(\tau_N(x))d\mu(x)\\ &\quad = \sum_{i=1}^{\infty}\int_X A_1(x)A_2(f^N(x))\cdot\rho(\tau_N(x))\mathbb{1}_{\mathcal{C}_i}(\tau_N(x))d\mu(x).\end{aligned} \tag{4.10}$$

Using that G is mixing with rate ψ_G on C^r, (4.10) shows that

$$\begin{aligned}&\left|\int_X A_1(x)A_2(f^N(x))\cdot\rho(\tau_N(x))d\mu(x)\right|\\ &\qquad \le C\|A_1\|_0\|A_2\|_0\|B_1\|_r\|B_2\|_r\sum_{i=1}^{\infty}[\sup_{t\in\mathcal{C}_i}\psi_G(t)]\mu(\{x\in X : \tau_N(x)\in\mathcal{C}_i\}).\end{aligned}$$

Together with the anticoncentration inequality, we have

$$\begin{aligned}(4.11)\qquad &\left|\int_X A_1(x)A_2(f^N(x))\cdot\rho(\tau_N(x))d\mu(x)\right| \\ &\qquad \le CD\cdot\|A_1\|_0\|A_2\|_0\|B_1\|_r\|B_2\|_r L_N^{-d}\sum_{i=1}^{\infty}\sup_{t\in\mathcal{C}_i}\psi_G(t).\end{aligned}$$

Now by (4.9)

$$(4.12)\qquad \sum_{i=1}^{\infty}\sup_{t\in\mathcal{C}_i}\psi_G(t)\le C'\int_{\mathbb{R}^d}\psi_G(t)dt<C''.$$

Summarizing, we get

$$\int_X A_1(x)A_2(f^N(x))\cdot\rho(\tau_N(x))d\mu(x)\le C'''\|A_1\|_0\|A_2\|_0\|B_1\|_r\|B_2\|_r L_N^{-d}$$

showing that F is mixing with rate $L_N^{-\min\{d,\alpha\}}$. ∎

Proof of Theorem 4.7. By the same argument in the proof of Theorem 4.6 we just need to estimate

$$\int_X A_1(x)A_2(f^N(x))\cdot\rho(\tau_N(x))d\mu(x).$$

To prove part (a) note that due to (2.1) for each fixed i,

$$\begin{aligned}\lim_{N\to\infty} L_N^d\int_X A_1(x)A_2(f^N(x))\cdot\rho(\tau_N(x))\mathbb{1}_{\mathcal{C}_i}(\tau_N(x))d\mu(x)\\ = \mathfrak{p}(0)\int_{\mathcal{C}_i}\rho(t)dt\mu(A_1)\mu(A_2).\end{aligned}$$

This together with the Dominated Convergence Theorem (note that in part (a) we assume the conditions of Theorem 4.6 whence (4.11) and (4.12) apply) shows that

$$\lim_{N\to\infty} L_N^d\int_X A_1(x)A_2(f^N(x))\cdot\rho(\tau_N(x))d\mu(x)=\mathfrak{p}(0)\mu(A_1)\mu(A_2)\int_{\mathbb{R}^d}\rho(t)dt$$

proving (4.5).

To prove part (b), split

$$\int_X A_1(x)A_2(f^N(x))\cdot\rho(\tau_N(x))d\mu(x)=S_1+S_2,$$

where

$$S_1:=\int_X A_1(x)A_2(f^N(x))\cdot\rho(\tau_N(x))\mathbb{1}_{[-L_N,L_N]^d}(\tau_N(x))d\mu(x)$$

and

$$S_2 := \int_X A_1(x)A_2(f^N(x)) \cdot \rho(\tau_N(x))\mathbb{1}_{\mathbb{R}^d\setminus[-L_N,L_N]^d}(\tau_N(x))d\mu(x).$$

To estimate S_2, notice that for x as in S_2,

$$\begin{aligned}\rho(\tau_N(x)) &\le C\|B_1\|_r\|B_2\|_r\psi(\tau_N(x)) \le C_0\|B_1\|_r\|B_2\|_r\psi(L_N)\\ &\le C_0\|B_1\|_r\|B_2\|_r L_N^{-\beta}.\end{aligned}$$

Therefore

$$S_2 \le C_0\|A_1\|_0\|A_2\|_0\|B_1\|_r\|B_2\|_r L_N^{-\beta}.$$

It remains to estimate S_1. We trivially have

$$\begin{aligned}|S_1| =&\left|\int_X A_1(x)A_2(f^N(x)) \cdot \rho(\tau_N(x))\mathbb{1}_{[-L_N,L_N]^d}(\tau_N(x))d\mu(x)\right|\\ \le&\|A_1\|_0\|A_2\|_0\int_X|\rho(\tau_N(x))|\mathbb{1}_{[-L_N,L_N]^d}(\tau_N(x))d\mu(x).\end{aligned}\tag{4.13}$$

Cover $[-L_N, L_N]^d$ with (at most) $([L_N] + 1)^d$ disjoint cubes $\{I_j\}$ of size 1 centered at t_j, so that I_j's are translates of the cube I_0. By the mixing LLT for $z_n = t_j$ (notice that $\|t_j\| \le dL_N$ and so t_j/L_N belongs to a compact set), and $A_0 = A_1 = 1$, we get (for sufficiently large N),

$$L_N^d\mu(\{x \in X : \tau_N(x) \in I_j\}) < 2\mathfrak{p}^*\mathrm{Vol}(I_0) = 2\mathfrak{p}^*$$

where $\mathfrak{p}^* = \sup_t \mathfrak{p}$. Therefore,

$$\begin{aligned}\int_X|\rho(\tau_N(x))|\mathbb{1}_{[-L_N,L_N]^d}(\tau_N(x))d\mu(x) &= \sum_j\int_X|\rho(\tau_N(x))|\mathbb{1}_{I_j}(\tau_N(x))d\mu(x)\\ &\le 2\mathfrak{p}^*L_N^{-d}\sum_j\sup_{t\in I_j}|\rho(t)|\\ &\le CL_N^{-d}\int_{[-L_N,L_N]^d}\rho(t)dt\\ &\le CL_N^{-d}L_N^{d-\beta} = CL_N^{-\beta},\end{aligned}$$

completing the proof of (b).

To prove part (c), fix a small δ and split

$$\int_X A_1(x)A_2(f^N(x)) \cdot \rho(\tau_N(x))d\mu(x) = S_1 + S_2 + S_3,$$

where the integrand in S_1 is multiplied by $\mathbb{1}_{[-\delta L_N,\delta L_N]^d}(\tau_N(x))$, the integrand in S_2 is multiplied by $\mathbb{1}_{[-L_N/\delta,L_N/\delta]^d\setminus[-\delta L_N,\delta L_N]^d}(\tau_N(x))$ and the integrand in S_3

is multiplied by $\mathbb{1}_{\mathbb{R}^d\setminus[-L_N/\delta,L_N/\delta]^d}(\tau_N(x))$. Arguing as in the proof of part (b) we obtain that

$$S_3 = O\Big(\Big(\frac{\delta}{L_N}\Big)^{\beta}\Big).$$

Since the integrand is bounded, we have

$$S_1 = O\Big(\Big(\frac{\delta}{L_N}\Big)^{d}\Big) = O\Big(\Big(\frac{\delta}{L_N}\Big)^{\beta}\Big).$$

To handle S_2 we divide the domain of integration into unit cubes I_j. Let t_j be the center of I_j. Using the homogenuity of Ψ we conclude from the mixing LLT that

$$\int A_1(x)A_2(f^N x)\rho(\tau_N(x))\mathbb{1}_{I_j}(\tau_N(x))d\mu(x)$$
$$= L_N^{-(d+\beta)}\mu(A_1)\mu(A_2)q(B_1,B_2)\mathfrak{p}\Big(\frac{t_j}{L_N}\Big)\Psi\Big(\frac{t_j}{L_N}\Big) + o(L_N^{-(d+\beta)}).$$

Summing over j and using (4.7) we obtain

$$S_2 = L_N^{-\beta}Q(\Phi_1,\Phi_2)\int_{\mathcal{T}_\delta}\mathfrak{p}(t)\Psi(t)dt + o(L_N^{-\beta}),$$

where the domain of integration is $\mathcal{T}_\delta = [-\frac{1}{\delta},\frac{1}{\delta}]^d \setminus [-\delta,\delta]^d$. Combining our estimates for S_1, S_2 and S_3 we obtain

$$\int \Phi_1(z)\Phi_2(F^n z)d\zeta(z) = L_N^{-\beta}Q(\Phi_1,\Phi_2)\int_{\mathcal{T}_\delta}\mathfrak{p}(t)\Psi(t)dt + o(L_N^{-\beta}) + O\Big(\Big(\frac{\delta}{L_N}\Big)^{\beta}\Big).$$

Letting $\delta \to 0$ we obtain (4.6) for product observables, which by Lemma 3.3 is sufficient to conclude the general case. ∎

Remark 4.10: Note that the fact that $\mathbb{B} = C^r$ was only used to decompose any $\Phi \in C^w(X\times Y)$ as

$$\Phi(x,y) = \sum_n A_n(x)B_n(y), \quad \text{where } \sum_n \|A_n\|_{C^r}\|B_n\|_{C^r} < \infty. \tag{4.14}$$

Therefore the conclusions of Theorems 4.6 and 4.7 remain valid if (3.1) holds on an arbitrary space $\mathbb{B}$ provided that Φ_1, Φ_2 admit decomposition (4.14).

Remark 4.11: The results of this section apply (with obvious modifications) to continuous time T, T^{-1} systems of the form

$$F^t(x,y) = (\phi^t(x), G_{\tau_t(x)}y), \tag{4.15}$$

where ϕ is a flow on X and

$$\tau_t(x) = \int_0^t \tau(\phi^s(x))ds. \tag{4.16}$$

Note that due to the fact that

$$\zeta(H_1(H_2 \circ F^{n+\delta})) = \zeta(H_1((H_2 \circ F^\delta) \circ F^n))$$

it is sufficient to control the correlation at integer times. Next F^1 is a T, T^{-1}-transformation corresponding to $f = \phi^1$, $\tau = \tau_1$. We note, however, that in several cases for time one maps of the flow the LLT is unknown (or false) unless the observable is the time integral given by (4.16). We refer the reader to [27] for the discussion of mixing LLT for continuous time systems.

Example 4.12: (a) Let g_t be an exponentially mixing Anosov flow on some manifold M. Consider a continuous T, T^{-1} system F_1^t with $X = Y = M$ and $\phi^t = G_t = g^t$. Then Theorem 4.7(a) shows that for smooth zero mean observables

$$\lim_{t\to\infty} \sqrt{t}\zeta(H_1(H_2 \circ F^t)) = Q_1(H_1, H_2),$$

where Q_1 is given by (4.5). Indeed, the condition (4.4) can be relaxed and the conclusion of Theorem 4.7(a) holds for all zero mean smooth observables assuming that $\alpha > d$ (in this example, α is arbitrarily large and $d = 1$).

(b) For any positive integer k, define inductively a continuous T, T^{-1} system F_k^t with $X = M$, $Y = M^k$, $\phi^t = g^t$ and $G_t = F_{k-1}^t$, where F_1^t is the flow from the part (a). Then Theorem 4.7(c) shows that for smooth zero mean observables

$$\lim_{t\to\infty} t^{2^{-k}} \zeta(H_1(H_2 \circ F^t)) = Q_k(H_1, H_2),$$

where Q_k is given in terms of Q_{k-1} by (4.6).

4.2. Multiple mixing.

Definition 4.13: G_t is **mixing** of **order** s with **rate** ψ on a space $\mathbb{B}$ if

$$\left|\nu\left(\prod_{j=1}^s B_j(G_{t_j} y)\right) - \prod_{j=1}^s \nu(B_j)\right| \leq C\psi(\delta(t_1, \ldots t_s)) \prod_{j=1}^s \|B_j\|_{\mathbb{B}}$$

where

$$\delta(t_1, \ldots, t_s) = \min_{i\neq j} \|t_i - t_j\|.$$

This definition extends to maps (such as to f and F) in the natural way.

THEOREM 4.14: *If τ satisfies mixing LLT with zero drift and f and G are mixing of order s with rate $t^{-\alpha}$ with $\alpha > d$, then F is mixing of order s with rate $\psi_F(N) = L_N^{-d}$.*

Proof. For $i = 1, \ldots, s$, let $\Phi_i(x,y) = A_i(x)B_i(y)$, where $A_i \in C^r(X)$ and $B_i \in C^r(Y)$. Let $\rho(t_1, t_2, \ldots, t_s) := \int_Y \prod_{i=1}^s B_i(G_{t_i} y) d\nu(y)$ (with $t_1 = 0$). We have

$$\begin{aligned} &\int_{X\times Y} \prod_{i=1}^s \Phi_i(F^{N_i}(x,y)) d(\mu \times \nu) \\ &\quad = \int_X \prod_{i=1}^s A_i(f^{N_i}x) \cdot \rho(\tau_{N_1}(x), \ldots, \tau_{N_s}(x)) d\mu(x) \\ &\qquad \times \int_X \prod_{i=1}^s A_i(f^{N_i}x) \cdot \left(\rho(\tau_{N_1}(x), \ldots, \tau_{N_s}(x)) - \prod_{i=1}^s \nu(B_i)\right) d\mu(x) \\ &\qquad + \prod_{i=1}^s \nu(B_i) \int_X \prod_{i=1}^s A_i(f^{N_i}x) d\mu(x). \end{aligned} \tag{4.17}$$

Note that since f is mixing of order s with rate $N^{-\alpha}$, the last term above is equal to $\prod_{i=1}^s \mu(A_i)\nu(B_i)$ up to an error of size at most

$$\mathrm{O}\left(\prod_{i=1}^s \|A_i\|_r \min_{i\neq j} |N_i - N_j|^{-\alpha}\right).$$

It is therefore enough to bound the first term. Notice, moreover, that since τ is bounded and satisfies mixing LLT with zero drift, we have $L_N \le C'N$ (see Definition 2.1).

Denote $\bar{N} := \min_{i\neq j} |N_i - N_j|$.

