



Flexibility of statistical properties for smooth systems satisfying the central limit theorem

Dmitry Dolgopyat¹ · Changguang Dong^{1,3} ·
Adam Kanigowski¹ · Péter Nándori²

Received: 15 June 2020 / Accepted: 26 April 2022 / Published online: 9 June 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of
Springer Nature 2022, corrected publication 2022

Abstract We exhibit new classes of smooth systems which satisfy the Central Limit Theorem (CLT) and have (at least) one of the following properties:

- Zero entropy;
- Weak but not strong mixing;
- (Polynomial) mixing but not K ;
- K but not Bernoulli and mixing at arbitrary fast polynomial rate.

We also give an example of a system satisfying the CLT where the normalizing sequence is regularly varying with index 1. All these examples are C^∞ except for a zero entropy diffeomorphism satisfying the CLT which can be made C^r for an arbitrary finite r .

✉ Dmitry Dolgopyat
dimadolgopyat@gmail.com

Changguang Dong
dongchangguang@gmail.com

Adam Kanigowski
adkanigowski@gmail.com

Péter Nándori
peter.nandori@yu.edu

¹ University of Maryland, College Park, USA

² Yeshiva University, New York, USA

³ Chern Institute of Mathematics and LPMC, Nankai University, Tianjin, China

Contents

Part I. Main results	33
1 Introduction	33
2 Specifying systems appearing in Theorems 1.3–1.5	38
2.1 Zero entropy flow	38
2.2 Zero entropy map	40
2.3 Flows with intermediate mixing properties	41
2.4 K not Bernoulli example	43
3 The main result	45
4 A criterion of CLT for \mathbb{R}^d actions	46
5 The CLT for skew products	48
5.1 A quenched CLT	48
5.1.1 Proof of Property (b)	49
5.1.2 Property (c)	50
5.2 From quenched to annealed CLT: proofs of Theorems 3.1 and 3.2	53
6 Theorem 2.2	55
6.1 Reduction of h2 to a mixing local limit theorem	55
6.2 Variance	63
6.3 Mixing local limit theorem for geodesic flow	66
7 Theorem 2.3	70
8 Surface flows in the base	77
8.1 Proofs of Theorems 2.4 and 2.5	77
8.2 Variance	79
8.3 Proof of Proposition 8.1	81
9 Background on symbolic dynamics	87
10 Weyl chamber flow	89
11 Relative atoms of the past partition	91
12 Non Bernoullicity under zero drift. Proof of Theorem 9.1	94
12.1 The main reduction	94
12.2 Hamming–Bowen closeness	95
12.3 Proof of Proposition 12.2	96
13 Consequence of exponential mixing	97
14 Construction of Ω_k	99
15 Proof of Proposition 12.5	102
Part VI. Appendices	106
Appendix A. Entropy of skew products	106
Appendix B. Ergodic sums over subshifts of finite type	107
B.1 CLT for (T, T^{-1}) transformations with SFT in the base	107
B.2 Visits to cones	109
B.3 Separation estimates for cocycles	110
Appendix C. The main results in general context	112
Appendix D. Open problems	115
References	116

Part I. Main results

1 Introduction

An important discovery made in the last century is that deterministic systems can exhibit chaotic behavior. The Central Limit Theorem (CLT) is a hallmark of chaotic behavior. There is a vast literature on the topic. In particular there are numerous methods of establishing CLT including the method of moments (cumulants) [10, 25], spectral method [54], the martingale method [51, 58, 75] (the list of references here is by no means exhaustive, we just provide a sample of papers which could be used for introducing non-experts to the corresponding techniques and their applications to dynamical systems). However, the above methods require strong mixing properties of the system. As a result, they apply only to systems which have strong statistical properties including Bernoulli property and summable decay of correlations. The only example going beyond strongly chaotic framework as manifested by the Bernoullicity and summable correlations is the product of an Anosov¹ diffeomorphism (called diffeo in the sequel) and a Diophantine rotation, which is shown in [27] to satisfy the CLT (see also [71, 93] or Theorem 3.1 below).

Thus the knowledge on possible ergodic behaviors of smooth systems satisfying CLT is very restricted. The main goal of this paper is to provide new classes of systems satisfying CLT with intermediate ergodic properties. In Appendix C we will describe how our results fit into a general program of flexibility of statistical properties in smooth dynamics.

In order to formulate our results we need a few definitions. Let (M, ζ) be a smooth orientable manifold with a smooth measure ζ . For an integrable function A on M , we denote $\zeta(A) = \int_M A(x)d\zeta(x)$. For $r \in (0, \infty]$ we denote by $C^r(M, \zeta)$ the space of C^r diffeomorphisms of M preserving the measure ζ .

Definition 1.1 Let $r \in (0, \infty]$. We say that F satisfies the *Central Limit Theorem (CLT)* on C^r if $F \in C^r(M, \zeta)$ and there is a sequence a_n such that for each $A \in C^r(M)$,

$$\frac{\sum_{0 \leq j < n} A \circ F^j(\cdot) - n \cdot \zeta(A)}{a_n}$$

converges in law as $n \rightarrow \infty$ to normal random variable with zero mean and variance $\sigma^2(A)$ (such normal random variable will be denoted $\mathcal{N}(0, \sigma^2(A))$ in the sequel) and, moreover, $\sigma^2(\cdot)$ is not identically equal to zero on $C^r(M)$.

¹ The methods of [27] apply to more general systems in the first factor, however, they seem insufficient to produce the examples described in Theorems 1.3–1.5.

We say that F satisfies the *CLT* if it satisfies the CLT on C^r for some $r > 0$. We say that F satisfies the *classical CLT* if one can take $a_n = \sqrt{n}$.

One can analogously define the CLT for a flow $(F_T) \in C^r(M, \zeta)$ replacing

$$\frac{1}{a_n} \left[\sum_{0 \leq j < n} A \circ F^j(\cdot) - n \cdot \zeta(A) \right] \text{ by } \frac{1}{a_S} \left[\int_0^S A \circ F_s(\cdot) ds - S \cdot \zeta(A) \right],$$

where a_S is now a real valued function.

Definition 1.2 Let $\psi : \mathbb{N} \rightarrow \mathbb{R}$ be a function. We say that F is mixing on C^r at the rate ψ if $F \in C^r(M, \zeta)$ and for any $A_1, A_2 \in C^r(M)$ the correlation function $\rho_n(A_1, A_2) = \zeta(A_1 \cdot (A_2 \circ F^n)) - \zeta(A_1)\zeta(A_2)$ satisfies

$$|\rho_n(A_1, A_2)| \leq \|A_1\|_{C^r} \|A_2\|_{C^r} \psi(n). \quad (1.1)$$

We say that F is mixing at the rate ψ if it is mixing with the rate ψ on C^r for some $r > 0$. In case $\psi(n) = Cn^{-\delta}$ for some $C, \delta > 0$, we say that F is *polynomially mixing*. If $\psi(n) = Ce^{-\delta n}$ for some $C, \delta > 0$, we say that F is *exponentially mixing*.²

The above definitions can be extended to flows in a straightforward way by replacing the discrete parameter $n \in \mathbb{N}$ with a continuous parameter $t \in \mathbb{R}$. We are ready to state the main results only using basic notions of ergodic theory and the above two definitions. A more detailed exposition of the main results as well as prerequisite earlier work will be presented shortly in Sect. 2. Our first main result deals with the CLT for zero entropy systems:

Theorem 1.3 (a) *There exists an analytic flow of zero entropy which satisfies the CLT with normalization $a_T = T / \ln^{1/4} T$.*

(b) *For each $r \in \mathbb{N}$ there is a smooth manifold (M_r, ζ_r) and a zero entropy diffeomorphism $F_r \in C^r(M_r, \zeta_r)$ which satisfies the classical CLT.*

We note that in all previous results on the CLT the normalization was regularly varying³ with index $\frac{1}{2}$.⁴ Theorem 1.3(a) is the first result for a CLT with a

² We note that a simple interpolation argument shows that if F is mixing with exponential (respectively polynomial rate) on C^r for some $r > 0$ then it is mixing with exponential (respectively polynomial) rate on C^r for all $r > 0$, however the exponent δ depends on r .

³ Recall that a real valued function $a(\cdot)$ defined on $[m, \infty)$ for some $m \in \mathbb{R}$ is *regularly varying in the sense of Karamata* with index α if for each $s > 0$, $\lim_{t \rightarrow \infty} \frac{a(st)}{a(t)} = s^\alpha$. A sequence a_n is regularly varying with index α if the function $a(t) = a_{[t]}$ is regularly varying with index α .

⁴ CLT with normalization $\sqrt{n} \ln n$ appears for expanding and hyperbolic maps with neutral fixed points [19, 53], as well as in several hyperbolic billiards [5, 6, 101]. In a followup paper we will show it also appears for generalized T, T^{-1} transformations with hyperbolic base and two parameter exponentially mixing flows in the fiber.

different regularly varying index, namely 1.⁵ We say that a system F is K if it has no non-trivial zero entropy factor, [98]. In the theorem below, we give examples of weakly mixing but not mixing as well as polynomially mixing but not K systems satisfying the CLT.

Theorem 1.4 (a) *There exists a weakly mixing but not mixing C^∞ -flow, which satisfies the classical CLT.*

(b) *There exists a polynomially mixing C^∞ -flow, which is not K and satisfies the classical CLT.*

Recall that a system is Bernoulli if it is isomorphic to a Bernoulli shift. Our next result shows existence of K non Bernoulli systems which satisfy the classical CLT and are mixing at arbitrary fast polynomial rate.

Theorem 1.5 *For each $m \in \mathbb{N}$ there exists a manifold (M_m, ζ_m) and $F_m \in C^\infty(M_m, \zeta_m)$ which is mixing at rate n^{-m} but is not Bernoulli. Moreover, F_m is K and satisfies the classical CLT.*

To the best of our knowledge, the first part of the theorem provides the first example of a system which has summable correlations but is not Bernoulli. The second (“moreover”) part answers a question that we heard from multiple sources, initially from J-P. Thouvenot.

All the systems in Theorems 1.3–1.5 belong to the class of generalized (T, T^{-1}) transformations which we now describe. The class of generalized (T, T^{-1}) transformations is a classical subject (see [59, 85, 106] and reference therein for some early work on this topic) with a rich range of applications in probability and ergodic theory. In fact, generalized (T, T^{-1}) transformations were used to exhibit examples of systems with unusual limit laws [28, 68], central limit theorem with non standard normalization [12], K but non Bernoulli systems in abstract [60] and smooth setting in various dimensions [62, 63, 99], very weak Bernoulli but not weak Bernoulli partitions [31], slowly mixing systems [32, 35, 80], systems with multiple Gibbs measures [45, 83]. To define (C^r -smooth) (T, T^{-1}) transformations, let X, Y be compact orientable manifolds, $f : X \rightarrow X$ be an ergodic C^r map preserving a smooth measure μ and $G_t : Y \rightarrow Y$ be a C^r -smooth \mathbb{R}^d action on the manifold Y preserving a smooth measure ν .

Definition 1.6 Let X, Y, f, G_t be as above. Let $\tau : X \rightarrow \mathbb{R}^d$ be a C^r -smooth function that will be called a *cocycle*. The map $F : X \times Y \rightarrow X \times Y$ defined

⁵ We note that the requirement that the limiting distribution is Gaussian is important. If we allow other limit distributions, then there are several examples in both non-uniformly hyperbolic and parabolic settings where normalization is different, see [1, 20, 48] and references therein. If we allow our system to preserve an infinite measure then, there is an additional freedom related to the rate of return times, see e.g. [30, 96].

by

$$F(x, y) = (f(x), G_{\tau(x)}y) \quad (1.2)$$

is called a (C^r -smooth) (T, T^{-1}) transformation.

Note that F is C^r (since so are f, G_t and τ) and it preserves the measure $\zeta = \mu \times \nu$. Moreover,

$$F^N(x, y) = (f^N x, G_{\tau_N(x)}y), \quad (1.3)$$

where

$$\tau_N(x) = \sum_{n=0}^{N-1} \tau(f^n x). \quad (1.4)$$

We also analogously define (T, T^{-1}) flows. Namely let h_t be a C^r -flow on X preserving μ . Set

$$F_T(x, y) = (h_T(x), G_{\tau_T(x)}y) \quad \text{where} \quad \tau_T(x) = \int_0^T \tau(h_t x) dt. \quad (1.5)$$

Note that if F_T is a (T, T^{-1}) flow, then for each t_0 , the time t_0 map of F is a generalized (T, T^{-1}) transformation.

In this paper we study (T, T^{-1}) systems whose fiber dynamics are very chaotic:

Definition 1.7 G_t is *exponentially mixing of all orders* if there is $r > 0$ such that for every $m \in \mathbb{N}$ there exist $C_m, \delta_m > 0$ such that for every $A_j \in C^r(Y)$, $j = 1, \dots, m$, we have

$$\begin{aligned} & \left| \int_Y \left(\prod_{j=1}^m A_j(G_{t_j}x) \right) d\nu(x) - \prod_{j=1}^m \nu(A_j) \right| \\ & \leq C_m \prod_{j=1}^m \|A_j\|_{C^r} e^{-\delta_m \min_{i \neq j} \|t_i - t_j\|}. \end{aligned} \quad (1.6)$$

We note that if G_t is exponentially mixing of all orders then so is its any subaction, that is, the action of any proper subgroup $V \subset \mathbb{R}^d$.

Throughout the paper we assume that the action G_t is exponentially mixing of all orders.

The main example that the reader should keep in mind is the following:

Example 1.8 Let $d \geq 1$ and let Γ be a co-compact lattice in $SL(d+1, \mathbb{R})$. Let D_+ be the group of diagonal matrices in $SL(d+1, \mathbb{R})$ with positive elements

on the diagonal. It is easy to see that D_+ is isomorphic to \mathbb{R}^d . The group D_+ acts on $SL(d+1, \mathbb{R})/\Gamma$ by left translation. When $d=1$, this one parameter flow is called the *geodesic flow*. When $d \geq 2$, we obtain a \mathbb{R}^d action (G_t) , which is called the *Weyl Chamber flow*. Then (G_t) is exponentially mixing of all orders (see [9]).

Notice that by Definition 1.1, for all the systems appearing in Theorems 1.3–1.5, the variance is not identically zero. One could ask for a stronger property (motivated by classical results in hyperbolic dynamics): the variance of an observable is zero if and only if the observable is a *coboundary* for the system. This stronger property does not hold in the setting of Theorems 1.3 and 1.4. Indeed one of the key steps in the skew product constructions of Theorems 1.3 and 1.4 is to choose base transformations (or flows) with some zero mean cocycle over them whose ergodic sums (or integrals) are not tight, while the growth of all cocycles with the same regularity is of order $o(a_N)$ ($o(a_S)$). This inevitably produces a class of observables, namely, those depending only on the base coordinate, that have zero asymptotic variance while not all of them are coboundaries. The situation is different in the setting of Theorem 1.5 where the base map f is hyperbolic (and hence in particular satisfies the above stronger property). This in fact allows to show that systems constructed in Theorem 1.5 satisfy this stronger version of CLT (see also Remark 5.6. in [35]). It is an open question if one can produce systems as in Theorems 1.3 and 1.4 satisfying this stronger version of CLT.

In order to construct our examples we need to extend significantly the existing methods for proving both the CLT and the non Bernoulli property of these maps. In fact, the main difficulty in Theorems 1.3 and 1.4 is to establish the CLT while other properties are rather straightforward. On the other hand, the main difficulty in Theorem 1.5 is to show non Bernoullicity. We note that even though the question about the CLT and the non Bernoulli properties seem quite different, the key tools needed to answer both questions are the same. Namely, the proofs of all theorems in the paper rely on the exponential mixing in the fiber and the fine recurrence properties of deterministic cocycles. More details on the general framework for proving the CLT for generalized (T, T^{-1}) transformations is presented in Sect. 3, while the precise results pertaining to the non Bernoullicity are described in Part V.

Outline of the paper: The rest of the paper is organized as follows. In Sect. 2 we showcase the examples realizing Theorems 1.3–1.5. Specifically, in §2.1 we discuss Theorem 1.3(a), in §2.2 we discuss Theorem 1.3(b), in §2.3 we discuss Theorem 1.4, and in §2.4 we discuss Theorem 1.5. There are no technical proofs in Sect. 2, we just formulate more specific results, namely Theorems 2.2–2.7 that imply the main Theorems 1.3–1.5. Parts II–IV complete the proof of the theorems from Sect. 2.

In Part II, we discuss the CLT for generalized (T, T^{-1}) transformations. In Sect. 3, we state two results (Theorems 3.1 and 3.2) that are the main tools in proving CLT for all our examples (in the discrete and continuous case respectively). The proofs of these results occupy the rest of Part II. Part III is devoted to the proof of Theorems 2.2 and 2.3. Part IV is devoted to the proof of Theorems 2.4 and 2.5. In Part V, we prove Theorem 2.7. Finally in Part VI we prove some technical results needed in the proof (Appendix A and Appendix B) as well as discuss the general context of flexibility of statistical properties (Appendix C) and state some open problems (Appendix D).

2 Specifying systems appearing in Theorems 1.3–1.5

We will now make precise what type of generalized (T, T^{-1}) transformations will be used in the proofs of our main results. We present the examples in four subsections below.

2.1 Zero entropy flow

We start with the following lemma which shows that the entropy of a generalized (T, T^{-1}) transformation is zero provided that the base map has entropy zero and the cocycle has zero mean. Recall Definition 1.6.

Lemma 2.1 *Let $F \in C^r(X \times Y, \zeta)$ with $r > 1$ be a generalized (T, T^{-1}) transformation such that f is ergodic, $\text{ent}_\mu(f) = 0$ and $\mu(\tau) = 0$. Then $\text{ent}_\zeta(F) = 0$. The same result holds for (T, T^{-1}) flows.*

Here $\text{ent}_\mu(\cdot)$ denotes the metric entropy. The proof is given in Appendix A.