Let $Z \subset X$ be defined by setting: $x \in Z$ iff $\min_{i\neq j} \|\tau_{N_i}(x) - \tau_{N_j}(x)\| \ge L_{\bar{N}}$. Using that G is mixing of order s with rate $\|t\|^{-\alpha}$, we get

$$\begin{aligned} &\int_Z \prod_{i=1}^s A_i(f^{N_i}x) \cdot \left(\rho(\tau_{N_1}(x), \ldots, \tau_{N_s}(x)) - \prod_{i=1}^s \nu(B_i)\right) d\mu(x) \\ &\qquad\qquad \le C \prod_{i=1}^s \|A_i\|_0 \prod_{i=1}^s \|B_i\|_r L_{\bar{N}}^{-\alpha}. \end{aligned} \tag{4.18}$$

So it remains to estimate the above integral on Z^c. By definition, for every $x \in Z^c$, there exists $i_x \neq j_x$ such that

$$\|\tau_{N_{i_x}}(x) - \tau_{N_{j_x}}(x)\| = \min_{i\neq j} \|\tau_{N_i}(x) - \tau_{N_j}(x)\| \le L_{\bar{N}}. \tag{4.19}$$

Let $Z_{ij} := \{x \in Z^c : (i_x, j_x) = (i,j)\}$ (if there are several pairs satisfying (4.19) we take the smallest with respect to the lexicographic order). Let $\{C_k\}_{k=1}^{\bar{M}}$ be a finite family of unit cubes centered at $\{c_k\}_{k=1}^{\bar{M}}$ in $\mathbb{R}^d$ such that

$$[-L_{\bar{N}}, L_{\bar{N}}]^d = \bigcup_k C_k.$$

Then

$$\begin{aligned}
&\Big| \int_{Z_{ij}} \prod_{l=1}^{s} A_l(f^{N_l}x) \cdot \Big(\rho(\tau_{N_1}(x), \dots, \tau_{N_s}(x)) - \prod_{i=1}^{s} \nu(B_i)\Big) d\mu(x)\Big| \\
&= \Big| \sum_{k=1}^{\bar{M}} \int_{Z_{ij}} \prod_{l=1}^{s} A_l(f^{N_l}x) \cdot \Big(\rho(\tau_{N_1}(x), \dots, \tau_{N_s}(x)) - \prod_{i=1}^{s} \nu(B_i)\Big) \\
&\qquad \times \mathbb{1}_{C_k}(\tau_{N_i}(x) - \tau_{N_j}(x)) d\mu(x)\Big|.
\end{aligned} \tag{4.20}$$

Using that G is mixing of order s with rate $\|t\|^{-\alpha}$, and

$$\min\{\sup_{C_k} \|t\|^{-\alpha}, 1\} \le C \inf_{C_k} \|t\|^{-\alpha},$$

we get that the LHS of (4.20) is bounded above by

$$\begin{aligned}
&C' \prod_{l=1}^{s} (\|A_l\|_0 \|B_l\|_r) \sum_{k=1}^{\bar{M}} \Big(\int_{C_k} \min\{\|t\|^{-\alpha}, 1\} dt\Big) \\
&\qquad \times \mu(\{x \in X : \tau_{N_i}(x) - \tau_{N_j}(x) \in C_k\}).
\end{aligned} \tag{4.21}$$

Note that $\tau_{N_i}(x) - \tau_{N_j}(x) = \tau_{N_i - N_j}(f^{N_j}x)$. Hence, by the mixing LLT with $A_0 = A_1 = 1$, $D_n \equiv 0$, we get (by preservation of measure)

$$\mu(\{x \in X : \tau_{N_i}(x) - \tau_{N_j}(x) \in C_j\}) \le 2L_{N_i - N_j}^{-d} \mathfrak{p}(c_i / L_{\bar{N}}) < C L_{\bar{N}}^{-d}.$$

Therefore, (4.21) (and hence also (4.20)) is bounded above by (recall that $\alpha > d$)

$$C'' \prod_{l=1}^{s} (\|A_l\|_0 \|B_l\|_r) L_{\bar{N}}^{-d}.$$

Summing over all i, j and using (4.18), we get that the LHS of (4.17) is bounded by

$$C''' \prod_{l=1}^{s} (\|A_l\|_0 \|B_l\|_r) L_{\bar{N}}^{-d}.$$

This finishes the proof. ∎

THEOREM 4.15: *If τ has non-zero drift and satisfies exponential large deviation bounds, and f and G are exponentially mixing of order s, then F is exponentially mixing of order s.*

Proof. For $i = 1, 2, \ldots, s$, let $\Phi_i(x, y) = A_i(x)B_i(y)$ be a C^r function on $X \times Y$. Let

$$\rho(t_1, \ldots, t_s) := \int_Y \prod_{i=1}^{s} B_i(G_{t_i} y) d\nu(y)$$

(with $t_1 = 0$). Since G is exponentially mixing, there exist a constant $C_1 > 0$ and $\kappa > 0$ such that

$$\left|\rho(t_1, \ldots, t_s) - \prod_{i=1}^{s} \nu(B_i)\right| \leq C_1 \|B_1\|_{C^r} \|B_2\|_{C^r} e^{-\kappa\delta(t_1,\ldots,t_s)}. \tag{4.22}$$

Fix $0 = N_1 \leq N_2 \leq \cdots \leq N_s$. We again use the decomposition (4.17). By exponential mixing of order s of f, the second term in (4.17) is exponentially close to $\prod_{i=1}^{s} \nu(B_i) \prod_{i=1}^{s} \mu(A_i)$, and hence we only need to estimate the first term.

Let

$$T_{ij} := \{x \in X : \|\tau_{N_i}(x) - \tau_{N_j}(x) - (N_i - N_j)\mu(\tau)\| \geq (N_i - N_j)\|\mu(\tau)\|/2\}.$$

Let $\bar{T} = \bigcup_{i \neq j} T_{ij}$. By exponential large deviation bounds (and preservation of measure),

$$\mu(\bar{T}) \leq s^2 \max_{ij} \mu(T_{ij}) \leq Ce^{-\delta N}.$$

Therefore it is enough to bound the integral of the first term in the RHS on $X \setminus \bar{T}$. By exponential mixing of G,

$$\begin{aligned}\int_{X\setminus\bar{T}} \prod_{i=1}^{s} A_i(f^{N_i} x) \cdot \left(\rho(\tau_{N_1}(x), \ldots, \tau_{N_s}(x)) - \prod_{i=1}^{s} \nu(B_i)\right) d\mu(x)& \\ \leq C \prod_{i=1}^{s} \|A\|_0 \prod_{i=1}^{s} \|B_i\|_r \min_{x\notin\bar{T}} e^{-\kappa\delta(\tau_{N_1}(x),\ldots,\tau_{N_s}(x))}&.\end{aligned}$$

By the definition of $\bar{T}$,

$$\delta(\tau_{N_1}(x), \ldots, \tau_{N_s}(x)) \geq \frac{\|\mu(\tau)\|}{2} \min_{i\neq j} |N_i - N_j|,$$

completing the proof. ∎

Let $n_1 \le n_2 \le \cdots \le n_s$ be an s tuple. A partition

$$\mathfrak{P} = \mathcal{P}_1 \cup \mathcal{P}_2 \cup \cdots \cup \mathcal{P}_k$$

of a set $\{n_1, n_2, \ldots, n_s\}$ (where an item may be listed more than once) is called **social** if for each $j \in \{1, \ldots, k\}$, $\mathrm{Card}(\mathcal{P}_j) > 1$. An element n_j is called **forward free** (**backward free**) for partition $\mathfrak{P}$ if it is the smallest (respectively, the largest) in its atom. We call n_j forward (or backward) fixed if it is not forward (backward) free. We let $F^{\pm}$ be the set of all forward (or backward) fixed elements. Let

$$\kappa^{\pm}(\mathfrak{P}) = \prod_{n_j \in F^{\pm}} L_{n_j - n_{j-1}}.$$

For $\mathfrak{P} = (P_1, \ldots, P_k)$, let $(n_{i_\ell})_{\ell=1}^k$ be the collection of forward free elements, i.e., n_{i_ℓ} is the smallest element of P_ℓ. Analogously we define $(n_{j_\ell})_{\ell=1}^k$ to be the collection of backward free elements. Notice that we have the following formula for $\kappa^{\pm}(\mathfrak{P})$:

$$\kappa^{+}(\mathfrak{P}) = \left(\prod_{j=1}^{s} L_{n_j - n_{j-1}}\right) \cdot \left(\prod_{\ell=1}^{k} L_{n_{i_\ell} - n_{i_\ell - 1}}\right)^{-1}, \tag{4.23}$$

with $n_0 = 0$, and analogously

$$\kappa^{-}(\mathfrak{P}) = \left(\prod_{j=1}^{s} L_{n_j - n_{j-1}}\right) \cdot \left(\prod_{\ell=1}^{k} L_{n_{j_\ell + 1} - n_{j_\ell}}\right)^{-1}, \tag{4.24}$$

with $n_{s+1} := n_1 + n_s$.

We have the following

Definition 4.16: We say that τ satisfies the **anticoncentration large deviation bound of order** s if there exist a constant K and a decreasing function Θ such that $\int_1^\infty \Theta(r) r^d < \infty$, and for any unit cubes $C_1, C_2, \ldots, C_s$ centered at $c_1, c_2, \ldots, c_s$

$$\mu(x : \tau_{n_j} \in C_j \text{ for } j = 1, \ldots, s) \le K \left(\prod_{j=1}^{s} L_{n_j - n_{j-1}}^{-d}\right) \Theta\left(\max_j \frac{\|c_j - c_{j-1}\|}{L_{n_j - n_{j-1}}}\right).$$

Remark 4.17: For $s = 2$ anticoncentration large deviation bounds were considered in [26].

Theorem 4.18: *If τ satisfies anticoncentration large deviation bounds of order s and f and G are exponentially mixing of order s, then*

$$\left|\int\left(\prod_{j=1}^{s} H_j(F^{n_j}z)\right)d\zeta(z)-\prod_{j=1}^{s}\zeta(H_j)\right|\leq C\prod_{j=1}^{s}\|H_j\|_{C^r}(\min_{\mathfrak{P}}\kappa(\mathfrak{P}))^{-d} \tag{4.25}$$

where

$$\kappa(\mathfrak{P})=\max\{\kappa^+(\mathfrak{P}),\kappa^-(\mathfrak{P})\}$$

and the minimum in (4.25) *is taken over all social partitions of* $\{n_1,\dots n_s\}$.

We first recall the following result, which simplifies our analysis.

Lemma 4.19 ([6]): *If G is exponentially mixing of order s, then for some $\eta>0$*

$$\left|\nu\left(\prod_{j=1}^{s} B_j(G_{t_j}y)\right)-\prod_{j=1}^{s}\nu(B_j)\right|\leq Ce^{-\eta\Delta(t_1,\dots,t_s)}\prod_{j=1}^{s}\|B_j\|_{\mathbb{B}}, \tag{4.26}$$

where

$$\Delta(t_1,\dots,t_s)=\max_j\min_{i\neq j}\|t_i-t_j\|.$$

With the above lemma, we prove Theorem 4.18.

Proof of Theorem 4.18. By Lemma 3.3 it is enough to show the statement for $H_j=A_j\times B_j\in C^r(M)$. Let

$$\rho(t_1,\dots,t_s):=\nu\left(\prod_{j=1}^{s} B_j(G_{t_j}y)\right)-\prod_{j=1}^{s}\nu(B_j).$$

Then

$$\begin{aligned}&\int\left(\prod_{j=1}^{s} H_j(F^{n_j}z)\right)d\zeta(z)\\ &\quad=\int\left(\prod_{j=1}^{s} A_j(f^{n_j}x)\right)\rho(\tau_{n_1}(x),\dots,\tau_{n_s}(x))d\mu(x)\\ &\qquad+\left(\prod_{j=1}^{s}\nu(B_j)\right)\mu\left(\prod_{j=1}^{s} A_j(f^{n_j}x)\right).\end{aligned} \tag{4.27}$$

Since f is exponentially mixing of order s,

$$\left|\mu\left(\prod_{j=1}^{s} A_j(f^{n_j}x)\right)-\prod_{j=1}^{s}\mu(A_j)\right|\leq C\prod_{j=1}^{s}\|A_j\|_r e^{-\eta\Delta}, \tag{4.28}$$

where $\Delta=\Delta(n_1,\dots,n_s)$.

Let $\mathcal{P}$ be the following partition of $n_1 < \ldots < n_s$. Let $i_1 \in \{2, \ldots, s-1\}$ be the smallest index i such that $|n_i - n_{i-1}| > \Delta$. Then the first atom of $\mathcal{P}$ is $\{n_0, \ldots, n_{i_1-1}\}$. Notice that $|n_{i_1} - n_{i_1+1}| \leq \Delta$ by the definition of Δ. Now recursively, let $i_{k+1} \in \{i_k+1, \ldots, s\}$ be the smallest index i such that $|n_i - n_{i-1}| > \Delta$. Then the $(k+1)$-th atom of $\mathcal{P}$ is $\{n_{i_k}, \ldots, n_{i_{k+1}-1}\}$. We continue until we partition all of $n_1 < \cdots < n_s$. Then by the definition of Δ, every atom of $\mathcal{P}$ has at least two elements, and so $\mathcal{P}$ is social. Moreover, all elements in one atom are at distance at most $s\Delta$ (since the number of elements is $\leq s$). Using that τ is bounded (and so $|L_n| < Cn$) together with (4.23) and (4.24), we conclude that

$$\min\{\kappa^+(\mathcal{P})^{-d}, \kappa^+(\mathcal{P})^{-d}\} \geq \left(\prod_{j=1}^{s} L_{n_j - n_{j-1}}\right)^{-d} \gg [s\Delta]^{-sd} \geq C\Delta^{-sd} \geq Ce^{-\eta\Delta}.$$

Combining this estimate with (4.28) we find that the second term in (4.27) equals $\prod_{j=1}^{s} \zeta(H_j)$ up to an error which is bounded by the RHS of (4.25). It remains to show that

$$\left| \int \left(\prod_{j=1}^{s} A_j(f^{n_j}x)\right) \rho(\tau_{n_1}(x), \ldots, \tau_{n_s}(x)) d\mu(x) \right| \\ \leq C \prod_{j=1}^{s} \|A_j \times B_j\|_{C^r} (\min_{\mathfrak{P}} \kappa(\mathfrak{P}))^{-d},$$

which will follow by showing that

$$\int |\rho(\tau_{n_1}(x), \ldots, \tau_{n_s}(x))| d\mu(x) \leq C \prod_{j=1}^{s} \|B_j\|_{C^r} (\min_{\mathfrak{P}} \kappa(\mathfrak{P}))^{-d}.$$

Let

$$C_0 := \prod_{j=1}^{s} \|B_j\|_{C^r}$$

and

$$D_m := \{x : |\rho(\tau_{n_1}(x), \ldots, \tau_{n_s}(x))| \in [C_0 2^{-m}, C_0 2^{-m+1})\}.$$

Then

$$\int |\rho(\tau_{n_1}(x), \ldots, \tau_{n_s}(x))| d\mu(x) \leq 2C_0 \sum_{m \geq 0} \frac{1}{2^m} \mu(D_m). \tag{4.29}$$

We will estimate the measure of D_m. Note that by Lemma 4.19, for some $C_\eta \in \mathbb{N}$,

$$D_m \subset A_m := \{x : \Delta(\tau_{n_1}(x), \dots, \tau_{n_s}(x)) \le C_\eta m\}.$$

We will therefore give an upper bound on the measure of A_m. By the definition of Δ it follows that there exists a social partition $\mathfrak{P} = (P_1, \dots, P_k)$ of $n_1 < n_2 < \cdots < n_s$ such that for any atom of $\mathfrak{P}$ and any two n_i, n_j in the same atom we have

$$|\tau_{n_i}(x) - \tau_{n_j}(x)| < C_\eta sm. \tag{4.30}$$

Let $A_{m,\mathfrak{P}} \subset A_m$ be the set of x for which $\mathfrak{P}$ is a social partition of $n_1 < n_2 < \cdots < n_s$ satisfying (4.30). Then

$$A_m = \bigcup_{\mathfrak{P} \text{ social}} A_{m,\mathfrak{P}},$$

and so we will estimate the measure of $A_{m,\mathfrak{P}}$.