Let Q be a hyperbolic surface of constant negative curvature of arbitrary genus $p \geq 1$. Let h_t be the (stable) horocycle flow on the unit tangent bundle $X = S\mathcal{Q}$, that is, h_t is moving $x \in X$ at unit speed along its stable horocycle

$$\mathcal{H}(x) = \left\{ \tilde{x} \in X : \lim_{t \rightarrow \infty} d(\mathbf{G}_t(x), \mathbf{G}_t(\tilde{x})) = 0 \right\} \quad (2.1)$$

where \mathbf{G}_t is the geodesic flow on X . Let $\gamma_1, \dots, \gamma_{2p}$ be the basis in homology of Q . Choose $i \in \{1, \dots, 2p\}$ and let λ be a closed form on Q such that

$$\int_{\gamma_j} \lambda = \delta_{ij}, \quad (2.2)$$

where δ is the Kronecker symbol. Set

$$\tau(q, v) = \lambda(q)(v^*), \quad (2.3)$$

where v^* is a unit vector obtained from v by the 90 degree rotation. Let (G_t, Y, v) be an \mathbb{R} action which is exponentially mixing of all orders. Consider the system (see (1.5))

$$F_T(x, y) = (h_T(x), G_{\tau_T(x)}y). \quad (2.4)$$

We have

$$\tau_T(x) = \int_{\mathfrak{h}(x, T)} \lambda, \quad (2.5)$$

where $\mathfrak{h}(x, T)$ is the projection of the horocycle starting from x and of length T , to \mathcal{Q} .

Theorem 1.3(a) follows immediately from the following theorem:

Theorem 2.2 *Let $(F_T)_{T \in \mathbb{R}}$ be the flow defined in (2.4). Then*

- h1.** $\text{ent}_\zeta(F_T) = 0$;
- h2.** *For every smooth observable $H \in C^\infty(X \times Y)$ with $\zeta(H) = 0$, there exists $\sigma^2(H) \geq 0$ such that*

$$\frac{(\ln T)^{1/4}}{T} \int_0^T H(F_t(\cdot)) dt$$

converges as $T \rightarrow \infty$ to the normal distribution with zero mean and variance $\sigma^2(H)$;

- h3.** *There exists $H \in C^\infty(X \times Y)$ with $\zeta(H) = 0$ such that $\sigma^2(H) > 0$.*

The proof of Theorem 2.2 is provided in Sect. 6. To demystify the normalization $T/(\ln T)^{1/4}$ let us write

$$\int_0^T H(F_t(\cdot)) dt = \sum_n Z_{n,T}, \quad (2.6)$$

where $Z_{n,T}$ contains the contribution of the times t when $\tau_t(x) \in [n, n+1]$. According to the analysis in [43] of

$$\text{mes}(t \in [0, T] : \tau_t(x) \in [n, n+1]), \quad (2.7)$$

(2.7) is significant if n is of order $\sqrt{\ln T}$ and we will show that in this case (2.7) is of order $T/\sqrt{\ln T}$.

Thus (2.6) is the sum of $O(\sqrt{\ln T})$ summands, each with variance $(T/\sqrt{\ln T})^2$. As we will see, these summands are only weakly correlated and so (2.6) will be approximately normal with variance $T^2/\sqrt{\ln T}$. This explains h2. In the proof, we will in fact need a more precise estimate related to (2.7) which, besides [43, Theorem 5.1] also uses some methods introduced in [40].

Another important step of the proof is to quantify the weak dependence among the summands to the extent that the CLT holds. To this end, we use a CLT by Björklund and Gorodnik [10] (see Sect. 4 for a precise statement). We will need some extensions of the main theorem of [10] not just here but also for the forthcoming Theorems 2.3–2.5. We refer the reader to Sect. 6 for a complete proof of Theorem 2.2.

2.2 Zero entropy map

We will now define the generalized (T, T^{-1}) transformations used in Theorem 1.3 (b). Let $\mathbf{m} \in \mathbb{N}$ and let $\|\cdot\|$ denote the distance to the nearest integer in $\mathbb{R}^{\mathbf{m}}$. For $\kappa > 0$, let

$$\mathbb{D}(\kappa) = \left\{ \alpha \in \mathbb{T}^{\mathbf{m}} : \exists D(\alpha) > 0 \text{ such that } \left\| \langle k, \alpha \rangle \right\| \geq \frac{D(\alpha)}{|k|^\kappa}, \text{ for every } k \in \mathbb{Z}^{\mathbf{m}} \setminus \{0\} \right\}.$$

Recall that from Khintchine's theorem ([70]) it follows that $\mathbb{D}(\kappa)$ is non-empty if $\kappa \geq \mathbf{m}$ and it has full measure if $\kappa > \mathbf{m}$. Theorem 1.3 (b) immediately follows from

Theorem 2.3 *Let $\mathbf{m} \in \mathbb{N}$ and $\kappa \in [\mathbf{m}, 2\mathbf{m}]$. Let μ be the Lebesgue measure on $\mathbb{T}^{\mathbf{m}}$. For every $\alpha \in \mathbb{D}(\kappa)$ and every $r \in (\kappa/2, \mathbf{m})$ there exists $d \in \mathbb{N}$ and a function $\tau \in C^r(\mathbb{T}^{\mathbf{m}}, \mathbb{R}^d)$ such that if (G_t, Y, ν) is a C^∞ smooth \mathbb{R}^d action which is exponentially mixing of all orders then the system*

$$F : (\mathbb{T}^{\mathbf{m}} \times Y, \mu \times \nu) \rightarrow (\mathbb{T}^{\mathbf{m}} \times Y, \mu \times \nu), \quad F(x, y) = (x + \alpha, G_{\tau(x)}y)$$

satisfies:

- r1.** $\text{ent}_{\mu \times \nu}(F) = 0$;
- r2.** For every $H \in C^r(\mathbb{T}^{\mathbf{m}} \times Y)$ with $(\mu \times \nu)(H) = 0$, there is $\sigma^2(H) \geq 0$ such that

$$\frac{1}{\sqrt{N}} H_N = \frac{1}{\sqrt{N}} \sum_{n \leq N} H(F^n(\cdot))$$

converges as $N \rightarrow \infty$ to the normal distribution with zero mean and variance $\sigma^2(H)$;

- r3.** There exists $H \in C^r(\mathbb{T}^{\mathbf{m}} \times Y)$ with $(\mu \times \nu)(H) = 0$ such that $\sigma^2(H) > 0$.

Since the rotation by α and the action (G_t, Y, ν) are both C^∞ and the cocycle τ is of class $C^r(\mathbb{T}^{\mathbf{m}}, \mathbb{R}^d)$ it follows that the map F is of class C^r and so indeed the above theorem implies Theorem 1.3 (b). We prove Theorem 2.3 in Sect. 7.

2.3 Flows with intermediate mixing properties

We will now describe the class of generalized (T, T^{-1}) transformations used to prove Theorem 1.4. The base flows in our construction are a subclass of the class of smooth flows on surfaces. For more details on smooth flows on surfaces we refer the reader to [4, 73, 74, 103, 104]. In particular it follows by Pesin entropy formula ([7]) that the entropy of any smooth flow on a surface is equal to 0. Let M be a surface and let (φ_t) be a C^∞ flow on M that preserves the area μ . Ergodic properties of smooth flows on surfaces have been successfully studied via their *special representation*. More precisely, one considers a one dimensional closed transversal \mathcal{T} on M and represents the flow as the special flow over the *first return map* to $\mathcal{T} \sim \mathbb{T}$ and under the roof function f which is the *first return time*. Since the flow is smooth, the return function is also smooth except for fixed points of the flow, at which f blows up. In particular, every point in $x \in M$ which is not a fixed point can be written as $x = \varphi_s \theta$, where $\theta \in \mathcal{T}$ and $0 \leq s < f(\theta)$.

In what follows we will always assume that the set of fixed points of (φ_t) is non-empty and finite. In the case of smooth flows on surfaces the first return map to \mathcal{T} is an interval exchange transformation or in some cases (which will be our main focus) an irrational rotation. For a more detailed discussion on special representation of (φ_t) we refer the reader to [49, 73, 74]. We will now describe what examples of smooth flows (φ_t) will be considered in this paper.

Let $\alpha \in \mathbb{T}$ be an irrational number. Let $f : \mathbb{T} \rightarrow \mathbb{R}_+$ be a function which is C^3 on $\mathbb{T} \setminus \{0\}$, satisfies $\int f dLeb = 1$ and

$$\lim_{\theta \rightarrow 0^+} \frac{f''(\theta)}{h''(\theta)} = A \quad \text{and} \quad \lim_{\theta \rightarrow 1^-} \frac{f''(\theta)}{h''(1 - \theta)} = B, \quad (2.8)$$

where $A^2 + B^2 \neq 0$ and the function h belongs to one of the classes specified below.

- (1) $h(\theta) = \log \theta$ and $A = B$, then for every $\alpha \in \mathbb{T} \setminus \mathbb{Q}$ there exists f satisfying (2.8) such that $R_\alpha \theta = \theta + \alpha$ is the first return map and f is the first return time of some C^∞ ergodic flow (φ_t) on a surface (M, μ) with genus ≥ 2 (see e.g. [73] or [49, Proposition 2]). Such flows are not mixing, [73], but are weakly mixing for every α , [49]. Let us denote $\mathcal{K}(\alpha, \text{logsym})$ the set of C^∞ area preserving flows (φ_t) for which R_α is the first return map and the corresponding first return time f satisfies (2.8) (with $h(\theta) = \log(\theta)$).
- (2) $h(x) = x^{-\gamma}$, $\gamma \in B_{\text{sing}}$ where B_{sing} is a non-empty set such that for every $\alpha \in \mathbb{T} \setminus \mathbb{Q}$ there exists f satisfying (2.8) with h such that $R_\alpha \theta = \theta + \alpha$ is the first return map and f is the first return time of some C^∞ ergodic flow (φ_t) on the torus (\mathbb{T}^2, μ) , [74]. In [74] it is shown that $\gamma = 1/3 \in B_{\text{sing}}$. Moreover by [74] (φ_t) is mixing for every α and by [47] if $\gamma \leq 2/5$,

then the flow is *polynomially* mixing for a.e. α . In what follows we will always assume that $\gamma \leq 2/5$. For $\gamma \in B_{sing}$, let us denote $\mathcal{K}(\alpha, \gamma)$ the set of smooth area preserving flows (φ_t) on \mathbb{T}^2 for which R_α is the first return map and the corresponding first return time f satisfies (2.8) with $h(x) = x^{-\gamma}$.

We will consider the continuous flow F_T given by (see (1.5)) $F_T(x, y) = (\varphi_T(x), G_{\tau_T}(y))$, where (φ_t) is as in (1) or (2) above and τ and G_t are defined in the theorems below.

Theorem 2.4 *Let (G_t, Y, v) be a C^∞ flow which is exponentially mixing of all orders and let $\tau : M \rightarrow \mathbb{R}$ be any C^∞ positive function. There exists $\mathcal{F} \subset \mathbb{T}$ with $Leb(\mathcal{F}) = 1$ such that if $\alpha \in \mathcal{F}$, $(\varphi_t) \in \mathcal{K}(\alpha, \text{logsym})$, and $F_T(x, y) := (\varphi_T(x), G_{\tau_T}(y))$, then*

- w1.** *$(F_T)_{T \in \mathbb{R}}$ is weakly mixing but not mixing;*
- w2.** *For every $H \in C^\infty(M \times Y)$ with $(\mu \times v)(H) = 0$, there exists $\sigma^2(H) \geq 0$ such that*

$$\frac{1}{\sqrt{T}} H_T = \frac{1}{\sqrt{T}} \int_0^T H(F_t(\cdot)) dt$$

converges as $T \rightarrow \infty$ to the normal distribution with zero mean and variance $\sigma^2(H)$;

- w3.** *There exists $H \in C^\infty(M \times Y)$ with $(\mu \times v)(H) = 0$ such that $\sigma^2(H) > 0$.*

The above theorem immediately implies Theorem 1.4 (a).

Theorem 2.5 *Let (G_t, Y, v) be a C^∞ flow which is exponentially mixing of all orders and let $\tau : \mathbb{T}^2 \rightarrow \mathbb{R}$ be any C^∞ positive function. There exists $\mathcal{F}' \subset \mathbb{T}$ with $Leb(\mathcal{F}') = 1$ such that if $\alpha \in \mathcal{F}'$, $(\varphi_t) \in \mathcal{K}(\alpha, \gamma)$ for $\gamma \in B_{sing}$ and $F_T(x, y) := (\varphi_T(x), G_{\tau_T}(y))$, then*

- n1.** *$(F_T)_{T \in \mathbb{R}}$ is polynomially mixing and not K;*
- n2.** *For every $H \in C^\infty(\mathbb{T}^2 \times Y)$ with $(\mu \times v)(H) = 0$, there exists $\sigma^2(H) \geq 0$ such that*

$$\frac{1}{\sqrt{T}} H_T = \frac{1}{\sqrt{T}} \int_0^T H(F_t(\cdot)) dt$$

converges as $T \rightarrow \infty$ to the normal distribution with zero mean and variance $\sigma^2(H)$;

- n3.** *There exists $H \in C^\infty(\mathbb{T}^2 \times Y)$ with $(\mu \times v)(H) = 0$ such that $\sigma^2(H) > 0$.*

Notice that the above theorem immediately implies Theorem 1.4 (b).

Remark 2.6 In the above two theorems we only need that the C^∞ cocycle τ is positive. The simplest case of our theorem is to take $\tau \equiv 1$. In this case the resulting (T, T^{-1}) transformation is just the direct product flow $(\varphi_t \times G_t)$.

2.4 K not Bernoulli example

We will now specify the (T, T^{-1}) transformations that we will use in the proof of Theorem 1.5. Let $f : (\mathbb{T}^m, \mu) \rightarrow (\mathbb{T}^m, \mu)$ be a volume preserving Anosov diffeomorphism.

Let $\tau : \mathbb{T}^m \rightarrow \mathbb{R}^d$ be a mean zero cocycle. We shall say that τ is *irreducible* if it is not cohomologous to a cocycle taking value in a proper linear subspace of \mathbb{R}^d .

Recall Example 1.8. Theorem 1.5 is a consequence of the following result.

Theorem 2.7 *Fix an integer $d \geq 1$. Let $f : (\mathbb{T}^m, \mu) \rightarrow (\mathbb{T}^m, \mu)$ be a volume preserving Anosov diffeomorphism. Let (G_t) be a geodesic flow on $SL(2, \mathbb{R})/\Gamma$ (if $d = 1$), or a Weyl chamber flow on $SL(d+1, \mathbb{R})/\Gamma$ (when $d \geq 2$). Let $\tau : \mathbb{T}^m \rightarrow \mathbb{R}^d$ be a mean zero irreducible Hölder cocycle. Then the map on $\mathbb{T}^m \times SL(d+1, \mathbb{R})/\Gamma$ defined by*

$$F_d(x, y) = (fx, G_{\tau(x)}y)$$

with the invariant measure $\mu \times \text{Haar}$ is non-Bernoulli.

The irreducibility assumption is not too restrictive.

First, it holds for most cocycles. To see this we shall use the following well known fact. Let $\tau_{(1)}(x), \dots, \tau_{(d)}(x)$ denote the components of the vector $\tau(x)$. Recall that by the CLT for Anosov diffeos (see e.g. [92, Chapter 4]) τ_N/\sqrt{N} converges in law as $N \rightarrow \infty$ to a normal random variable with zero mean and covariance matrix with components

$$D_{i,j}^2(\tau) = \sum_{n=0}^{\infty} \mu(\tau_{(i)}(\tau_{(j)} \circ f^n)). \quad (2.9)$$

Proposition 2.8 *Let $f : (\mathbb{T}^m, \mu) \rightarrow (\mathbb{T}^m, \mu)$ be a volume preserving Anosov diffeomorphism and $\tau : \mathbb{T}^m \rightarrow \mathbb{R}^d$ be a zero mean Hölder cocycle. Then the following are equivalent.*

- (i) *There is a measurable function $h : \mathbb{T}^m \rightarrow \mathbb{R}^d$ such that $\tau - h + h \circ f$ takes values in a proper linear subspace;*
- (ii) *There is a Hölder function $h : \mathbb{T}^m \rightarrow \mathbb{R}^d$ such that $\tau - h + h \circ f$ takes values in a proper linear subspace;*

- (iii) The diffusion matrix $D^2(\tau)$ is degenerate, i.e. there is a unit vector \mathbf{u} such that $D^2(\tau)\mathbf{u} = 0$;
- (iv) There is a unit vector \mathbf{u} such that if x is a periodic point of period p then $\tau_p(x) \perp \mathbf{u}$.

Thus if τ is reducible, then for any collection of d periodic points x_1, \dots, x_d of periods p_1, \dots, p_d the determinant of the matrix with components $Q_{ij} = (\tau_{(i)})_{p_j}(x_j)$ is zero. Since there are infinitely many periodic orbits, τ must satisfy infinitely many algebraic equations. Thus the set of reducible cocycles is contained in an algebraic submanifold of infinite codimension.

Second, if τ is reducible we still can apply Theorem 2.7 to a lower rank subaction. Namely suppose $\tilde{\tau} = \tau - h + h \circ f$ takes values in a proper subspace V . Then the transformations defined by τ and by $\tilde{\tau}$ are conjugated via the change of variables $(x, y) \mapsto (x, G_h(y))$. Thus to understand the (T, T^{-1}) map defined by τ one can study the (T, T^{-1}) map defined by $\tilde{\tau}$ which is associated to the lower rank subaction of $V \subset \mathbb{R}^d$.

Proof of Proposition 2.8 If $\tau - h + h \circ f \in V$ where V is a proper linear subspace of \mathbb{R}^d , then taking a unit vector \mathbf{u} orthogonal to V we get $\langle \tau(x), \mathbf{u} \rangle = \hat{h}(x) - \hat{h} \circ f(x)$ where $\hat{h}(x) = \langle h(x), \mathbf{u} \rangle$. Conversely, if for some unit vector \mathbf{u} we have that $\langle \tau, \mathbf{u} \rangle = \hat{h} - \hat{h} \circ f$ then $\tau - [\hat{h} - \hat{h} \circ f] \mathbf{u}$ belongs to the orthogonal complement of \mathbf{u} . Also denoting $\hat{\tau}_{\mathbf{u}} = \langle \tau, \mathbf{u} \rangle$ we have that

$$\langle D^2(\tau)\mathbf{u}, \mathbf{u} \rangle = \sum_{n=-\infty}^{\infty} \mu(\hat{\tau}_{\mathbf{u}}(\hat{\tau}_{\mathbf{u}} \circ f^n)) =: \sigma^2(\hat{\tau}_{\mathbf{u}}).$$

The foregoing discussion shows that for $y \in \{i, ii, iii, iv\}$ we have that (y) holds iff there exists a unit vector $\mathbf{u} \in \mathbb{R}^d$ such that $(\widehat{y})_{\mathbf{u}}$ holds where

- $\widehat{(i)}_{\mathbf{u}}$ The equation $\hat{\tau}_{\mathbf{u}} = \hat{h} - \hat{h} \circ f$ has a measurable solution;
- $\widehat{(ii)}_{\mathbf{u}}$ The equation $\hat{\tau}_{\mathbf{u}} = \hat{h} - \hat{h} \circ f$ has a Hölder solution;
- $\widehat{(iii)}_{\mathbf{u}} \sigma^2(\hat{\tau}_{\mathbf{u}}) = 0$;
- $\widehat{(iv)}_{\mathbf{u}}$ For each periodic point x of period p , $\hat{\tau}_p(x) = 0$.