Let $\{\tilde{C}_j\}$ be a disjoint cover of $\mathbb{R}^d$ by cubes of side length $C_\eta s \cdot m$ centered and $\tilde{c}_j$. Note that by the anticoncentration large deviation bounds of order s (decomposing $\tilde{C}_j$ into unit cubes),

$$\begin{aligned}\mu(x : \tau_{n_j}(x) \in \tilde{C}_j \text{ for } j = 1, \dots, s) \\ \le K'(sm)^{sd} \Big(\prod_{j=1}^{s} L^{-d}_{n_j - n_{j-1}} \Big) \Theta\Big(\max_j \frac{\|\tilde{c}_j - \tilde{c}_{j-1}\|}{m \cdot L_{n_j - n_{j-1}}} \Big).\end{aligned} \tag{4.31}$$

It follows by the definition of $\mathfrak{P}$ and (4.30) that all the $\{\tau_{n_j}(x)\}_{n_j \in P_\ell}$ belong to one cube $\tilde{C}_{r_\ell}$. Below, we use the notation $\tau_{P_\ell}(x) \in C_{r_\ell}$ which means that for every $n_j \in P_\ell$, $\tau_{n_j}(x) \in \tilde{C}_{r_\ell}$. Therefore, we have

$$\mu(A_{m,\mathfrak{P}}) \le \sum_{r_1, \dots, r_k} \mu(\{x : \tau_{P_\ell}(x) \in \tilde{C}_{r_\ell}, \ell \le k\}).$$

Let n_{i_ℓ} (and n_{j_ℓ}) be the smallest (the largest) element of P_ℓ, $\ell \le k$. Below we will argue with (n_{i_ℓ}) (analogous reasoning can be done for (n_{j_ℓ})). Let $u(\ell)$ be such that $n_{i_\ell - 1} \in P_{u(\ell)}$. By (4.31), monotonicity of Θ and the above discussion (using that n_{i_ℓ} and $n_{i_\ell - 1}$ are in different atoms), we obtain

$$\mu(\{x : \tau_{P_\ell}(x) \in C_{r_\ell}, \ell \le k\}) \le K' m^{sd} \Big(\prod_{j=1}^{s} L^{-d}_{n_j - n_{j-1}} \Big) \Theta\Big(\max_{\ell \le k} \frac{\|\tilde{c}_{r_\ell} - \tilde{c}_{r_{u(\ell)}}\|}{m \cdot L_{n_{i_\ell} - n_{i_\ell - 1}}} \Big).$$

Therefore

$$\mu(A_{m,\mathfrak{P}}) \leq K' m^{sd}\Big(\prod_{j=1}^{s} L_{n_j-n_{j-1}}^{-d}\Big)\sum_{r_1,\dots r_k}\Theta\Big(\max_{\ell\leq k}\frac{\|\tilde{c}_{r_\ell}-\tilde{c}_{r_{u(\ell)}}\|}{m\cdot L_{n_{i_\ell}-n_{i_\ell-1}}}\Big).$$

Note that

$$\begin{aligned}\sum_{r_1,\dots r_k}&\Theta(\max_{\ell\leq k}\frac{\|\tilde{c}_{r_\ell}-\tilde{c}_{r_{u(\ell)}}\|}{m\cdot L_{n_{i_\ell}-n_{i_\ell-1}}})\\ &\leq\sum_{\ell}\Theta(\ell)\cdot|\{(r_1,\dots,r_k):\|\tilde{c}_{r_\ell}-\tilde{c}_{r_{u(\ell)}}\|\leq\ell\cdot m\cdot L_{n_{i_\ell}-n_{i_\ell-1}}\text{ for every }\ell\leq k\}|\\ &\leq\sum_{\ell}\Theta(\ell)\ell^d\cdot m^d\cdot\Big(\prod_{\ell\leq k}L_{n_{i_\ell}-n_{i_\ell-1}}\Big)^d.\end{aligned}$$

Therefore, by the decay assumptions on Θ and (4.23),

$$\mu(A_{m,\mathfrak{P}}) \leq K' m^{sd+d}\Big(\prod_{j=1}^{s} L_{n_j-n_{j-1}}^{-d}\Big)\cdot\Big(\prod_{\ell\leq k}L_{n_{i_\ell}-n_{i_\ell-1}}\Big)^d = K' m^{sd+d}\kappa^+(\mathfrak{P})^{-d}.$$

Analogously we have that

$$\mu(A_{m,\mathfrak{P}}) \leq K' m^{sd+d}\kappa^+(\mathfrak{P})^{-d}.$$

Therefore,

$$\mu(A_{m,\mathfrak{P}}) \leq K' m^{sd+d}\kappa(\mathfrak{P})^{-d}.$$

Using that $A_m = \bigcup_{\mathfrak{P}} A_{m,\mathfrak{P}}$, we get

$$\mu(A_m) \leq K' C_s m^{sd+d}(\min_{\mathfrak{P}}\kappa(\mathfrak{P}))^{-d},$$

for some constant $C_s > 0$. Summarizing, by (4.29) (since $D_m \subset A_m$), we get

$$\begin{aligned}\int|\rho(\tau_{n_1}(x),\dots,\tau_{n_s(x)})|d\mu(x) &\leq 2K'C_s\prod_{j=1}^{s}\|B_j\|_{C^r}(\min_{\mathfrak{P}}\kappa(\mathfrak{P}))^{-d}\sum_{m\geq 0}2^{-m}m^{sd+d}\\ &\leq C_{s,d}\prod_{j=1}^{s}\|B_j\|_{C^r}(\min_{\mathfrak{P}}\kappa(\mathfrak{P}))^{-d}.\end{aligned}$$

This finishes the proof. ∎

5. The central limit theorem

Let $H(x,y)$ be a C^r function not cohomologous to a constant function. Let

$$\Sigma_N(H) := \sum_{n=0}^{N-1} H(F^n(x,y)).$$

Assume that $\zeta(H)=0$. Let $Z = X \times Y$.

Theorem 5.1: *Suppose that F satisfies (4.25) and $\sum_{n=1}^{\infty} L_n^{-d}$ converges. Then $\frac{\Sigma_N(H)}{\sqrt{N}}$ converges as $N \to \infty$ to the normal distribution with zero mean and variance σ^2 given by formula (5.1) below.*

Corollary 5.2: *If F satisfies either the assumptions of Theorem 4.15 or the assumptions of Theorem 4.18 with $L_N \geq c\sqrt{N}$ and $d \geq 3$, then F satisfies the CLT.*

Proof. In the case of Theorem 4.15, this follows from the CLT for exponentially mixing systems ([11, 6]). In the case of Theorem 4.18, the result follows from Theorem 5.1. ∎

Proof of Theorem 5.1. By (4.25) with $n_1 = 0, n_2 = n$,

$$\sigma^2 := \sum_{n=-\infty}^{\infty} \zeta(H(H \circ F^n)) \tag{5.1}$$

exists and is finite. Hence

$$\begin{aligned}\zeta\Big(\frac{\Sigma_N^2(H)}{N}\Big) &= \frac{1}{N} \sum_{1 \leq i,j \leq N} \zeta((H \circ F^i)(H \circ F^j)) \\ &= \sum_{k=-N+1}^{N-1} \frac{N-|k|}{N} \zeta(H(H \circ F^k)) \to \sum_{n=-\infty}^{\infty} \zeta(H(H \circ F^n)).\end{aligned}$$

To finish our proof, we need to estimate the asymptotics of moments $\zeta(\Sigma_N^m(H))$, for any $m \geq 3$. Denote

$$\Omega(k_1, \dots, k_m) = \int_Z \Big(\prod_{i=1}^m H(F^{k_i} z)\Big) d\zeta(z)$$

so that

$$\zeta(\Sigma_N^m(H)) = \sum_{k_1,\dots,k_m=1}^{N} \Omega(k_1,\dots,k_m). \tag{5.2}$$

For the vector $(k_1, \ldots, k_m)$ we associate another vector $(n_1, \ldots, n_m)$ which is the permutation of the elements of $(k_1, \ldots, k_m)$ in increasing order, that is $n_1 \leq n_2 \leq \cdots \leq n_m$ Noting that Ω is symmetric, we have

$$\Omega(k_1, \ldots, k_m) = \Omega(n_1, \ldots, n_m).$$

We rewrite the above sum into two terms as $I_1 + I_2$, where I_1 is the sum of terms, whose social partition minimizing the RHS of (4.25) is not pairing (i.e., at least one atom contains more than two elements), and I_2 is the sum of terms, whose corresponding social partition is pairing. (If there are more than one partition minimizing κ, at least one of which is not pairing then we put the corresponding term into I_1.)

We need two auxiliary estimates. Let $Q = \{Q_1, \ldots, Q_r\}$ be a fixed social partition of the set $\{1, 2, \ldots, m\}$. We say that $\mathcal{Q}(n_1, \ldots, n_m) = Q$ if the partition $\mathfrak{P}$ minimizing the RHS of (4.25) for the given numbers $n_1, \ldots, n_m$ is of the form $\mathfrak{P} = \{P_1, \ldots, P_r\}$ with $\{i : n_i \in P_k\} = Q_k$ for all $k = 1, \ldots, r$. Next we write

$$I_Q = \sum_{k_1, \ldots, k_m : \mathcal{Q}(n_1, \ldots, n_m) = Q} \Omega(n_1, \ldots, n_m).$$

Lemma 5.3: (a) $I_Q = \mathrm{O}(N^r)$.

(b) *If $Q = Q_1 \cup \cdots \cup Q_r$ is not pairing, then the sum $I_Q = \mathrm{O}(N^{(m-1)/2})$.*

Proof. Since $1/\kappa_Q(n_1, \ldots, n_m) \leq 1/\kappa_Q^+(n_1, \ldots, n_m)$, by (4.25) it suffices to estimate

$$\sum_{n_1, \ldots, n_m} \frac{1}{(\kappa_Q^+(n_1, \ldots, n_m))^d}. \tag{5.3}$$

Let $n_1' < n_2' < \cdots < n_r'$ be the forward free elements among $\{n_1, \ldots, n_m\}$ and $n_1'', \ldots, n_{m-r}''$ be the forward fixed elements. For each fixed element n_j'', let $\bar{n}_j$ be the previous element in $\{n_1, \ldots, n_m\}$. Rewrite (5.3) as

$$\sum_{n_1', \ldots, n_r'} \left[\sum_{n_1'', \ldots, n_{m-r}''} \left(\frac{1}{\prod_{j=1}^{m-r} L_{n_j'' - \bar{n}_j}} \right)^d \right]. \tag{5.4}$$

Since L_n^{-d} is summable, the inner sum is uniformly bounded, so that (5.4) is bounded by N^r. This proves (a).

(b) follows from (a) because if Q is not pairing, then $r < \lfloor m/2 \rfloor$. ∎

Since there are finitely many partitions of $\{1,\dots,m\}$, Lemma 5.3 implies that $|I_1|$ is bounded above by $O(N^{(m-1)/2})$. In particular, for odd m,

$$\zeta(\Sigma_N^m(H)) = \mathrm{O}(N^{(m-1)/2}).$$

Now let m be even and Q be a pairing, that is $Q = \{Q_1,\dots,Q_{m/2}\}$ with all atoms Q_k containing exactly two numbers. By a forward (backward) step we mean $n_j - n_{j-1}$ where n_j is forward (backward) fixed in the partition $\mathcal{Q}(n_1,\dots,n_m)$. Let $\Gamma_Q(n_1,\dots,n_m)$ be a largest among all forward and backward steps in the partition Q and let $\Gamma(n_1,\dots,n_m) = \Gamma_{\mathcal{Q}(n_1,\dots,n_m)}(n_1,\dots,n_m)$.

Lemma 5.4: *For any $\epsilon > 0$, there exists $M > 0$ such that*

$$\left|\sum_{k_1,\dots,k_m:\Gamma(n_1,\dots,n_m)>M} \Omega(n_1,\dots,n_m)\right| \le N^{m/2}\epsilon.$$

Proof. It is enough to prove the lemma for Γ replaced by Γ^+ and also for Γ replaced by Γ^-, where Γ^+ is a largest among all forward steps and Γ^- is a largest among all forward steps. We only consider Γ^+ as Γ^- is similar. The proof for Γ^+ proceeds in the same way as the proof of Lemma 5.3 except we estimate the inner sum in (5.4) by

$$C\left(\sum_{n=1}^{\infty} L_n^{-d}\right)^{m-r-1}\left(\sum_{n=M}^{\infty} L_n^{-d}\right). \tag{5.5}$$

Indeed there are $m-r$ factors in the inner sum in (5.4), and by our assumptions one of them should be greater than M. As the second factor can be made as small as we wish by taking M large and since $r = m/2$, the result follows. ∎

Lemma 5.5: *Let Q be a pairing which is different from*

$$\bar{Q} := [(12),(34),\dots,((m-1)m)]. \tag{5.6}$$

Then the number of m-tuples $(k_1,\dots,k_m)$ with $\Gamma_{\bar{Q}}(n_1,n_2,\dots,n_m) < L$ is

$$\mathrm{O}(N^{(m/2)-1}),$$

where the implicit constant depends on L.

Proof. We claim that if $Q \ne \bar{Q}$, then the sets of forward fixed and backward fixed edges are different. If follows that if both $\Gamma^+_{\bar{Q}}(n_1,\dots,n_m) < M$ and $\Gamma^-_{\bar{Q}}(n_1,\dots,n_m) < M$, then there are at least $m/2+1$ edges which are shorter that M. The number of such tuples is $O(N^{(m/2)-1})$ and the result follows.