However, for each fixed \mathbf{u} the properties $\widehat{(i)}_{\mathbf{u}}$, $\widehat{(ii)}_{\mathbf{u}}$, $\widehat{(iii)}_{\mathbf{u}}$, and $\widehat{(iv)}_{\mathbf{u}}$ are equivalent. Indeed the equivalence of $\widehat{(i)}_{\mathbf{u}}$, $\widehat{(ii)}_{\mathbf{u}}$, and $\widehat{(iv)}_{\mathbf{u}}$ follows from Livsic Theorem [82], while the equivalence of $\widehat{(i)}_{\mathbf{u}}$ and $\widehat{(iii)}_{\mathbf{u}}$ follows from the L_2 -Gottschalk-Hedlund Theorem ([23]). This completes the proof of the proposition. \square

Proof of Theorem 1.5 The K property for F_d with any $d \geq 1$ follows from Corollary 2 in [57], the classical CLT for any $d \geq 3$ follows from [35, Theorem 5.1] (since the proof of Theorem 5.1 in [35] is relatively long we provide a different proof of the CLT in §B.1 using the tools developed in Part II) and

mixing of F_d with rate $n^{-d/2}$ follows from [35, Theorem 4.6]. Finally non-Bernoullicity follows from Theorem 2.7. \square

The proof of Theorem 2.7 is carried out in Part V.

Part II. Central Limit Theorem for (T, T^{-1}) transformations

3 The main result

Here we present sufficient conditions for generalized (T, T^{-1}) transformations defined by (1.2) (and (1.5)) to satisfy the CLT. Namely, Theorems 3.1 and 3.2 below give such conditions for discrete and continuous time (T, T^{-1}) -transformations, respectively.

Recall (1.4).

Theorem 3.1 *Let $r \in \mathbb{R}_+$ and $f \in C^r(M)$ satisfy the following: for each $A \in C^r(M)$ with $\mu(A) = 0$, there is a number $\sigma^2(A) \geq 0$ such that*

$$\frac{1}{\sqrt{N}} \sum_{0 \leq n < N} A(f^n \cdot) \Rightarrow \mathcal{N}(0, \sigma^2(A)) \quad (3.1)$$

as⁶ $N \rightarrow \infty$, where the left hand side is understood as a random variable with respect to the measure μ . Let $\tau : M \rightarrow \mathbb{R}^d$ be a C^r cocycle satisfying the following: there are $\varepsilon > 0$ and $C > 0$ so that for every $N \geq 2$,

$$\mu(x \in M : |\tau_N(x)| < \log^{1+\varepsilon} N) < \frac{C}{N^5}. \quad (3.2)$$

Let (G_t, Y, ν) be a $C^\infty \mathbb{R}^d$ action which is exponentially mixing of all orders and let $F(x, y) = (fx, G_{\tau(x)}y)$. Then for every $H \in C^r(M \times Y)$ with $(\mu \times \nu)(H) = 0$, there is $\Sigma^2(H) \geq 0$ such that

$$\frac{1}{\sqrt{N}} \sum_{0 \leq n < N} H(F^n(\cdot, \cdot)) \Rightarrow \mathcal{N}(0, \Sigma^2(H))$$

as $N \rightarrow \infty$. Moreover, if $\sigma^2(A) = 0$ for all $A \in C^r(M)$, then

$$\Sigma^2(H) = \sum_{k=-\infty}^{\infty} \int_M \int_Y \tilde{H}(x, y) \tilde{H}(f^k x, G_{\tau_k(x)}y) d\nu(y) d\mu(x), \quad (3.3)$$

where $\tilde{H}(x, y) = H(x, y) - \int_Y H(x, y) d\nu(y)$.

⁶ Here, and in the sequel, \Rightarrow denotes weak convergence of random variables. Note that in contrast with Definition 1.1, we do not require $\sigma^2(A) > 0$.

Next, we extend Theorem 3.1 to continuous time. Below, $\tau_T(x) := \int_0^T \tau(f_t x) dt$.

Theorem 3.2 *Let $r \in \mathbb{R}_+$ and $f \in C^r(M)$ satisfy the following: for each $A \in C^r(M)$ with $\mu(A) = 0$, there is a number $\sigma^2(A) \geq 0$ such that*

$$\frac{1}{\sqrt{T}} \int_0^T A(f_t \cdot) dt \Rightarrow \mathcal{N}(0, \sigma^2(A)) \quad (3.4)$$

as $T \rightarrow \infty$. Let $\tau : M \rightarrow \mathbb{R}^d$ be a C^r cocycle satisfying: there are $\varepsilon > 0$ and $C > 0$ so that for every $T \geq 2$,

$$\mu \left(x \in M : |\tau_T(x)| < \log^{1+\varepsilon} T \right) < \frac{C}{T^\varepsilon}. \quad (3.5)$$

Let (G_t, Y, ν) be a C^∞ , \mathbb{R}^d action which is exponentially mixing of all orders and let $F_t(x, y) = (f_t x, G_{\tau_t(x)} y)$. Then for every $H \in C^r(M \times Y)$ with $(\mu \times \nu)(H) = 0$ there is $\Sigma^2(H) \geq 0$ such that

$$\frac{1}{\sqrt{T}} \int_0^T H(F_t(\cdot, \cdot)) dt \Rightarrow \mathcal{N}(0, \Sigma^2(H))$$

as $T \rightarrow \infty$. Moreover, if $\sigma^2(A) = 0$ for all $A \in C^r(M)$, then

$$\Sigma^2(H) = \int_{-\infty}^{\infty} \int_M \int_Y \tilde{H}(x, y) \tilde{H}(f_t x, G_{\tau_t(x)} y) d\nu(y) d\mu(x) dt \quad (3.6)$$

where $\tilde{H}(x, y) = H(x, y) - \int_Y H(x, y) d\nu(y)$.

Remark 3.3 We remark that in Theorems 3.1 and 3.2, the conditions (3.2) and (3.5) hold trivially in case τ is bounded below by a positive constant c (indeed, in this case $\min_x \{|\tau_N(x)|\} \geq cN \geq \log^2 N$ for N sufficiently large).

4 A criterion of CLT for \mathbb{R}^d actions

In the proof of Theorem 3.1, we will use the strategy of [12] except that we replace the Feller Lindenberg CLT for iid random variables used in [12] by a CLT for \mathbb{R}^d actions which are exponentially mixing of all orders. This CLT for such \mathbb{R}^d actions was proven by Björklund and Gorodnik in [10]. Since it is a key tool in our argument, we devote this section to recalling it.

Proposition 4.1 (Theorem 1.5 in [10]) *Let (G_t, Y, ν) be an \mathbb{R}^d action which is exponentially mixing of all orders. Let $(\mathbf{m}_L)_{L \in \mathbb{R}}$ be a sequence of non-negative measures on \mathbb{R}^d . For $t \in \mathbb{R}^d$, let $\mathbf{A}_t \in C^1(Y)$ be a family of functions*

satisfying: $v(\mathbf{A}_t) = 0$ for every $t \in \mathbb{R}^d$ and $\sup_{t \in \mathbb{R}^d} \|\mathbf{A}_t\|_{C^1(Y)} < +\infty$. For any $L \in \mathbb{R}$, let $\mathcal{S}_L(y) := \int_{\mathbb{R}^d} \mathbf{A}_t(G_t y) d\mathfrak{m}_L(t)$. Suppose that

- (a) $\lim_{L \rightarrow \infty} \mathfrak{m}_L(\mathbb{R}^d) = \infty$.
- (b) For each $r \in \mathbb{N}$, $r \geq 3$, and each $K > 0$,

$$\lim_{L \rightarrow \infty} \int \mathfrak{m}_L^{r-1} \left(B(t, K \ln \mathfrak{m}_L(\mathbb{R}^d)) \right) d\mathfrak{m}_L(t) = 0,$$

where $B(t, v)$ denotes a ball in \mathbb{R}^d of radius $v > 0$ centered at t .

- (c) There exists $\sigma^2 = \sigma^2(\mathbf{A}_t) \geq 0$ so that $\lim_{L \rightarrow \infty} V_L = \sigma^2$, where

$$V_L := \int \mathcal{S}_L^2(y) d\nu(y) = \iiint \mathbf{A}_{t_1}(G_{t_1} y) \mathbf{A}_{t_2}(G_{t_2} y) d\mathfrak{m}_L(t_1) d\mathfrak{m}_L(t_2) d\nu(y).$$

Then $\mathcal{S}_L(\cdot)$ converges as $L \rightarrow \infty$ to normal distribution with zero mean and variance σ^2 .