It remains to prove the claim. That is, we show that if the sets of forward fixed and backward fixed edges are the same, then $Q = \bar{Q}$. We proceed by induction. If $m = 0$ or 2 then there are no pairings different from $\bar{Q}$. Suppose $m > 2$. Then (n_{m-1}, n_m) is forward fixed, so it should be backward fixed, but this is only possible if $(m-1)$ is paired to m. Likewise (n_1, n_2) is backward fixed, hence it is forward fixed. But this is only possible if 1 is paired to 2. Removing $1, 2, (m-1)$ and m from Q we obtain a partition of $m-4$ elements for which the set of forward fixed and backward fixed edges coincide. By induction 3 is paired to 4, 5 to $6, \dots, (m-3)$ to $(m-2)$. The proof is complete. ∎

By the above lemmas, it suffices to consider indices $k_1, \dots, k_m$ so that

$$\forall i = 1, \dots, m/2 : \; M_i := n_{2i} - n_{2i-1} \le M$$

(5.7) $\qquad$ and

$$\forall i = 1, \dots, m/2 - 1 : \; n_{2i+1} - n_{2i} > L$$

for some large M and $L = L(M)$. Indeed, by choosing $M = M(\varepsilon)$ and $N > N_0$, $N_0 = N_0(L)$, the above lemmas give that the contribution of other terms is $< \varepsilon N^{m/2}$. Now we choose L so that for any fixed $M_1, \dots, M_{m/2}$ (finitely many choices), the RHS of (4.25) with $s = m/2$ and $H_j = H(H \circ T^{M_j})$ is less than ε. We conclude that

$$\left| \zeta(\Sigma_N^m) - \sum_{k_1, \dots, k_m \text{ satisfying (5.7)}} \prod_{i=1}^{m/2} \left(\int_Z (H(H \circ T^{M_i})) d\zeta(z) \right) \right| \le 2\varepsilon N^{m/2}.$$

Let us write

$$A_\ell = \int_Z (H(H \circ T^\ell)) d\zeta(z).$$

Now we claim that

$$\sum_{k_1, \dots, k_m \text{ satisfying (5.7)}} \prod_{i=1}^{m/2} A_{M_i} = (m-1)!! N^{m/2} (1 + o(1)) \left[\sum_{\ell=0}^{M} (A_\ell (1 + \mathbb{1}_{\ell > 0})) \right]^{m/2}.$$

To prove the claim, first note that

$$\sum_{M_1, \dots, M_{m/2} = 0}^{M} A_{M_1} \cdots A_{M_{m/2}} = \left(\sum_{\ell=0}^{M} A_\ell \right)^{m/2}.$$

Now it remains to count the number of tuples $(k_1, \dots, k_m)$ corresponding to the values $M_1, \dots, M_{m/2}$. Assume, for example, that $M_i > 0$ for all i. To count the number of possibilities, we first fix a pairing of indices $1, \dots, m$ which can be done in $(m-1)!!$ different ways. Then we have $\approx N^{m/2}$ choices to prescribe exactly one element of each pair. Let us say these values are $s_1 < s_2 < \cdots < s_{m/2}$. Except for an $o(N^{m/2})$ of these choices, we have $s_i - s_{i-1} > 2M + L$ and so for each remaining index k_j we have two choices: if it is paired to s_i, then either $k_j = s_i - M_i$ or $k_j = s_i + M_i$. Thus the total number of choices is

$$(m-1)!!2^{m/2}N^{m/2}(1+o(1)),$$

which verifies the claim for the case $M_i > 0$ for all i. If $M_i = 0$ for some i, then we only have one choice for the corresponding k_j and so we lose a factor of 2. The claim follows.

To finish the proof, notice that

$$\sum_{\ell=0}^{M} A_\ell(1+\mathbb{1}_{\ell>0}) = \sum_{\ell=-M}^{M} \zeta(H(H\circ F^\ell)) \to \sigma^2 \quad \text{as } M\to\infty.$$

Thus we have verified

$$\zeta(\Sigma_N^m(H)) = \begin{cases} o(N^{m/2}), & m \text{ is odd},\\ (m-1)!!N^{m/2}\sigma^m + o(N^{m/2}), & m \text{ is even}, \end{cases}$$

completing the proof of the theorem. ∎

Remark 5.6: The asymptotic variance given by (5.1) is typically non-zero. In particular, if either the drift is non-zero, or $d \geq 5$, then a direct calculation shows that

$$\lim_{N\to\infty} \zeta(\Sigma_N^2) - N\sigma^2 = -\sum_{n=-\infty}^{\infty} n\zeta(H(H\circ F^n))$$

(the convergence of the right hand side follows from the assumptions imposed above). Thus if $\sigma^2 = 0$ then $\zeta(\Sigma_N^2)$ is bounded, so by the L^2-Gotshalk–Hedlund Theorem H is an L_2 coboundary. It is an open question if the same conclusion holds if $\mu(\tau) = 0$ and d is 3 or 4. However, by assumption, f is exponentially mixing, so if H does not depend on y then $\sigma^2 > 0$ unless H is an L^2 coboundary. Thus in many (possibly all) cases σ^2 is a positive semidefinite quadratic form which is not identically equal to zero, and so its null set is a a linear subspace of positive (or infinite) codimension.

6. Mixing rates for ergodic fibers

6.1. Results.

Definition 6.1: We say that (f,τ) satisfies a **mixing averaged Edgeworth expansion of order** r if there are constants k_1, k_2 and a sequence $\delta_N \to 0$ so that for any function $\phi = \phi_N \in C^{k_2}(\mathbb{R}^d, \mathbb{R})$ supported on the box $J = J_N$, the expression

$$\mathcal{I}_{A_1,A_2,\phi}(N) := \mu(A_1(x)A_2(f^N x)\phi(\tau_N(x)))$$

satisfies

$$\left|\mathcal{I}_{A_1,A_2,\phi}(N) - N^{-d/2}\int_{s\in\mathbb{R}^d}\phi(s)\mathcal{E}_r^{A_1,A_2}(s/\sqrt{N})ds\right| \\ \leq \|A_1\|_{C^{k_1}}\|A_2\|_{C^{k_1}}\|\phi\|_{C^{k_2}}\mathrm{Vol}(J)\delta_N N^{-(d+r)/2},$$

where

$$\mathcal{E}_r(s) = \mathcal{E}_r^{A_1,A_2}(s) = \mathfrak{g}(s)\sum_{p=0}^{r}\frac{P_p^{A_1,A_2}(s)}{N^{p/2}},$$

with $\mathfrak{g}(\cdot)$ is a centered Gaussian density with positive-definite covariance matrix and $P_p(s)$ are polynomials in s whose coefficients are bilinear forms in (A_1, A_2), bounded in absolute value by $C\|A_1\|_{C^{k_1}}\|A_2\|_{C^{k_1}}$, and $P_0^{A_1,A_2}(s) = \mu(A_1)\mu(A_2)$.

Definition 6.2: We say that (f,τ) satisfies a **mixing averaged double Edgeworth expansion of order** r if there are constants k_1, k_2 and a sequence $\delta_N \to 0$ so that for any functions $\phi_i = \phi_i(N_i) \in C^{k_2}(\mathbb{R})$ supported on the interval $J_i = J_i(N_i)$ $(i = 1, 2)$, the expression

$$\mathcal{I}_{A_1,A_2,A_3,\phi_1,\phi_2}(N_1,N_2) := \mu(A_1(x)A_2(f^{N_1}(x))A_3(f^{N_2}(x))\phi_1(\tau_{N_1}(x))\phi_2(\tau_{N_2}(x)))$$

satisfies

$$\Bigg|\mathcal{I}_{A_1,A_2,A_3,\phi_1,\phi_2}(N_1,N_2) - \iint \phi_1(s_1)\mathfrak{g}\Big(\frac{s_1}{\sqrt{N_1}}\Big)\phi_2(s_2)\mathfrak{g}\Big(\frac{s_2-s_1}{\sqrt{N_2-N_1}}\Big)N_1^{-d/2}N_2^{-d/2} \\ \times\sum_{p_1,p_2=0}^{r}\frac{P_{p_1,p_2}^{A_1,A_2,A_3}(s_1/\sqrt{N_1},(s_2-s_1)/\sqrt{N_2-N_1})}{N_1^{\frac{p_1}{2}}(N_2-N_1)^{\frac{p_2}{2}}}ds_1 ds_2\Bigg| \\ \leq \Big(\prod_{j=1}^{3}\|A_j\|_{C^{k_1}}\Big)\Big(\prod_{i=1,2}\|\phi_i\|_{C^{k_2}}\mathrm{Vol}(J_i)\Big) \\ \times\delta_{\min\{N_1,N_2-N_1\}}(\max\{N_1,N_2-N_1\})^{-d/2}(\min\{N_1,N_2-N_1\})^{-(d+r)/2}$$

where $P^{A_1,A_2,A_3}_{p_1,p_2}(s_1,s_2)$ are polynomials in s_1, s_2 whose coefficients are bounded trilinear forms in (A_1, A_2, A_3), bounded in absolute value by

$$C\prod_{j=1}^{3}\|A_j\|_{C^{k_1}},$$

and

$$P^{A_1,A_2,A_3}_{0,0}(s) = \mu(A_1)\mu(A_2)\mu(A_3).$$

We will use the following hypotheses:

(A1) (f,τ) satisfies a mixing averaged Edgeworth expansion of order r_1;

(A1') (f,τ) satisfies a mixing averaged double Edgeworth expansion of order r_1;

(A2) for each $\delta > 0$, we have $\mu(|\tau_N| > N^{1/2+\delta}) = \mathrm{O}_\delta(N^{-r_2})$;

(A3) there are constants $\beta < 1$ and $k_3 \in \mathbb{R}^+$ such that if $B \in C^{k_3}(Y)$ has zero mean, then for any $T \in \mathbb{R}_+$,

$$S^B_T(y) := \int_{s\in[0,T]^d} B(G_s y)ds$$

satisfies

$$\nu(\max_{t\in\mathbb{R},|t|<T} |S^B_t| > T^{d\beta}) < \frac{C\|B\|_{C^{k_3}}}{T^{r_3}};$$

(A3') there exist constants $\beta < 1$, $k_3 \in \mathbb{R}^+$ so that if $B \in C^{k_3}(Y)$ has zero mean, then for any positive integer M there is some constant $C = C_M$ so that for any $T \in \mathbb{R}_+$,

$$\nu(y : |S^B_T| > T^{d\beta}) \le CT^{-M};$$

(A4) $\mu(A_1(x)A_2(f^N x)) - \mu(A_1)\mu(A_2) = \mathrm{O}(\|A_1\|_{C^{k_1}}\|A_2\|_{C^{k_1}}N^{-r_4})$.

Given $H, H_1, H_2 : X \times Y \to \mathbb{R}$ let

$$\rho_{H_1,H_2}(N) = \zeta(H_1(H_2 \circ F^N)) - \zeta(H_1)\zeta(H_2). \tag{6.1}$$

Theorem 6.3: *For $i = 1, 2, 3, 4$, assume* (A$\,i$) *with*

$$r_i > d(1-\beta) \tag{6.2}$$

(noting that r_1 is an integer). Then there exists K such that if $H_j \in C^K(X\times Y)$, then for any $\delta > 0$ there is some C_δ so that

$$|\rho_{H_1,H_2}(N)| \le C_\delta\|H_1\|_{C^K}\|H_2\|_{C^K}N^{d\frac{\beta-1}{2}+\delta}.$$

THEOREM 6.4: *Assume* (A1') *with*

$$r_1 \in \mathbb{N}, \quad r_1 > 2d(1-\beta) \tag{6.3}$$

and (A2), (A3'), (A4) *with* r_2, r_4 *satisfying* (6.2). *Then there exists* K *such that if* $H_j \in C^K(X \times Y)$, *then for any* $\delta > 0$ *there is some* C_δ *so that*

$$|\rho_{H_1,H_2}(N)| \le C_\delta \|H_1\|_{C^K} \|H_2\|_{C^K} N^{d(\beta-1)+\delta}. \tag{6.4}$$

The proofs of the above results use integrations by parts combined with various versions of (A1) and (A3). The exponents and the ideas of the proofs are similar to those appearing in [25], section 4.

Proof of Theorem 6.3. Case of $d = 1$. Let ψ be a C^∞ function such that $0 \le \psi(s) \le 1$, $\psi(0) = 0$ and $\psi(1) = 1$. Given $L > 0$, let

$$\psi_L(s) = \begin{cases} \psi(s+L+1) & \text{if } s \in [-L-1, -L], \\ 1 & \text{if } s \in (-L, L), \\ 1 - \psi(s-L) & \text{if } s \in [L, L+1], \\ 0 & \text{otherwise.} \end{cases}$$

By Corollary 3.4 and (A4), it suffices to consider the case

$$H_j(x,y) = A_j(x) B_j(y) \quad \text{where } \nu(B_j) = 0, \tag{6.5}$$

with $A_j, B_j \in C^{k_3}$. Without loss of generality we can assume $k_3 \ge k_2$, where k_2 is given by (A1).