Proposition 4.1 is proven in [10, Theorem 1.5] in case \mathbf{A}_t does not depend on t . Since the proof directly extends to the case of t -dependent observables \mathbf{A}_t (with uniform C^1 norm), we do not repeat it here. (The proof uses the method of moments (cumulants). This method requires showing that

$$\int \mathcal{S}_L^p(y) d\nu(y) = \begin{cases} o_{L \rightarrow \infty}(1) & \text{if } p \text{ is odd} \\ (2m-1)!! \sigma^{2m} (1 + o_{L \rightarrow \infty}(1)) & \text{if } p = 2m \text{ is even.} \end{cases} \quad (4.1)$$

The proof of (4.1) proceeds by expanding

$$\int \mathcal{S}_L^p(y) d\nu(y) = \int \dots \int_{t_1, \dots, t_p} \left(\int_Y \left[\prod_j \mathbf{A}_{t_j}(G_{t_j} y) \right] d\mu(y) \right) d\mathfrak{m}_L(t_1) \dots d\mathfrak{m}_L(t_p) \quad (4.2)$$

and showing that the main contribution to (4.2) comes from the region where (t_1, \dots, t_{2m}) can be partitioned into m pairs so that the points in each pair are close to one another but all pairs are far apart. The contribution of points (t_1, \dots, t_{2m}) that do not satisfy this property is estimated by exponential mixing and that estimate only uses the fact that the functions \mathbf{A}_t have uniformly bounded norms. Thus the proof extends to the case when \mathbf{A}_t depends on t .)

In the case of discrete (T, T^{-1}) transformations, we will only need Proposition with $L \in \mathbb{N}$. In this case, we will replace L by N and write $\mathfrak{m}_N, \mathcal{S}_N$, etc.

5 The CLT for skew products

5.1 A quenched CLT

In this section, we use Proposition 4.1 to derive a quenched CLT (Lemma 5.2). In the next section, we will use this quenched CLT to prove Theorems 3.1 and 3.2.

It is convenient to make the following definition. Let $F \in C^r(X \times Y, \zeta)$ be a skew product of the form $F(x, y) = (fx, g(x, y))$. Thus we assume that f preserves a probability measure μ on X and for each x , $g(x, \cdot)$ preserves a probability measure ν on Y , so that F preserves the measure $\zeta = \mu \times \nu$.

Definition 5.1 F satisfies a quenched CLT on C^r if for each function $H \in C^r(X \times Y)$ satisfying (5.1) there exist a constant $\sigma(H)$ and sets $X_N \subset X$ such that $\lim_{N \rightarrow \infty} \mu(X_N) = 1$ and for each $x_N \in X_N$ the sequence of random variables $\frac{H_N(x_N, y)}{\sqrt{N}}$, where y is distributed according to ν , converges in law as $N \rightarrow \infty$ to the normal random variable with zero mean and variance $\sigma^2(H)$.

In this definition, we regard the fiber Y as the phase space of a random dynamical system and X as a source of external noise. From this point of view the *quenched limit theorem* means the theorem which holds for a typical but fixed realization of the noise, while *annealed limit theorem* is the limit theorem when both x and y are assumed random.

Our first step is to obtain a quenched CLT (Lemma 5.2) under the assumptions of Theorem 3.1. Thus we suppose that $f \in C^r(M)$, $\tau : M \rightarrow \mathbb{R}^d$ and (G_t, Y, ν) satisfy the assumptions of Theorem 3.1. Let $H \in C^r(M \times Y)$ be such that

$$\int H(x, y) d\nu(y) = 0 \quad (5.1)$$

for each $x \in M$. Given $x \in M$, we define the measure $\mathfrak{m}_N(x)$ and the observable $\mathbf{A}_{t,x}$ for all $t \in \mathbb{R}^d$ as

$$\mathfrak{m}_N(x) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \delta_{\tau_n(x)}, \quad \mathbf{A}_{t,x}(y) = \frac{1}{\#\{n \leq N : \tau_n(x) = t\}} \sum_{n \leq N : \tau_n(x) = t} H(f^n x, y). \quad (5.2)$$

Then

$$S_N = S_{N,x}(y) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} H(f^n x, G_{\tau_n(x)} y) = \frac{1}{\sqrt{N}} H_N(x, y),$$

where H_N is the ergodic sum of H .

One may think about \mathfrak{m}_N as the rescaled local time of the deterministic random walk τ_N . It is enough to verify the conditions of Proposition 4.1:

Lemma 5.2 *Assume that (5.1) holds and that the assumptions of Theorems 3.1 are satisfied. Then the number*

$$\sigma^2(H) := \sum_{k=-\infty}^{\infty} \int_M \int_Y H(x, y) H(f^k x, G_{\tau_k(x)} y) d\nu(y) d\mu(x)$$

is non-negative and finite. Furthermore, there are subsets $\hat{X}_N \subset M$ such that

$$\lim_{N \rightarrow \infty} \mu(\hat{X}_N) = 1$$

and for any sequence $x_N \in \hat{X}_N$ the measures $\{\mathfrak{m}_N(x_N)\}$ and the functions $(\mathbf{A}_{t,x_N})_{t \in \mathbb{R}^d}$ defined by (5.2) satisfy the conditions of Proposition 4.1, with $\sigma^2 = \sigma^2(H)$ in part (c).

The rest of Sect. 5.1 contains the proof of Lemma 5.2. In Sect. 5.2 we will show how Lemma 5.2 implies Theorems 3.1 and 3.2.

To prove Lemma 5.2, we need to check properties (a)–(c) of Proposition 4.1.

Property (a) is clear since for every $x \in M$, $\mathfrak{m}_N(x)(\mathbb{R}^d) = \sqrt{N}$. Other properties are less obvious and will be checked in separate subsections below.

5.1.1 Proof of Property (b)

Let

$$X_N = \left\{ x \in M : \text{Card}\{n : |n| < N \text{ and } \|\tau_n(x)\| \leq \ln^{1+\varepsilon/2} N\} \geq N^{0.23} \right\}, \quad (5.3)$$

where ε is from (3.2).

Lemma 5.3 *If τ satisfies (3.2), then $\lim_{N \rightarrow \infty} N\mu(X_N) = 0$.*

Proof First observe that for large N

$$X_N \subset X_N^* := \{x : \mathcal{L}(x, N) \geq N^{0.22}\},$$

where

$$\mathcal{L}(x, N) = \text{Card}\{n : N^{0.21} < |n| < N, \|\tau_n(x)\| \leq \ln^{1+\varepsilon/2} N\}.$$

Next, note that if τ satisfies (3.2), then for every $n \geq N^{0.21}$, we have

$$\mu(\|\tau_n\| < \ln^{1+\varepsilon/2} N) \leq \mu(\|\tau_n\| < \ln^{1+\varepsilon} n) < C|n|^{-5}$$

for N sufficiently large, where C is the constant from (3.2). We conclude by the Markov inequality that

$$\begin{aligned}\mu(X_N) &\leq \mu(X_N^*) \leq N^{-0.22} \mu(\mathcal{L}(x, N)) \\ &= N^{-0.22} \sum_{n: N^{0.21} < |n| < N} \mu(\|\tau_n\| < \ln^{1+\varepsilon/2} N) < CN^{-1.06}\end{aligned}$$

for N sufficiently large, where C is the constant from (3.2). \square

Lemma 5.4 *There are sets $\tilde{X}_N \subset M$ such that $\mu(\tilde{X}_N) \rightarrow 1$ and for all $x_N \in \tilde{X}_N$ the measures $\mathfrak{m}_N(x_N)$ satisfy property (b).*

Proof Let $\tilde{X}_N = \{x : f^n x \notin X_N \text{ for all } n = 1, \dots, N\}$. By Lemma 5.3

$$\mu(\tilde{X}_N) \geq 1 - N\mu(X_N) \rightarrow 1$$

as $N \rightarrow \infty$. Thus for each K we have that for N large enough and for each $x \in \tilde{X}_N$

$$\begin{aligned}&\int \mathfrak{m}_N^{r-1}(x)(B(t, K \ln N)) d\mathfrak{m}_N(x)(t) \\ &= \frac{1}{N^{r/2}} \sum_{n=0}^{N-1} \text{Card}^{r-1}\{j < N : \|\tau_j(x) - \tau_n(x)\| \leq K \ln N\} \\ &\leq \frac{1}{N^{r/2}} \sum_{n=0}^{N-1} \text{Card}^{r-1}\{j < N : \|\tau_{j-n}(f^n x)\| \leq \ln^{1+\varepsilon/2} N\} \\ &\leq N^{0.23(r-1)-\frac{r}{2}+1} = N^{0.77-0.27r} \rightarrow 0.\end{aligned}$$

Here, in the last line we used that $x \in \tilde{X}_N$ and that $r \geq 3$. Property (b) follows. \square

5.1.2 Property (c)

Note that by definition of \mathcal{S}_N ,

$$V_N(x) = \frac{1}{N} \int \mathcal{S}_N^2(x, y) d\nu(y) = \frac{1}{N} \sum_{n_1, n_2=0}^{N-1} \sigma_{n_1, n_2}(x), \quad (5.4)$$

where

$$\sigma_{n_1, n_2}(x) = \int H(f^{n_1}x, G_{\tau_{n_1}(x)}y) H(f^{n_2}x, G_{\tau_{n_2}(x)}y) d\nu(y). \quad (5.5)$$

Notice that since ν is G_t invariant,

$$\sigma_{n_1, n_2}(x) = \sigma_{0, n_2 - n_1}(f^{n_1}x). \quad (5.6)$$

To prove property (c) we need to show that $\sigma^2(H)$ is indeed finite and that there exist subsets $\hat{X}_N \subset X$ of measure close to 1 such that for any sequence $x_N \in \hat{X}_N$ we have $\lim_{N \rightarrow \infty} V_N(x_N) = \sigma^2(H)$. We first study $\sigma^2(H)$. We have

$$\begin{aligned} \int_M V_N(x) d\mu(x) &= \frac{1}{N} \sum_{n_1, n_2=0}^{N-1} \int_M \sigma_{n_1, n_2}(x) d\mu(x) \\ &= \sum_{k=-N+1}^{N-1} \frac{N-|k|}{N} \int H(x, y) H(f^k x, G_{\tau_k(x)} y) d\mu(x) d\nu(y) \\ &= \sum_{k=-N+1}^{N-1} \frac{N-|k|}{N} \int_M \sigma_{0, k}(x) d\mu(x) \\ &= \sum_{k=-N+1}^{N-1} \int_M \sigma_{0, k}(x) d\mu(x) - \frac{1}{N} \sum_{k=-N+1}^{N-1} |k| \int_M \sigma_{0, k}(x) d\mu(x). \end{aligned}$$

Due to (5.1) and exponential mixing of G_t , there are constants c, C so that for all x

$$|\sigma_{0, k}(x)| \leq C \|H\|_{C^r}^2 e^{-c\|\tau_k(x)\|}. \quad (5.7)$$

If τ satisfies (3.2), then by (5.7) there are constants $\beta > 1$ and $\bar{C} > 0$ such that

$$\int_M |\sigma_{0, k}(x)| d\mu(x) \leq \bar{C} k^{-\beta} \quad (5.8)$$

(in fact, (5.8) holds for each $\beta < 5$).

In particular, (5.8) implies that the following limit exists

$$\sigma^2(H) := \lim_{N \rightarrow \infty} \int_M V_N(x) d\mu(x) = \sum_{k=-\infty}^{\infty} \int \sigma_{0, k}(x) d\mu(x). \quad (5.9)$$

This shows that $\sigma^2(H)$ is finite.

The next result shows that property (c) holds with probability close to 1.

Lemma 5.5 *Let F be an ergodic (T, T^{-1}) transformation and H be a function satisfying (5.1) and (5.8) with $\beta > 1$. Let V_N be given by (5.4) and $\sigma^2(H)$ be given by (5.9). Then V_N converges to $\sigma^2(H)$ in probability as $N \rightarrow \infty$.*

Lemma 5.5 completes the proof of Lemma 5.2. Indeed given N let ε_N be the smallest number ε such that $\mu(X_{N,\varepsilon}^*) \geq 1 - \varepsilon$ where $X_{N,\varepsilon}^* = \{x \in X : |V_N(x) - \sigma^2(H)| \leq \varepsilon\}$. By Lemma 5.5, $\lim_{N \rightarrow \infty} \varepsilon_N = 0$. Therefore the set $\hat{X}_N = \tilde{X}_N \cap X_{N,\varepsilon_N}^*$, where \tilde{X}_N is from Lemma 5.4, satisfies the conclusions of Lemma 5.2.

Thus it remains to prove Lemma 5.5.

Proof Recall (5.8). Given $\varepsilon > 0$ let k_ε be the smallest number such that

$$\sum_{|k| > k_\varepsilon} \mu(|\sigma_{0,k}|) \leq \varepsilon^2.$$

By ergodicity for large N we have for $|k| \leq k_\varepsilon$

$$\mu \left(x : \left| \frac{1}{N} \sum_{n=k_\varepsilon}^{N-1-k_\varepsilon} \sigma_{0,k}(f^n x) - \mu(\sigma_{0,k}) \right| \geq \frac{\varepsilon}{2k_\varepsilon} \right) \leq \frac{\varepsilon}{2k_\varepsilon}. \quad (5.10)$$

Next, we write

$$\begin{aligned} V_N(x) &= \frac{1}{N} \sum_{k=-N+1}^{N-1} \sum_{n=\max\{-k, 0\}}^{\min\{N-1, N-1-k\}} \sigma_{0,k}(f^n x) \\ &= \frac{1}{N} \sum_{|k| \leq k_\varepsilon} \sum_{n=k_\varepsilon}^{N-1-k_\varepsilon} \sigma_{0,k}(f^n x) + \frac{1}{N} \sum_{|k| > k_\varepsilon} \sum_{n=\max\{-k, 0\}}^{\min\{N-1, N-1-k\}} \sigma_{0,k}(f^n x) \\ &\quad + \frac{1}{N} \sum_{|k| \leq k_\varepsilon} \left(\sum_{n=\max\{-k, 0\}}^{k_\varepsilon-1} + \sum_{n=N-k_\varepsilon}^{\min\{N-1, N-1-k\}} \right) \sigma_{0,k}(f^n x) \\ &=: V'_N + V''_N + V'''_N. \end{aligned}$$

Definition of k_ε and the Markov inequality imply that $\mu(x : |V''_N| \geq \varepsilon) \leq \varepsilon$. Next, $|V'''_N| \leq 2k_\varepsilon \|H\|_\infty^2 / N$ and so V'''_N is negligible. Also by (5.10)

$$\mu \left(x : \left| V'_N - \sum_{|k| \leq k_\varepsilon} \mu(\sigma_{0,k}) \right| > \varepsilon \right) \leq \varepsilon$$

for N sufficiently large. Using the definition of k_ε again we see that

$$\left| \left(\sum_{|k| \leq k_\varepsilon} \mu(\sigma_{0,k}) \right) - \sigma^2(H) \right| \leq \varepsilon^2.$$

Combining the above estimates we obtain

$$\mu(x : |V_N(x) - \sigma^2(H)| \geq 3\varepsilon) \leq 2\varepsilon + \varepsilon^2$$

for N sufficiently large. Since ε is arbitrary, the lemma follows. \square

5.2 From quenched to annealed CLT: proofs of Theorems 3.1 and 3.2

In this section we give the proof of Theorem 3.1 using the quenched CLT (Lemma 5.2). We do not give a separate proof of Theorem 3.2 because the proof of Theorem 3.1 with trivial modifications applies in continuous time.

Recall Definition 5.1. Lemma 5.2 and Proposition 4.1 tell us that F satisfies a quenched CLT on C^r . Thus Theorem 3.1 follows immediately from Lemma 5.6 below.

Lemma 5.6 *Let $F \in C^r(X \times Y, \zeta)$ be a skew product such that f satisfies the CLT on C^r and F satisfies the quenched CLT on C^r . Then F satisfies the CLT on C^r .*

Proof of Lemma 5.6 Split

$$H(x, y) = \tilde{H}(x, y) + \bar{H}(x) \quad \text{where} \quad \bar{H}(x) = \int H(x, y) d\nu(y). \quad (5.11)$$

We will show that for each $\xi \in \mathbb{R}$

$$\lim_{N \rightarrow \infty} \int e^{i\xi H_N(x, y) / \sqrt{N}} d\zeta = e^{-\sigma^2(H)\xi^2/2}$$

where $\sigma^2(H) = \sigma^2(\bar{H}) + \sigma^2(\tilde{H})$, $\sigma^2(\bar{H})$ is the limiting variance in the CLT for \bar{H}_N and $\sigma^2(\tilde{H})$ is the limiting variance in the quenched CLT for \tilde{H} . Let X_N be the sets from Definition 5.1 for \tilde{H} . Split

$$\begin{aligned} & \int_{X \times Y} e^{i\xi H_N(x, y) / \sqrt{N}} d\zeta \\ &= \int_{X_N \times Y} e^{i\xi H_N(x, y) / \sqrt{N}} d\zeta \\ &+ \int_{X_N^c \times Y} e^{i\xi H_N(x, y) / \sqrt{N}} d\zeta. \end{aligned}$$

The second integral converges to 0 since $\lim_{N \rightarrow \infty} \mu(X_N^c) = 0$. On the other hand for $x \in X_N$ $\lim_{N \rightarrow \infty} \int e^{i\xi \tilde{H}_N(x, y) / \sqrt{N}} d\nu(y) = e^{-\xi^2 \sigma^2(\tilde{H})/2}$ uniformly on X_N

(otherwise we could take a subsequence $x_{N_j} \subset X_{N_j}$ such that the distribution of $\frac{\tilde{H}_N(x_{N_j}, \cdot)}{\sqrt{N}}$ does not converge to the normal distribution with zero mean and variance $\sigma^2(\tilde{H})$ which is a contradiction with the assumption that F satisfies a quenched CLT on C^r). Hence

$$\begin{aligned} & \lim_{N \rightarrow \infty} \int_{X_N \times Y} e^{i\xi \frac{H_N(x, y)}{\sqrt{N}}} d\mu(x) d\nu(y) \\ &= \lim_{N \rightarrow \infty} \int_{X_N} e^{i\xi \frac{\tilde{H}_N(x)}{\sqrt{N}}} \left[\int_Y e^{i\xi \frac{\tilde{H}_N(x, y)}{\sqrt{N}}} d\nu(y) \right] d\mu(x) \\ &= \lim_{N \rightarrow \infty} e^{-\xi^2 \sigma^2(\tilde{H})/2} \int_{X_N} e^{i\xi \frac{\tilde{H}_N(x)}{\sqrt{N}}} d\mu(x) \\ &= \lim_{N \rightarrow \infty} e^{-\xi^2 \sigma^2(\tilde{H})/2} \int_X e^{i\xi \frac{\tilde{H}_N(x)}{\sqrt{N}}} d\mu(x) = e^{-\xi^2 \sigma^2(H)/2}, \end{aligned}$$

where the last equation follows from the assumption that f satisfies the CLT on C^r . This completes the proof of the lemma. \square

Proof of Theorem 3.1 The CLT for F follows from Lemma 5.6. So it remains to establish the formula for the variance. If $\sigma^2(\tilde{H}) = 0$, then $\sigma^2(H) = \sigma^2(\tilde{H})$, where

$$\sigma^2(\tilde{H}) = \sum_{k=-\infty}^{\infty} \int_M \int_Y \tilde{H}(x, y) \tilde{H}(f^k x, G_{\tau_k(x)} y) d\nu(y) d\mu(x)$$

by Lemma 5.2. This gives (3.3) and hence completes the proof of the theorem. \square

Remark 5.7 We note that the proof of Theorem 3.1 proceeds in two steps. The first step is to establish a quenched CLT for $\frac{S_N}{\sqrt{V_N(x)}}$ by verifying the conditions of Proposition 4.1. The second step is to prove a weak law of large numbers for $V_N(x)$. In a forthcoming paper we consider (T, T^{-1}) transformations with 1 dimensional action in the fiber, in which case $V_N(x)$ has a nontrivial limit law. This leads to the limit distributions of the form $\mathfrak{S}Z$ where \mathfrak{S} and Z are independent, Z has a standard normal distribution and \mathfrak{S} is the limiting distribution of the quenched standard deviation (cf. [80]). We note that similar distributions appear in limit theorems for \mathbb{Z} and \mathbb{Z}^2 extensions of hyperbolic systems ([96]) in which case \mathfrak{S} stands for the limiting distribution of the local time. However, the approach of [96] is quite different from ours and relies on the symmetry of the infinite measure system with respect to fiber translations.