Let $L = N^{1/2+\delta}$. Then

$$\begin{aligned} \rho_{H_1,H_2}(N) = & \iint A_1(x) A_2(f^N x) B_1(y) B_2(G_{\tau_N(x)} y) d\mu(x) d\nu(y) \\ = & \iint A_1(x) A_2(f^N x) B_1(y) B_2(G_{\tau_N(x)} y) \psi_L(\tau_N(x)) d\mu(x) d\nu(y) \qquad (6.6) \\ & + \iint A_1(x) A_2(f^N x) B_1(y) B_2(G_{\tau_N(x)} y)(1 - \psi_L(\tau_N(x))) d\mu(x) d\nu(y). \end{aligned}$$

The integrand in the last line is zero unless $|\tau_N(x)| \ge L$, so by (A2) the last line is

$$O(\|H_1\|_{C^0} \|H_2\|_{C^0} N^{-r_2})$$

and so we need only bound (6.6). First, observe that we can restrict the integral to $\bar{Y}$, the set of points where

$$|S_t^{B_2}(y)| < L^{d\beta} = L^\beta \quad \text{for } t \in [-L, L].$$

Indeed, by (A3), the integral over $Y \setminus \bar{Y}$ is in

$$\mathrm{O}(\|H_1\|_{C^0}\|H_2\|_{C^0}L^{-r_3}) \tag{6.7}$$

and so is negligible. Next observe that (6.6), restricted to $\bar{Y}$, is of the form

$$\int_{\bar{Y}} \mathcal{I}_{A_1,A_2,\phi_y}(N)d\nu(y) \quad \text{with } \phi_y(s) = B_1(y)B_2(G_s y)\psi_L(s).$$

Now by (6.2), $r_1 \geq 1$ and so by (A1), the above expression can be replaced by

$$N^{-1/2}\int_{\bar{Y}}\left(\int_{-\bar{L}}^{\bar{L}} \phi_y(s)\mathcal{E}_1(s/\sqrt{N})ds\right)d\nu(y)$$

with error

$$o(\|A_1\|_{C^{k_1}}\|B_1\|_{C^{k_0}}\|A_2\|_{C^{k_1}}\|B_2\|_{C^{k_2}}\bar{L}N^{-1}) = o(N^{\frac{\beta-1}{2}+\delta}), \tag{6.8}$$

where $\bar{L} = L+1$. Integrating by parts, we obtain

$$\begin{aligned}
&\int_{\bar{Y}}\left(\int_{-\bar{L}}^{\bar{L}} \phi_y(s)\mathcal{E}_1(s/\sqrt{N})\frac{ds}{\sqrt{N}}\right)d\nu(y)\\
&\quad = -\int_{\bar{Y}}\left(\int_{-\bar{L}}^{\bar{L}} \mathcal{E}_1'(s/\sqrt{N})\tilde{S}_y(s)\frac{ds}{N}\right)d\nu(y) + \mathrm{O}(\|H_1\|_{C^0}\|H_2\|_{C^0}L\mathfrak{g}(L/\sqrt{N})),
\end{aligned}$$

where $\tilde{S}_s(y) = B_1(y)\int_0^s \psi_L(u)B_2(G_u y)du$. Since

$$\tilde{S}_s 1_{|s|\leq L} = B_1(y)S_s^{B_2}(y)1_{|s|\leq L}$$

it follows from the definition of $\bar{Y}$ that the last integral is

$$\mathrm{O}(\|A_1\|_{C^{k_1}}\|B_1\|_{C^0}\|A_2\|_{C^{k_1}}\|B_2\|_{C^{k_3}}\frac{L^{1+\beta}}{N}).$$

This completes the proof of the theorem. ∎

Proof of Theorem 6.3. Case of $d \geq 2$. We follow the approach of the one-dimensional case. Let us assume (6.5) (the general case follows from Corollary 3.4). Now $\tau \in \mathbb{R}^d$ and so we define

$$\psi_L(s) = \prod_{j=1}^{d}\psi_L(s_j) \quad \text{for } s = (s_1,\ldots,s_d).$$

Let $\bar{Y}$ be defined as

$$\bar{Y} = \{y : |S_t^{B_2}(y)| < L^{d\beta} \quad \text{for } t \in [-L,L]\}.$$

Next we claim that

$$\rho_{H_1,H_2}(N) \approx N^{-d/2} \int_{\bar{Y}} \left(\int_{s\in[-\bar{L},\bar{L}]^d} \phi_y(s)\mathcal{E}_{r_1}(s/\sqrt{N})ds \right) d\nu(y),$$

where $a_N \approx b_N$ means

$$|a_n - b_N| = o(\|H_1\|_{C^{k_1}} \|H_2\|_{C^{k_3}} N^{d\frac{\beta-1}{2}+\varepsilon}).$$

Indeed, repeating the argument for $d = 1$, the error term (6.7) remains valid and the error term corresponding to (6.8) is $\mathrm{O}(\bar{L}^d N^{-(d+r_1)/2})$ which is in $o(N^{d(\beta-1)/2+\delta})$ by the assumption (6.2).

Performing d integrations by parts, one in each coordinate direction, we conclude that

$$\rho_{H_1,H_2}(N) \approx -N^{-d} \int_{\bar{Y}} \left(\int_{s\in[-\bar{L},\bar{L}]^d} \tilde{S}_s(y) \frac{\partial^d}{\partial s_1 \cdots \partial s_d} \mathcal{E}_{r_1}(s/\sqrt{N})ds \right) d\nu(y).$$

Now by the definition of $\bar{Y}$,

$$\rho_{H_1,H_2}(N) = \mathrm{O}(\|H_1\|_{C^{k_1}} \|H_2\|_{C^{k_3}} N^{-d} L^{d(1+\beta)}),$$

and the theorem follows. ∎

Proof of Theorem 6.4. Case of $d = 1$. Assume (6.5) (the general case follows from Corollary 3.4).

For fixed y, let us write

$$\sigma_N = \sigma_N(y) = \int H_1(F^N(x,y))H_2(F^{2N}(x,y))d\mu(x)$$

so that

$$\rho_{H_1,H_2}(N) = \zeta(H_1(H_2 \circ F^N)) = \int \sigma_N(y)d\nu(y).$$

We will prove that for any $\delta > 0$ and for any $y \in \bar{Y}$,

$$\sigma_N = o(N^{\beta-1+\delta}) \tag{6.9}$$

where $\bar{Y}$ (to be defined later) satisfies

$$\nu(\bar{Y}) > 1 - N^{-100} \tag{6.10}$$

(and so the contribution of its complement is negligible). As in the case of Theorem 6.3, the constant in the convergence in (6.9) can be bounded above by

$$C_\delta \|A_1\|_{C^{k_1}} \|A_2\|_{C^{k_1}} \|B_1\|_{C^{k_3}} \|B_2\|_{C^{k_3}}.$$

To simplify formulas, we do not indicate this dependence in the sequel.

Denote

$$Y_{L,\eta} = \{y \in Y : \exists t \in \mathbb{R} : |t| \in [L^\eta, L] : |S_t^B| > t^{\beta+\eta}\}.$$

Next we claim that for any $\eta > 0$ and for any M there is some C so that $\nu(Y_{L,\eta}) < CL^{-M}$. To prove this claim, observe that for $y \in Y_{L,\eta}$ there is some $t_* = t_*(y)$ with $|t_*| \in [L^\eta, L]$ and $|S_{t_*}^B(y)| > t_*^{\beta+\eta}$. Then

$$|S_{\lfloor t_* \rfloor}^B(y)| > \frac{1}{2} \lfloor t_* \rfloor^{\beta+\eta}$$

and so

$$Y_{L,\eta} \subset \bigcup_{k=\lfloor L^\eta \rfloor}^{\lceil L \rceil} Y_{L,\eta,k},$$

where

$$Y_{L,\eta,k} = \left\{ y \in Y : |S_k^B(y)| > \frac{1}{2} k^{\beta+\eta} \text{ or } |S_{-k}^B(y)| > \frac{1}{2} k^{\beta+\eta} \right\}.$$

Now we apply (A3') with M replaced by $(M+1)/\eta$ to conclude that

$$\nu(Y_{L,\eta,k}) < 2Ck^{-(M+1)/\eta} < CL^{-M-1}$$

for all $k \geq \lfloor L^\eta \rfloor$. The claim follows.

Next, define

$$\bar{Y} = Y \setminus \bigcup_{l=0,1,\ldots,\lfloor N \rfloor} G_l^{-1}(Y_{N^{1/2+\varepsilon},\delta/4})$$

with a small $\varepsilon = \varepsilon(\delta)$. By the previous claim, $\bar{Y}$ satisfies (6.10).

Denote $L_1 = N^{1/2+\varepsilon}$, $L_2 = 2N^{1/2+\varepsilon}$ and $\bar{L}_i = L_i + 1$. We start by computing

$$\begin{aligned}\sigma_N &= e_1 + \int A_1(f^N(x)) A_2(f^{2N}(x)) B_1(G_{\tau_N}(y)) B_2(G_{\tau_{2N}}(y)) \psi_{L_1}(\tau_N) \psi_{L_2}(\tau_{2N}) d\mu(x) \\ &= e_1 + \mathcal{I}_{1,A_1,A_2,\phi_{y,1},\phi_{y,2}}(N, 2N)\end{aligned}$$

where

$$\phi_{y,i}(s) = B_i(G_s(y)) \psi_{L_i}(s),$$

and the error term e_1 satisfies

$$|e_1| = \mathrm{O}(N^{-r_2}) = o(N^{\beta-1}) \tag{6.11}$$

by (A2).

Now using (A1'), we derive

$$\sigma_N = e_1 + e_2 + \sum_{p_1,p_2=0}^{r_1} \frac{1}{N^{\frac{p_1+p_2+2}{2}}} \mathcal{J},$$

where

$$\mathcal{J}=\int_{-\bar{L}_1}^{\bar{L}_1}\phi_{y,1}(s_1)\mathfrak{g}\Big(\frac{s_1}{\sqrt{N}}\Big)\int_{-\bar{L}_2}^{\bar{L}_2}\phi_{y,2}(s_2)\mathfrak{g}\Big(\frac{s_2-s_1}{\sqrt{N}}\Big)P^{1,A_1,A_2}_{p_1,p_2}\Big(\frac{s_1}{\sqrt{N}},\frac{s_2-s_1}{\sqrt{N}}\Big)ds_2ds_1,$$

and where by the error term in (A1') and by (6.3), e_2 satisfies

$$|e_2| = \mathrm{O}(\bar{L}_1\bar{L}_2N^{-1/2}N^{-(1+r_1)/2}) = \mathrm{O}(N^{2\varepsilon-r_1/2}) = o(N^{\beta-1+\delta}). \tag{6.12}$$

Next, we write the integral w.r.t. s_2 in $\mathcal{J}$ as

$$\mathcal{J}_1+\mathcal{J}_2=\int_{s_1-N^{1/2+\varepsilon}}^{s_1+N^{1/2+\varepsilon}}(\dots)ds_2+\int_{s_2\in[-\bar{L}_2,\bar{L}_2]\setminus[s_1-N^{1/2+\varepsilon},s_1+N^{1/2+\varepsilon}]}(\dots)ds_2.$$

The integrand in $\mathcal{J}_2$ is bounded by a polynomial term times $\mathfrak{g}(N^\varepsilon)$ and so $\mathcal{J}_2$ is negligible. Now let us write

$$\partial_2(P\mathfrak{g})(x,y)=\frac{\partial}{\partial y}(P(x,y)\mathfrak{g}(y)).$$

Then using integration by parts in $\mathcal{J}_1$ we conclude that

$$\sigma_N\approx-\sum_{p_1,p_2=0}^{r_1}\frac{1}{N^{\frac{p_1+p_2+3}{2}}}\int_{-\bar{L}_1}^{\bar{L}_1}\phi_{y,1}(s_1)\mathfrak{g}\Big(\frac{s_1}{\sqrt{N}}\Big)\mathcal{K}_{p_1,p_2}(s_1)ds_1, \tag{6.13}$$

where

$$\begin{aligned}\mathcal{K}(s_1)&=\mathcal{K}_{p_1,p_2}(s_1)\\&:=\int_{s_1-N^{1/2+\varepsilon}}^{s_1+N^{1/2+\varepsilon}}S^{B_2}_{s_2-s_1}(G_{s_1}y)\Big[\partial_2(P^{1,A_1,A_2}_{p_1,p_2}\mathfrak{g})\Big(\frac{s_1}{\sqrt{N}},\frac{s_2-s_1}{\sqrt{N}}\Big)\Big]ds_2\\&=\int_{-N^{1/2+\varepsilon}}^{N^{1/2+\varepsilon}}S^{B_2}_{u}(G_{s_1}y)\Big[\partial_2(P^{1,A_1,A_2}_{p_1,p_2}\mathfrak{g})\Big(\frac{s_1}{\sqrt{N}},\frac{u}{\sqrt{N}}\Big)\Big]du\end{aligned}$$

and $\approx$ means that the difference between the two sides is in $o(N^{\beta-1+\delta})$.

Using the fact that $y\in\bar{Y}$ and assuming that $\varepsilon=\varepsilon(\delta)$ is small enough, we have

$$\mathcal{K}_{p_1,p_2}(s_1)=\mathrm{O}(N^{\frac{1+\beta}{2}+\delta/2}) \tag{6.14}$$

for any p_1,p_2. If $p_1+p_2\geq 1$, then by (6.14) the term corresponding to p_1,p_2 in (6.13) is

$$\mathrm{O}(N^{-2}N^{1/2+\varepsilon}N^{\frac{1+\beta}{2}+\delta/2})=o(N^{\beta-1+\delta}).$$

Next, we claim that

$$\mathcal{K}'_{0,0}(s_1)=\mathrm{O}(N^{\frac{\beta}{2}+\delta/2}). \tag{6.15}$$

Note that by (A1'), $P^{1,A_1,A_2}_{0,0}(x,y) = \mu(A_1)\mu(A_2)$ and so

$$\begin{aligned}\mathcal{K}'_{0,0}(s_1) =&\mu(A_1)\mu(A_2)\int_{-N^{1/2+\varepsilon}}^{N^{1/2+\varepsilon}} \Big[\frac{\partial}{\partial s_1} S_u^{B_2}(G_{s_1}y)\Big]\mathfrak{g}'\Big(\frac{u}{\sqrt{N}}\Big)du\\ =&\mu(A_1)\mu(A_2)\int_{-N^{1/2+\varepsilon}}^{N^{1/2+\varepsilon}} B_2(G_{s_1+u}y)\mathfrak{g}'\Big(\frac{u}{\sqrt{N}}\Big)du\\ &-\mu(A_1)\mu(A_2)\int_{-N^{1/2+\varepsilon}}^{N^{1/2+\varepsilon}} B_2(G_{s_1}y)\mathfrak{g}'\Big(\frac{u}{\sqrt{N}}\Big)du.\end{aligned}$$

The integral in the penultimate line is $\mathrm{O}(N^{\frac{\beta}{2}+\delta/2})$ since we can perform one more integration by parts with respect to u. The integral in the last line is equal to

$$\sqrt{N}B_2(G_{s_1}y)[\mathfrak{g}(N^\varepsilon) - \mathfrak{g}(-N^\varepsilon)],$$

which decays rapidly (i.e., faster than any polynomial) in N and so is negligible. Thus we have verified (6.15).

Now we use (6.15) and an integration by parts with respect to s_1 to conclude that the term corresponding to $p_1 = p_2 = 0$ in (6.13) is

$$\approx N^{-3/2}\int_{-\bar{L}_1}^{\bar{L}_1} S_{s_1}^{B_1}(y)\frac{\partial}{\partial s_1}\Big(\mathfrak{g}\Big(\frac{s_1}{\sqrt{N}}\Big)\mathcal{K}_{0,0}(s_1)\Big)ds_1.$$

Now the definition of $\bar{Y}$ together with (6.14) and (6.15) imply that the last expression is $O(N^{\beta-1+\delta})$, which completes the proof of (6.9). ∎

We remark that the bound (6.15) can be derived in case $p_1 + p_2 \geq 1$ as well. This was not needed in case $d = 1$ but will be needed in case $d \geq 2$, which we discuss next.

Proof of Theorem 6.4. Case of $d \geq 2$. Assume (6.5) (the general case follows from Corollary 3.4).

We proceed as in the case of $d = 1$. That is, we need to show that

$$\sigma_N = o(N^{d(\beta-1)+\delta}) \tag{6.16}$$

for $y \in \bar{Y}$, where $\bar{Y}$ satisfies

$$\nu(\bar{Y}) > 1 - N^{-100d}. \tag{6.17}$$

First, we obtain $|e_1| = \mathrm{O}(N^{-r_2}) = o(N^{d(\beta-1)})$ as in (6.11). Similarly, (6.12) reads as

$$|e_2| = \mathrm{O}(\bar{L}_1^d\bar{L}_2^d N^{-d/2}N^{-(d+r_1)/2}) = \mathrm{O}(N^{d\varepsilon-r_1/2}) = o(N^{d(\beta-1)+\delta})$$

by (6.3) and by assuming that $\varepsilon = \varepsilon(\delta, d)$ is small. Next, we write

$$\bar{\partial}_2(P\mathfrak{g})(x,y) = \frac{\partial^d}{\partial y_1 \cdots \partial y_d}(P(x,y)\mathfrak{g}(y)).$$

Then as in (6.13), we derive

$$\sigma_N \approx -\sum_{p_1,p_2=0}^{r_1} N^{-\frac{p_1+p_2+3d}{2}} \mathcal{J}_{p_1,p_2}, \tag{6.18}$$

where $\approx$ means that the difference between the two sides is in $o(N^{d(\beta-1)+\delta})$ and

$$\mathcal{J}_{p_1,p_2} = \int_{s_1\in[-\bar{L}_1,\bar{L}_1]^d} \phi_y(s_1)\mathfrak{g}\Big(\frac{s_1}{\sqrt{N}}\Big)\mathcal{K}_{p_1,p_2}(s_1)ds_1,$$

where

$$\mathcal{K}_{p_1,p_2}(s_1) = \int_{u\in[-N^{1/2+\varepsilon},N^{1/2+\varepsilon}]^d} S_u^{B_2}(G_{s_1}y)\Big[\bar{\partial}_2(P^{1,A_1,A_2}_{p_1,p_2}\mathfrak{g})\Big(\frac{s_1}{\sqrt{N}}, \frac{u}{\sqrt{N}}\Big)\Big]du,$$

and for $u \in \mathbb{R}^d$,

$$S_u^B(\tilde{y}) = \int_{0\le v_i\le |u_i|} B(G_{v_1\,\mathrm{sgn}(u_1),\dots,v_d\,\mathrm{sgn}(u_d)}(\tilde{y}))dv_1\cdots dv_d$$

where sgn is the sign function ($\mathrm{sgn}(w) = -1$ if $w < 0$ and $\mathrm{sgn}(w) = 1$ if $w > 0$). For $I = \{i_1, \dots, i_{|I|}\} \subset \{1, 2, \dots, d\}$, let us write

$$\partial^I = \frac{\partial}{\partial s_{1,i_1}\cdots \partial s_{1,i_{|I|}}}, \quad \bar{\partial} = \partial^{\{1,\dots,d\}}.$$

We use d integrations by parts with respect to the variables $s_{11}, \dots, s_{1d}$ to write

$$\mathcal{J}_{p_1,p_2} = \int_{s_1\in[-\bar{L}_1,\bar{L}_1]^d} S_{s_1}^{B_1}(y)\bar{\partial}\Big[\mathfrak{g}\Big(\frac{s_1}{\sqrt{N}}\Big)\mathcal{K}_{p_1,p_2}(s_1)\Big]ds_1. \tag{6.19}$$

We will show that for any $I \subset \{1, \dots, d\}$ and for any p_1, p_2,

$$|\partial^I \mathcal{K}_{p_1,p_2}| \lesssim N^{\frac{d}{2}(\beta+1)-\frac{|I|}{2}} \tag{6.20}$$

where $a_N \lesssim b_N$ means that $a_N < b_N N^{\delta/2}$ (assuming that $\varepsilon = \varepsilon(\delta)$ is small enough). Assume first that (6.20) holds. Then observe that

$$\Big|\bar{\partial}\Big[\mathfrak{g}\Big(\frac{s_1}{\sqrt{N}}\Big)\mathcal{K}_{p_1,p_2}(s_1)\Big]\Big| \lesssim N^{\frac{d\beta}{2}}.$$

Substituting this estimate in (6.19), we obtain

$$|\mathcal{J}_{p_1,p_2}| \lesssim N^{d/2}N^{\frac{d\beta}{2}}N^{\frac{d\beta}{2}},$$

which implies (6.16). Thus it remains to prove (6.20).

Assume that $\mathfrak{g}$ is the standard Gaussian density (if this is not the case, we can compute all integrals on a parallelepiped of side length $cN^{1/2+\varepsilon}$, then apply a linear change of variables to reduce to the case of standard Gaussian). To prove (6.20) we write

$$h = \bar{\partial}_2(P^{1,A_1,A_2}_{p_1,p_2}\mathfrak{g}).$$

Recall that $I = \{i_1, \ldots, 1_{|I|}\}$, the set of indices i such that we are differentiating with respect to $s_{1,i}$, is given. We need to differentiate the integrand in $\mathcal{K}$, which is a product. Let $I' = \{i'_1, \ldots, i'_{|I'|}\} \subset I$ denote the set of indices i' so that we differentiate the term $S^{B_2}_u(G_{s_1}(y))$ with respect to $s_{1,i'}$. For $i \in I \setminus I'$, we differentiate h with respect to $s_{1,i}$. We also write

$$J = \{1, \ldots, d\} \setminus I \quad \text{and} \quad J' = \{1, \ldots, d\} \setminus I'.$$

Performing the differentiation, we find that

$$\begin{aligned}\partial^I \mathcal{K}_{p_1,p_2} = &\sum_{I':I'\subset I} \int_{u\in[-N^{1/2+\varepsilon},N^{1/2+\varepsilon}]^d} \int_{w_{j'}\in[0,|u_{j'}|] \text{ for } j'\in J'} \\ &\sum_{\delta_{i'}\in\{0,1\} \text{ for } i'\in I'} (-1)^{|I'|-\sum\delta_{i'}} B_2(G_{(i':s_{1i'}+\delta_{i'}u_{i'};j':s_{1j'}+w_{j'}\operatorname{sgn}(u_{j'}))}(y)) \\ &\times \Big[\partial^{I\setminus I'} h\Big(\frac{s_1}{\sqrt{N}}, \frac{u}{\sqrt{N}}\Big)\Big] dw_{j'}\, du,\end{aligned} \tag{6.21}$$

where in the subscript of G the notation $(i' : a_{i'}; j' : b_{j'})$ means that for coordinates $i' \in I'$ we use $a_{i'}$ and for $j' \in J'$ we use $b_{j'}$. Note that

$$\partial^{I\setminus I'} h\Big(\frac{s_1}{\sqrt{N}}, \frac{u}{\sqrt{N}}\Big) = N^{-\frac{|I|-|I'|}{2}} \tilde{h}\Big(\frac{s_1}{\sqrt{N}}, \frac{u}{\sqrt{N}}\Big), \tag{6.22}$$

where

$$\tilde{h}(x, y) = \frac{\partial^{|I|-|I'|}}{\partial x_{i'_1} \cdots \partial x_{i'_{|I'|}}} \frac{\partial^d}{\partial y_1 \cdots \partial y_d}(P(x, y)\mathfrak{g}(y)).$$

Now assume there is some i' so that $\delta_{i'} = 0$. Then $B_2(\ldots)$ does not depend on $u_{i'}$, and so performing the integral with respect to $u_{i'}$ first we obtain

$$\begin{aligned}&\int_{u_i\in[-N^{1/2+\varepsilon},N^{1/2+\varepsilon}]} \tilde{h}\Big(\frac{s_1}{\sqrt{N}}, \frac{u}{\sqrt{N}}\Big) du_i \\ &= \sqrt{N} \sum_{a=1,2} (-1)^a \tilde{h}_i\Big(\frac{s_1}{\sqrt{N}}, \Big(\frac{u_1}{\sqrt{N}}, \ldots, \frac{u_{i-1}}{\sqrt{N}}, (-1)^a N^{\varepsilon}, \frac{u_{i+1}}{\sqrt{N}}, \ldots, \frac{u_d}{\sqrt{N}}\Big)\Big),\end{aligned} \tag{6.23}$$

where

$$\tilde{h}_i(x, y) = \frac{\partial^{|I|-|I'|}}{\partial x_{i'_1} \cdots \partial x_{i'_{|I'|}}} \frac{\partial^{d-1}}{\partial y_1 \cdots \partial y_{i-1} \partial y_{i+1} \cdots \partial y_d}(P(x, y)\mathfrak{g}(y)).$$

Recalling that

$$\mathfrak{g}(y) = \frac{1}{(2\pi)^{d/2}} \exp\Big(- \sum_{i=1}^{d} y_i^2/2 \Big),$$

we see that $\tilde{h}_i(x, y)$ decays rapidly as $y_i \to \infty$ (i.e., faster than any polynomial). Since we have $|y_i| = N^{\varepsilon}$, (6.23) decays rapidly as $N \to \infty$. Thus this term, even when integrated with respect to all other variables, decays rapidly and consequently we can neglect all terms in (6.21) where there is some i' so that $\delta_{i'} = 0$.

It remains to study the case when $\delta_{i'} = 1$ for all $i' \in I'$. Then we perform the integrals in (6.21) with respect to $w_{j'}, j' \in J'$ and we integrate by parts with respect to $u_{i'}, i' \in I'$ to obtain that

$$|\partial^I \mathcal{K}_{0,0} - \mathcal{I}|$$

decays rapidly as $N \to \infty$, where

$$\mathcal{I} = \int_{u_{j'}, j' \in J'} \int_{u_{i'}, i' \in I'} S_b^{B_2}(y) \Big[\partial^I h\Big(\frac{s_1}{\sqrt{N}}, \frac{u}{\sqrt{N}} \Big) \Big] du_{i'} du_{j'}$$

and

$$b = (i' : N^{1/2+\varepsilon}, j' : u_{j'}).$$

As in (6.22), we have

$$\partial^I h\Big(\frac{s_1}{\sqrt{N}}, \frac{u}{\sqrt{N}} \Big) = N^{-\frac{|I|}{2}} \hat{h}\Big(\frac{s_1}{\sqrt{N}}, \frac{u}{\sqrt{N}} \Big) \tag{6.24}$$

where

$$\hat{h}(x, y) = \frac{\partial^{|I|}}{\partial x_{i_1} \cdots \partial x_{i_{|I|}}} \frac{\partial^d}{\partial y_1 \cdots \partial y_d} (P(x, y)\mathfrak{g}(y)).$$

Note that we can assume $|S_b^{B_2}| \lesssim N^{d\beta/2}$. Indeed, we can subdivide the rectangular box with opposite corners 0 and b into small cubes of side length N^{ε} and we can assume that the integral of $G_s(y)$ over all of the boxes is smaller than $N^{d\varepsilon\beta}$ for $y \in \bar{Y}$ by (A3') ($\bar{Y}$ satisfies (6.17) similarly to the case $d = 1$). Combining this observation with (6.24), we conclude that

$$|\mathcal{I}| \leq N^{\frac{d\beta - |I|}{2}} \int_{u \in [-N^{1/2+\varepsilon}, N^{1/2+\varepsilon}]} \|\hat{h}\|_{\infty} du \leq C N^{\frac{d(\beta+1)-|I|}{2} + \delta/2}$$

if $\varepsilon(\delta)$ is small enough. This completes the proof of (6.20) and so the theorem follows. ∎

7. Toral translations and related systems

7.1. Rapid mixing. Let f be an Axiom A diffeomorphism and μ be a Gibbs measure with Hölder potential. Let $Y = \mathbb{T}^m$ and G_t be a d-parameter flow:

$$G_{(t_1,\dots,t_d)}(y) = y + \sum_{j=1}^{d} \alpha_j t_j$$

for some $\alpha_1, \dots, \alpha_d \in \mathbb{R}^m$. Note that G_t has discrete spectrum, so it is far from being mixing. However, according to [21] the mixing properties of the corresponding skew products are typically much better than the results obtained in Section 4 for the case of the mixing fibers. Namely, let Π be the linear subspace generated by $\alpha_1, \dots, \alpha_d$. We say that Π is **Diophantine** if there exist numbers K, s such that for any unit vector $v \in \Pi$ for any $k \in \mathbb{Z}^m$ we have

$$|\langle v, k\rangle| \geq K|k|^{-s}.$$

Proposition 7.1 ([21]): *If Π is Diophantine, then F is rapidly mixing except for the set $\tau : X \to \Pi$ lying in an infinite codimension submanifold.*

Next, we describe an application of this result.

7.2. Constant suspensions in the fiber. Again we take f as in §7.1, but now we consider constant suspensions acting in the fiber. That is, let $\mathcal{G}^{\mathbf{n}}$ be a $\mathbb{Z}^d$ exponentially mixing action on a manifold $\mathcal{Y}$ preserving a measure $\tilde{\nu}$, let $Y = \mathcal{Y} \times \mathbb{R}^d / \sim$ where $\sim$ is the identification

$$(\tilde{y}, z + \mathbf{n}) \sim (\mathcal{G}^{\mathbf{n}} \tilde{y}, z).$$

Let G^t be the action $(\tilde{y}, z) \to (\tilde{y}, z + t)$. It preserves measure $d\nu = d\tilde{\nu} dz$.

Given a T, T^{-1} map as above, consider an associated action $\mathcal{F}$ on $X \times \mathbb{T}^d$ given by

$$\mathcal{F}(x, \theta) = (fx, \theta + \tau(x)).$$

Proposition 7.2: *Suppose that $\mathcal{F}$ is rapidly mixing. Then* (4.5) *holds.*

Proof. Split $H = \bar{H} + \tilde{H}$ where $\bar{H}(x, z) = \int H(x, \tilde{y}, z) d\tilde{\nu}(\tilde{y})$. Note that G_t, and hence F, preserves this splitting and that $\bar{H}$ is $\mathbb{Z}^d$ invariant, because $\mathcal{G}^{\mathbf{n}}$ preserves $\tilde{\nu}$ and

$$\int H(x, \tilde{y}, z + \mathbf{n}) d\tilde{\nu}(\tilde{y}) = \int H(x, \mathcal{G}^{\mathbf{n}} \tilde{y}, z) d\tilde{\nu}(\tilde{y}) = \bar{H}(x, z).$$

It follows that

$$\rho_{H_1,H_2}(n) = \rho_{\bar{H}_1,\bar{H}_2}(n) + \rho_{\tilde{H}_1,\tilde{H}_2}(n).$$

The first term decays faster than any polynomial, because $\mathcal{F}$ is rapidly mixing and the second term is $O(n^{-d/2})$ due to Remark 4.10. However, to apply the remark, we need to check that G_t is exponentially mixing on the space $\mathbb{B}$ of C^L functions such that

$$\int H(x,(\tilde{y},z))d\tilde{\nu}(y) = 0 \quad \text{for all } (x,z).$$

To check mixing, we write $t = \mathbf{n} + \hat{t}$, where $\mathbf{n} \in \mathbb{Z}^d$ and $\hat{t}$ belongs to the unit cube. Then

$$\int H_1(x_1,(\tilde{y},z))H_2(x_2,G_t(\tilde{y},z))d\nu = \iint H(x_1,(\tilde{y},z-\hat{t}))H_2(x_2,(\mathcal{G}^n\tilde{y},z))d\tilde{\nu}(\tilde{y})dz.$$

Integrating first with respect to $\tilde{y}$, we see that the RHS decays exponentially as needed. ∎

8. Deviations of ergodic averages

8.1. Mixing and deviations. Here we recall some results about the relations of mixing and deviations of ergodic averages.

Lemma 8.1: *Let $X_1, X_2, \ldots$ be a stationary sequence of random variables on a probability space (Ω, P) and $S_N = \sum_{k=1}^N X_k$. Assume that there are constants C and ρ such that for every n*

$$E(S_n^2) < Cn^{2\rho}. \tag{8.1}$$

Then $S_n/n^{\max\{\rho,\frac{1}{2}\}+\varepsilon}$ converges to zero almost surely for all $\varepsilon > 0$.

Proof. Let us assume $\rho > 1/2$ (the case $\rho \le 1/2$ is a simple consequence). For a positive integer m, let D_m denote the collection of intervals of the form $I_{i,j} = [j2^i + 1, (j+1)2^i]$ for all non-negative integers i, j so that $(j+1)2^i \le 2^m$. By the stationarity assumption,

$$E\Bigg(\sum_{I\in D_m}\bigg(\sum_{k\in I} X_n\bigg)^2\Bigg) \le \sum_{i=0}^m 2^{m-i}E(S_{2^i}^2) \le C\sum_{i=0}^m 2^{m-i}2^{2i\rho} \le \tilde{C}2^{2m\rho}.$$

Now for a given positive integer n, let m be so that $2^{m-1} < n \le 2^m$. Then the interval $[1,n]$ can be written as a disjoint union of at most $2m$ intervals from the family D_m. Let us denote this collection of intervals by $D(n)$. Then by the Cauchy–Schwartz inequality,

$$S_n^2 = \left(\sum_{I\in D(n)}\sum_{k\in I} X_k\right)^2 \le 2m\sum_{I\in D(n)}\left(\sum_{k\in I} X_k\right)^2 \le 2m\sum_{I\in D_m}\left(\sum_{k\in I} X_k\right)^2.$$

Thus we have

$$\begin{aligned} &P(\exists n = 2^{m-1}+1,\dots,2^m : S_n^2 > \eta n^{2\rho+\varepsilon}) \\ &\qquad \le P\left(2m\sum_{I\in D_m}\left(\sum_{k\in I} X_k\right)^2 > \eta 2^{(m-1)(2\rho+\varepsilon)}\right) \\ &\qquad \le 2m\eta^{-1}2^{-(m-1)(2\rho+\varepsilon)}E\left(\sum_{I\in D_m}\left(\sum_{k\in I} X_k\right)^2\right) \\ &\qquad \le \tilde{C}\eta^{-1}m2^{-m\varepsilon}. \end{aligned}$$

Using the Borel–Cantelli lemma and the fact that $\eta > 0$ is arbitrary, Lemma 8.1 follows. ∎

Lemma 8.2: *Under the assumptions of Lemma 8.1 suppose that*

$$|E(X_iX_j)| \le C|i-j|^{-\beta}.$$

Then (8.1) *is satisfied with*

$$\rho = \begin{cases} \frac{1}{2}, & \text{if } \beta > 1, \\ 1-\frac{\beta}{2} & \text{if } \beta < 1. \end{cases}$$

Proof. (8.1) follows since $E(S_N^2) = NE(X_0^2) + 2\sum_{n=0}^{N-1}(N-n)E(X_0X_n)$. ∎

8.2. Examples and open questions. Here we describe several classes of systems satisfying our assumptions on the base and the fiber dynamics made in previous sections. We also present several open questions pertaining to establishing those properties in several new cases.

Mixing of the base system is required in all our results. In addition, the results of Section 4 require mixing in the fiber, so we begin with reviewing known results for mixing.

Exponential mixing is known in the following cases: uniformly hyperbolic diffeomorphisms with Gibbs measures ([9, 54]); nonuniformly hyperbolic systems admitting Young towers with exponential tails ([61]); partially hyperbolic translations on homogeneous spaces ([47, 5]); contact Anosov flows [49] as well as Anosov flows with suitable assumptions on the Lyapunov spectrum [1, 60]; some singular hyperbolic flows [2]; ergodic automorphisms of tori [45] and of nilmanifolds ([38]). In all the examples of $\mathbb{R}$ or $\mathbb{Z}$ actions listed above, we also have multiple exponential mixing (see, e.g., [22]) while in higher rank the multiple exponential mixing is only known for partially hyperbolic translations on homogeneous spaces ([5]) (partial results for some $\mathbb{Z}^d$ actions are obtained in [39]).

Rapid mixing is known for generic Axiom A flows with Gibbs measures ([19, 20, 32]), hyperbolic flows having Young towers with exponential tails (see [52] and references therein), some singular hyperbolic flows [3], and generic compact group extensions of uniformly hyperbolic systems ([21]).

Polynomial mixing is known for nonuniformly hyperbolic diffeomorphisms and flows having Young towers with polynomial tails ([58, 40, 4]), unipotent actions ([47, 5], time changes of nilflows ([37]), and some flows on surfaces with degenerate singularities ([30]).

Additional assumptions imposed on base dynamics in various results include large deviations, anticoncentration, LLT and Edgeworth expansions.

The easiest way to get large deviation is to have unique ergodicity, since in that case the set in LHS of (3.2) is empty. A relative version of unique ergodicity is the so-called **Uunique ergodicity** (see [22] for a definition), which holds for partially hyperbolic systems with unique measure absolutely continuous with respect to the unstable foliation. In this case (3.2) holds due to [22]. Exponential large deviations also hold for non-uniformly hyperbolic systems admitting Young towers with exponential tails for return times [53, 56], while in case the tail is polynomial, polynomial large deviations hold [51, 41] (see also [26] where the large deviations are discussed under a quasiindependence assumption).

Anticoncentration inequality is established for systems admitting Young towers provided that the return time tail has second moment [55].

The LLT is known for Axiom A diffeomorphisms with Gibbs measures ([54]), the systems admitting Young towers under the assumptions that the tails admit the second moment ([59]) as well as flows which can be represented as suspensions of flows admitting nice symbolic dynamics [27] including Axiom A flows and certain Lorenz type attractors. The results of [27] can be applied to continuous time T, T^{-1} systems given by (4.15).

Mixing averaged Edgeworth expansions are obtained in [31] for systems admitting Young towers with exponential tails. It seems that the methods of [31] as well as [28] could be used to obtain the multiple expansions as well, but this remains an open problem.

For fiber dynamics we require control on ergodic averages. For mixing systems such control can be obtain using moment estimates (cf. Lemma 8.1).

Systems satisfying assumption (A3) (or (A3')) for $d = 1$ include exponentially mixing systems described above, as well as toral translations (see ,e.g., [24]), products of the last two examples [13], horocycle flows [33], translation flows (those flows are not smooth, however, the results of Section 6 apply provided that we consider the observables which vanish near the singularities), typical area preserving flows on surfaces (with non-degenerate singularities) [35] and nilflows ([34], [36]). Higher dimensional examples include Cartan and unipotent actions on homogeneous spaces of semisimple Lie groups ([5]) and multidimensional niltranslations [15].

The results of this paper motivate the study of the statistical properties discussed above for a wider class of dynamical systems. In particular, it is of interest to

(a) construct an example of systems satisfying mixing multiple Edgeworth expansion;
(b) prove mixing LLTs for partially hyperbolic systems;
(c) investigate mixing LLTs and anticoncentration bounds for parabolic systems.

8.3. Deviations of ergodic averages for generalized T, T^{-1} transformations. Here we illustrate the information that the results obtained in this paper provide about the growth of ergodic sums in several special cases. In the examples below we assume that the base dynamics f is given by an Anosov diffeomorphism equipped with a Gibbs measure and for each fiber flow (1–10) we give an exponent α such that with probability one the ergodic sums of the corresponding generalized T, T^{-1} transformation grow slower than $N^{\alpha+\varepsilon}$ for every $\varepsilon > 0$. This is going to be a simple consequence of Lemmas 8.1 and 8.2. For each example we list the result that implies the assumption of Lemma 8.2 with a suitable β. In case we use the results of Section 6, we also assume that (f, τ) satisfies the mixing double averaged Edgeworth expansion of any order.

Currently no examples of such systems are known but we expect this property to hold for a large class of map (cf., e.g., the computations in [28]).

(1) Anosov diffeomorphisms. In this case we have exponential mixing ([9, 54]);
   (a) zero drift: $\alpha = 3/4$ (Theorem 4.7);
   (b) positive drift: $\alpha = \frac{1}{2}$ (Theorem 4.1).
(2) Diophantine toral translations—here (A3') holds for any $\beta > 0$ and so $\alpha = 1/2$ by Theorem 6.4 (cf. also Proposition 7.1).
(3) Product of Anosov diffeomorphisms and toral translation: $\alpha = 3/4$ (Theorem 6.4).
(4) Horocycle flows (see [33]): Theorem 6.4 gives
   (a) no small eigenvalues of Δ, zero drift—(A3) holds for any $\beta > 1/2$, so $\alpha = \rho_1(\beta) = 3/4$;
   (b) smallest eigenvalue of Δ is $\lambda \in (0, \frac{1}{4})$—(A3) holds for any $\beta > \frac{1+\sqrt{1-4\lambda}}{2}$, so
$$\alpha = \rho_1(\beta) = \frac{1+\sqrt{1-4\lambda}}{2}.$$
(5) Translations flows—(A3') holds for any $\beta > \lambda_2$ ([35]) where λ_2 is the second exponent of the Kontsevich–Zorich cocycle. So
$$\alpha = \rho_1(\beta) = \frac{\lambda_2 + 1}{2}$$
(Theorem 6.4).
(6) Partially hyperbolic translations on homogenous spaces. In this case we have exponential mixing ([47, 5]);
   (a) zero drift: $\alpha = 3/4$ (Theorem 4.7);
   (b) positive drift: $\alpha = \frac{1}{2}$ (Theorem 4.1).
(7) Multidimensional Cartan actions on homogenous spaces: $\frac{1}{2}$ (Theorems 4.7 and 4.1).
(8) Constant suspensions of Cartan actions on tori: $\frac{1}{2}$ (Proposition 7.2).
(9) Continuous time T, T^{-1} system given by (4.15) with both base flow ϕ^t and fiber flow G_t given by geodesic flow on a unit tangent bundle over a negatively curve manifold: $\alpha = \frac{7}{8}$ by Example 4.12(b) with $k = 2$. In fact, Example 4.12(b) shows that for all positive integers k, we can obtain a system with $\alpha = 1 - 2^{-k-1}$.
(10) Generic higher rank actions on Heisenberg nilmanifolds: $\frac{1}{2}$ ([15] and Theorem 6.4).

Appendix A. Anticoncentration large deviation bounds for subshifts of finite type

We follow the argument in [26].

Let (Σ, σ) be a subshift of finite type, μ be a Gibbs measure and $\tau : \Sigma \to \mathbb{R}^d$ be a Hölder function of zero mean. We assume that for each $\mathbf{a} \in \mathbb{R}^d \backslash \{\mathbf{0}\}$ the function $\langle \mathbf{a}, \tau \rangle$ is not a coboundary.

Lemma A.1 ([54]): *There are constants c_1, δ_0 such that for $|\xi| < \delta_0$,*

$$\mu(e^{\langle \xi, \tau_N \rangle}) \leq e^{c_1 N \xi^2}; \tag{A.1}$$

$$|\Phi_N(\xi)| \leq e^{-c_1 N \xi^2}, \quad \textit{where } \Phi_N(\xi) = \mu(e^{i\langle \xi, \tau_N \rangle}). \tag{A.2}$$

Corollary A.2: *There are constants C_2, c_2 such that*

$$\mu(|\tau_N| > L) \leq C_2 e^{-c_2 L^2/N}, \tag{A.3}$$

and for each unit cube $\mathcal{Q}$

$$\mu(\tau_N \in \mathcal{Q}) \leq \frac{C_2}{N^{d/2}}. \tag{A.4}$$

Proof. To prove the first inequality we may assume without loss of generality that $d = 1$ and that $\sqrt{N} \leq L \leq 2c_1\delta_0 N$ (we obtain the general result by increasing C_2 and decreasing c_2). We estimate $\mu(\tau_N > L)$, the bound for $\mu(\tau_N < -L)$ being similar. We have that for each $\xi \in (0, \delta_0)$

$$\mu(\tau_N > L) = \mu(e^{\xi\tau_N} > e^{\xi L}) \leq e^{-\xi L}\mu(e^{\xi\tau_N}) \leq e^{-\xi L + c_1 N \xi^2}.$$

Taking $\xi = \frac{L}{2c_1 N}$ we obtain the result.

It is enough to prove (A.4) for cubes of any fixed size ρ since the unit cube can be covered by a finite number of cubes of size ρ. Let

$$g(x) = \prod_{l=1}^{d} \left(\frac{1 - \cos(\hat{\delta} x_{(l)})}{\hat{\delta}^2 x_{(l)}^2} \right),$$

where $\hat{\delta} = \delta_0/d$ and δ_0 is the constant from Lemma A.1. Then

$$\hat{g}(\xi) = (\pi\hat{\delta})^d \prod_{l=1}^{d} \left(\left(1 - \frac{|\xi|}{\hat{\delta}}\right) 1_{|\xi| \leq \hat{\delta}} \right).$$

Hence for each a

$$\mathbb{E}(g(\tau_N - a)) = \int_{\mathbb{R}^d} \hat{g}(-\xi) e^{i\xi a} \Phi_N(\xi) d\xi \leq \int_{|s| < \delta_0} \hat{g}(s) |\Phi_N(s)| ds$$

since $\hat{g}$ is real, positive, and supported inside the ball of radius δ_0. Thus (A.2) implies that there is a constant $\hat{D}$ such that

$$\mathbb{E}(g(\tau_N - a)) \leq \frac{\hat{D}}{N^{d/2}}.$$

On the other hand, $g(0) = \frac{1}{2^d}$ so there is a constant ρ such that $g(x) > \frac{1}{4^d}$ on the cube of size ρ centered at 0. Hence if $\mathcal{Q}$ is a cube of size ρ centered at a, then

$$\mathbb{E}(g(\tau_N - a)) \geq \frac{\mathbb{P}(S_N \in \mathcal{Q})}{4^d}.$$

Combining the last two displays we obtain the result. ∎

We now prove the anticoncentration large deviation estimate with

$$\Theta(r) = e^{-c_4 r^2}.$$

Lemma A.3: *If $\mathcal{Q}$ is a unit cube centered at z, then*

$$\mu(\tau_N \in \mathcal{Q}) \leq \frac{C_3}{N^{d/2}} e^{-c_3 z^2/N}.$$

Proof. There is a constant R such that

$$\mu(\tau_N \in \mathcal{Q}) \leq \mu\Big(\tau_N \in \mathcal{Q}, |\tau_{N/2}| > \frac{|z|}{2} - R\Big) + \mu\Big(\tau_N \in \mathcal{Q}, |\tau_N - \tau_{N/2}| > \frac{|z|}{2} - R\Big).$$

We will estimate the first term; the estimate of the second is obtained by replacing σ by σ^{-1}. We have

$$\mu\Big(\tau_N \in \mathcal{Q}, |\tau_{N/2}| > \frac{|z|}{2} - R\Big) \leq \sum_{\mathcal{C}', \mathcal{C}''} \mu(\mathcal{C}'\mathcal{C}''),$$

where the sum is over all pairs of cylinders $(\mathcal{C}', \mathcal{C}'')$ such that

(i) $\text{length}(\mathcal{C}') = \text{length}(\mathcal{C}'') = N/2$,
(ii) there exists $\omega' \in \mathcal{C}'$ such that $|\tau_{N/2}(\omega')| > \frac{|z|}{2} - R$,
(iii) there exists $\omega'' \in \mathcal{C}''$ such that $|\tau_{N/2}(\omega') + \tau_{N/2}(\omega'') - z| < 2R$.

By the Gibbs property

$$\sum_{\mathcal{C}', \mathcal{C}''} \mu(\mathcal{C}'\mathcal{C}'') \leq K \sum_{\mathcal{C}', \mathcal{C}''} \mu(\mathcal{C}')\mu(\mathcal{C}'').$$

By (A.4), for each $\mathcal{C}'$ the sum of $\mu(\mathcal{C}'')$ over the cylinders $\mathcal{C}''$ satisfying (iii) is smaller than $(2R)^d C_2/N^{d/2}$. Summing over $\mathcal{C}'$ satisfying (ii) and using (A.3), we obtain the result. ∎

LEMMA A.4: *Let $\mathcal{Q}_1, \dots, \mathcal{Q}_s$ be unit cubes centered at $z_1, \dots, z_s$. Then with the notation $z_0 = 0 \in \mathbb{R}^d$, $n_0 = 0$,*

$$\mu(\tau_{n_j} \in \mathcal{Q}_j \text{ for } j = 1, \dots, s) \le \prod_{j=1}^{s} \Big[\Big(\frac{C_4}{(n_j - n_{j-1})^{d/2}} \Big) e^{-c_4 \frac{|z_j - z_{j-1}|^2}{n_j - n_{j-1}}} \Big].$$

Proof. The LHS can be bounded by $\sum(\mu(\mathcal{C}_1\mathcal{C}_2 \cdots \mathcal{C}_s))$, where the sum is over all tuples of cylinders such that

(i) $\text{length}(\mathcal{C}_j) = n_j - n_{j-1}$, and
(ii) on $\mathcal{C}_j$, $\tau_{n_j - n_{j-1}}$ is contained in a cube of size R centered at $z_j - z_{j-1}$.

Using the Gibbs property the last can be bounded by

$$K \prod_{j=1}^{s} \Big[\sum_{\mathcal{C}_j:\ \text{(i) and (ii) hold}} \mu(\mathcal{C}_j) \Big].$$

Now the result follows by Lemma A.3. ∎

References

[1] V. Araujo, O. Butterley and P. Varandas, *Open sets of axiom A flows with exponentially mixing attractors*, Proceedings of the American Mathematical Society **144** (2016) 2971–2984.

[2] V. Araujo and I. Melbourne, *Exponential decay of correlations for nonuniformly hyperbolic flows with a $C^{1+\alpha}$ stable foliation, including the classical Lorenz attractor*, Annales Henri Poincaré **17** (2016), 2975–3004.

[3] V. Araujo and I. Melbourne, *Mixing properties and statistical limit theorems for singular hyperbolic flows without a smooth stable foliation*, Advances in Mathematics **349** (2019), 212–245.

[4] P. Bálint, O. Butterley and I. Melbourne, *Polynomial decay of correlations for flows, including Lorentz gas examples*, Communications in Mathematical Physics **368** (2019), 55–111.

[5] M. Björklund, M. Einsiedler and A. Gorodnik, *Quantitative multiple mixing*, Journal of the European Mathematical Society **22** (2020), 1475–1529.

[6] M. Björklund and A. Gorodnik, *Central limit theorems for group actions which are exponentially mixing of all orders*, Journal d'Analyse Mathématiques **141** (2020), 457–482.

[7] E. Bolthausen, *A central limit theorem for two-dimensional random walks in random sceneries*, Annals of Probability **17** (1989), 108–115.

[8] C. Bonatti, L. J. Diaz and M. Viana, *Dynamics Beyond Uniform Hyperbolicity*, Encyclopaedia of Mathematical Sciences, Vol. 102, Springer, Berlin, 2005.

[9] R. Bowen, *Equilibrium States and Ergodic Theory of Anosov Diffeomorphisms*, Lecture Notes in Mathematics, Vol. 470, Springer, New York, 1975.

[10] E. Breuillard, *Distributions diophantiennes et theoreme limite local sur* $\mathbb{R}^d$, Probability Theory and Related Fields **132** (2005), 39–73.

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

[12] N. Chernov and R. Markarian, *Chaotic Billiards*, Mathematical Surveys and Monographs, Vol. 127, American Mathematical Society, Providence, RI, 2006.

[13] G. Cohen and J.-P. Conze, *The CLT for rotated ergodic sums and related processes*, Discrete and Continuous Dynamical Systems. Series A **33** (2013), 3981–4002.

[14] G. Cohen and J.-P. Conze, *CLT for random walks of commuting endomorphisms on compact abelian groups*, Journal of Theoretical Probability **30** (2017), 143–195.

[15] S. Cosentino and L. Flaminio, *Equidistribution for higher-rank abelian actions on Heisenberg nilmanifolds*, Journal of Modern Dynamics **9** (2015), 305–353.

[16] F. den Hollander, M. S. Keane, J. Serafin and J. E. Steif, *Weak Bernoullicity of random walk in random scenery*, Japanese Journal of Mathematics **29** (2003), 389–406.

[17] F. den Hollander and J. E. Steif, *Mixing properties of the generalized* T, T^{-1}*-process*, Journal d'Analyse Mathématique **72** (1997), 165–202.

[18] F. den Hollander and J. E. Steif, *Random walk in random scenery: a survey of some recent results*, in *Dynamics & stochastics*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, Vol. 48, Institute of Mathematical Statistics, Beachwood, OH, 2006, pp. 53–65.

[19] D. Dolgopyat, *On decay of correlations in Anosov flows*, Annals of Mathematics **147** (1998), 357–390.

[20] D. Dolgopyat, *Prevalence of rapid mixing in hyperbolic flows*, Ergodic Theory and Dynamical Systems **18** (1998), 1097–1114.

[21] D. Dolgopyat, *On mixing properties of compact group extensions of hyperbolic systems*, Israel Journal of Mathematics **130** (2002), 157–205.

[22] D. Dolgopyat, *Limit theorems for partially hyperbolic systems*, Transactions of the American Mathematical Society **356** (2004), 1637–1689.

[23] D. Dolgopyat, C. Dong, A. Kanigowski and P. Nándori, *Flexibility of statistical properties for smooth systems satisfying the central limit theorem*, `https://arxiv.org/abs/2006.02191`.

[24] D. Dolgopyat and B. Fayad, *Limit theorems for toral translations*, in *Hyperbolic Dynamics, Fluctuations and Large Deviations*, Proceedings of Symposia in Pure Mathematics, Vol. 89, American Mathematical Society, Providence, RI, 2015, pp. 227–277.

[25] D. Dolgopyat, M. Lenci and P. Nándori, *Global observables for random walks: law of large numbers*, Annales de l'Institut Henri Poincaré Probabilités et Statistiques **57** (2021), 94–115.

[26] D. Dolgopyat and P. Nándori, *Infinite measure renewal theorem and related results*, Bulletin of the London Mathematical Society **51** (2019), 145–167.

[27] D. Dolgopyat and P. Nándori, *On mixing and the local central limit theorem for hyperbolic flows*, Ergodic Theory and Dynamical Systems **40** (2020), 142–174.

[28] D. Dolgopyat, P. Nándori and F. Pène, *Asymptotic expansion of correlation functions for* $\mathbb{Z}^d$ *covers of hyperbolic flows*, Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques, to appear, `https://arxiv.org/abs/1908.11504`.

[29] M. Einsiedler and D. Lind, *Algebraic $\mathbb{Z}^d$–actions of entropy rank 1*, Transactions of the American Mathematical Society **356** (2004), 1799–1831.

[30] B. Fayad, G. Forni and A. Kanigowski, *Lebesgue spectrum of countable multiplicity for conservative flows on the torus*, Journal of the American Mathematical Society **34** (2021), 747–813.

[31] K. Fernando and F. Pene, *Expansions in the local and the central limit theorems for dynamical systems*, Communications in Mathematical Physics, to appear, `https://doi.org/10.1007/s00220-021-04255-z`.

[32] M. Field, I. Melbourne and A. Török, *Stability of mixing and rapid mixing for hyperbolic flows*, Annals of Mathematics **166** (2007), 269–291.

[33] L. Flaminio and G, Forni, *Invariant distributions and time averages for horocycle flows*, Duke Mathematical Journal **119** (2003), 465–526.

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

[35] G. Forni, *Deviation of ergodic averages for area-preserving flows on surfaces of higher genus*, Annals of Mathematics **155** (2002), 1–103.

[36] G. Forni, *Effective equidistribution of nilflows and bounds on Weyl sums*, in *Dynamics and Analytic Number Theory*, London Mathematical Society Lecture Notes Series, Vol. 437, Cambridge University Press, Cambridge, 2016, pp. 136–188.

[37] G. Forni and A. Kanigowski, *Time changes of Heisenberg nilflows*, Asterisque **416** (2020), 253–299.

[38] A. Gorodnik and R. Spatzier, *Exponential mixing of nilmanifold automorphisms*, Journal d'Analyse Mathématique **123** (2014), 355–396.

[39] A. Gorodnik and R. Spatzier, *Mixing properties of commuting nilmanifold automorphisms*, Acta Mathematica **215** (2015), 127–159.

[40] S. Gouëzel, *Sharp polynomial estimates for the decay of correlations*, Israel Journal of Mathematics **139** (2004), 29–65.

[41] S. Gouëzel and I. Melbourne, *Moment bounds and concentration inequalities for slowly mixing dynamical systems*, Electronic Journal of Probability **19** (2014), Article no. 93.

[42] S. A. Kalikow, *T, T^{-1} transformation is not loosely Bernoulli*, Annals of Mathematics **115** (1982), 393–409.

[43] A. Kanigowski, F. Rodriguez Hertz and K. Vinhage, *On the non-equivalence of the Bernoulli and K properties in dimension four*, Journal of Modern Dynamics **13** (2018), 221–250.

[44] A. Katok, *Smooth non-Bernoulli K-automorphisms*, Inventiones Mathematicae **61** (1980), 291–299.

[45] Y. Katznelson, *Ergodic automorphisms of $\mathbb{T}^n$ are Bernoulli shifts*, Israel Journal of Mathematics **10** (1971), 186–195.

[46] H. Kesten and F. Spitzer, *A limit theorem related to a new class of self-similar processes*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **50** (1979), 5–25.

[47] D. Y. Kleinbock and G. A. Margulis, *Logarithm laws for flows on homogeneous spaces*, Inventiones Mathematicae **138** (1999), 451–494.

[48] S. Le Borgne, *Exemples de systèmes dynamiques quasi-hyperboliques a decorrelations lentes*, Comptes Rendus Mathématique. Académie des Sciences. Paris **343** (2006), 125–128.

[49] C. Liverani, *On contact Anosov flows*, Annals of Mathematics **159** (2004), 1275–1312.

[50] B. Marcus and S. Newhouse, *Measures of maximal entropy for a class of skew products*, in *Ergodic Theory (Proc. Conf., Math. Forschungsinst., Oberwolfach, 1978)* Lecture Notes in Mathematics, Vol. 729, Springer, Berlin, 1979, pp. 105–125.

[51] I. Melbourne, *Large and moderate deviations for slowly mixing dynamical systems*, Proceedings of the American Mathematical Society **137** (2009), 1735–1741.

[52] I. Melbourne, *Superpolynomial and polynomial mixing for semiflows and flows*, Nonlinearity **31** (2018), R268–R316.

[53] I. Melbourne and M. Nicol, *Large deviations for nonuniformly hyperbolic systems*, Transactions of the American Mathematical Society **360** (2008), 6661–6676.

[54] W. Parry and M. Pollicott, *Zeta Functions and Periodic Orbit Structure of Hyperbolic Dynamics*, Asterisque **187–188** (1990).

[55] F. Pène, *Planar Lorentz process in a random scenery*, Annales de l'Institut Henri Poincaré Probabilités et Statistique **45** (2009), 818–839.

[56] L. Rey–Bellet and L.-S. Young, *Large deviations in non-uniformly hyperbolic dynamical systems*, Ergodic Theory Dynamical Systems **28** (2008), 587–612.

[57] D. Rudolph, *Asymptotically Brownian skew products give non-loosely Bernoulli K-automorphisms*, Inventiones Mathematicae **91** (1988), 105–128.

[58] O. Sarig, *Subexponential decay of correlations*, Inventiones Mathematicae **150** (2002), 629–653.

[59] D. Szász and T. Varjú, *Local limit theorem for the Lorentz process and its recurrence in the plane*, Ergodic Theory and Dynamical Systems **24** (2004), 257–278.

[60] M. Tsujii, *Exponential mixing for generic volume-preserving Anosov flows in dimension three*, Journal of the Mathematical Society of Japan **70** (2018), 757–821.

[61] L.-S. Young, *Statistical properties of dynamical systems with some hyperbolicity*, Annals of Mathematics **147** (1998), 585–650.