Part III. CLT for systems of zero entropy, Theorem 1.3 By the discussion in Part I, it is enough to prove Theorem 2.2 (which implies Theorem 1.3 (a)) and Theorem 2.3 (which implies Theorem 1.3 (b)). We provide the proofs of these two theorems in the next two sections.

6 Theorem 2.2

Here we prove Theorem 2.2. We need to show **h1–h3**. **h1** is an immediate consequence of Lemma 2.1 since the entropy of the horocycle flow (h_t) is zero and τ given by (2.3) has mean zero. ($\mu(\tau) = 0$ since μ is invariant under the involution given by $I(q, v) = (q, -v)$, while $\tau \circ I = -\tau$.) We will prove **h2** in Sect. 6.1 and **h3** in Sect. 6.2. Finally, Sect. 6.3 contains the proof of the key technical result: mixing temporal local limit theorem for horocycle windings.

6.1 Reduction of h2 to a mixing local limit theorem

Proof of h2. As in Sect. 5, it suffices to give a proof under the assumption (5.1). Indeed we can split arbitrary H as $H(x, y) = \bar{H}(x) + \tilde{H}(x, y)$ where \bar{H} satisfies (5.1) and use the fact that due to [20], the ergodic integrals of \bar{H} satisfy $\bar{H}_T(x) = O(T^\alpha)$ for some⁷ $\alpha < 1$.

Since $|H_{T_1}(x, y) - H_{T_2}(x, y)| \leq \|H\|_{C^0}|T_1 - T_2|$ it suffices to consider the case when T is an integer. Analogously to (5.2), we define

$$\begin{aligned} \mathfrak{m}_T(x) &= \frac{(\ln T)^{1/4}}{T} \sum_{n=0}^{T-1} \delta_{\tau_n(x)}, \\ A_{t,x}(y) &= \frac{1}{\#\{n \leq N : \tau_n(x) = t\}} \sum_{n \leq T : \tau_n(x) = t} \int_0^1 H(h_{n+s}x, G_s y) ds. \end{aligned}$$

As before we check properties (a)–(c) of Proposition 4.1. Property (a) is immediate as $\|\mathfrak{m}_T\| = (\ln T)^{1/4}$.

To prove (b) and (c) we need some preliminary information. Let

$$\tilde{\mathfrak{m}}_T(x) = \frac{(\ln T)^{1/4}}{T} \int_0^T \delta_{\tau_t(x)} dt.$$

⁷ Let λ_0 be the smallest eigenvalue of the Laplacian on Q . According to [20, Theorem 1.2 and Corollary 1.3] (which relies on [48]) one can take $\alpha = \frac{1 + \sqrt{1 - 4\lambda_0}}{2}$ if $\lambda_0 < \frac{1}{4}$. If $\lambda_0 \geq \frac{1}{4}$ one can take $\alpha = \frac{1}{2} + \varepsilon$ for any $\varepsilon > 0$. The precise value of α is not important for our purposes.

Note that for each set $I \subset \mathbb{R}$, we have

$$\mathfrak{m}_T(x)(I) \leq \tilde{\mathfrak{m}}_T(x)(\tilde{I}), \quad \tilde{\mathfrak{m}}_T(x)(I) \leq \tilde{\mathfrak{m}}_T(x)(\tilde{I}),$$

where \tilde{I} is the unit neighborhood of I . Therefore it suffices to check (b) with $\tilde{\mathfrak{m}}_T$ in place of \mathfrak{m}_T . Thus for part (b), we need to control

$$\begin{aligned} & \tilde{\mathfrak{m}}_T(B(s, K \ln \ln T)) \\ &= (\ln^{1/4} T) \times \frac{\text{mes}(\{t \in [0, T] : |\tau_t(x) - s| \leq K \ln \ln T\})}{T}. \end{aligned}$$

The second factor here is the probability that $\tau_t(x)$ is within distance $K \ln \ln T$ from s when x is fixed and t is uniformly distributed on $[0, T]$. Such results are referred to in [42] as *temporal limit theorems* (in contrast to more classical spatial limit theorems where t is fixed and x is random). The study of temporal limit theorems goes back to [46]. By now there are several systems where the temporal limit theorem is proven (see [18, 42] and the references therein). However, there is only one such system which involves a smooth observable, namely, horocycle windings, and this is the main reason for the choice of the base map in Theorem 2.2. The availability of the temporal limit theorem is crucial in our construction. In fact, we need to extend the temporal limit theorem for horocycle windings ([42, Theorem 5.1]) in two ways. First, the results of [42] concern the probability that $\tau_t(x)$ belong to an interval of length $\sqrt{\ln T}$ whereas we need to consider intervals of unit size (to verify part (b) it is sufficient to handle intervals of size $O(\ln \ln T)$ but for part (c) we need to consider shorter intervals). It is natural to call this extension a *local temporal limit theorem*. Secondly we need a *mixing limit theorem* which claims, roughly speaking, that the values of $\tau_t(x)$ and $\mathbf{G}_t(x)$ are asymptotically independent.

To state our temporal limit theorem we need some notation. Write $x = (q, v) \in X$ and say that q is the *configurational component* of x .

Let $q_0 \in Q$ be an arbitrary reference point and for each $q \in Q$ let Γ_q be a shortest geodesic from q_0 to q . Define $\beta(q) = \int_{\Gamma_q} \lambda$ and let

$$\xi_T(x) = \tau_T(x) - \beta(h_T x) + \beta(x).$$

Equation (2.5) shows that $\xi_T(x)$ is an integral of λ over a curve starting and ending at q_0 , so by (2.2) it is an integer.

Let $\mathfrak{g}_T(x)$ be the configurational component of the geodesic of length $\ln T$ starting at q with speed $-v$. Denote $s_T(x) = \left(\int_{\mathfrak{g}_T(x)} \lambda \right) + \beta(x) - \beta(\bar{x})$, where $\bar{x} = \mathbf{G}_{-\ln T} x$ and \mathbf{G}_t denotes the geodesic flow. \square

Proposition 6.1 *There is a constant $C > 0$ and a zero mean Gaussian density \mathfrak{p} , so that the following statements are true for all $x \in X$.*

(a) *For each $z \in \mathbb{R}$,*

$$\frac{1}{T} \text{mes} \left(t \leq T : \frac{\xi_t - s_T(x)}{\sqrt{\ln T}} \leq z \right) = \int_{-\infty}^z \mathfrak{p}(s) ds + o_{T \rightarrow \infty}(1).$$

(b) *For any set $A \subset X$ whose boundary is a finite union of proper compact submanifolds (with boundary), we have*

$$\frac{\sqrt{\ln T}}{T} \int_0^T 1_{\xi_t(x)=k} 1_{h_t(x) \in A} dt = \mu(A) \mathfrak{p} \left(\frac{k - s_T(x)}{\sqrt{\ln T}} \right) + o_{T \rightarrow \infty}(1), \quad (6.1)$$

where the convergence is uniform when $\frac{k - s_T(x)}{\sqrt{\ln T}}$ varies over a compact set.

(c) *For any $k \in \mathbb{Z}$, we have*⁸

$$\text{mes}(\{t \leq T : \xi_t(x) = k\}) \leq \frac{CT}{\sqrt{\ln T}}. \quad (6.2)$$

Part (a) of Proposition 6.1 is proven in [42]. Parts (b) and (c) are new but they could be established by the methods of [42]. To focus on the new ideas first, we complete the proof of **h2** in Sect. 6.1 and **h3** in Sect. 6.2 assuming Proposition 6.1. Finally, in the separate Sect. 6.3, we prove Proposition 6.1.

Thus we proceed with the proof of property **h2**. Recall, that it remains to verify properties (b) and (c) of Proposition 4.1. Property (b) of Proposition 4.1 reduces to showing that for each K and each $r \geq 3$,

$$\int \tilde{\mathfrak{m}}_T^{r-1}(B(t, K \ln \ln T)) d\tilde{\mathfrak{m}}_T(t) \rightarrow 0.$$

Observe that by Proposition 6.1 (c), for each unit segment $I \subset \mathbb{R}$, $\tilde{\mathfrak{m}}_T(I) \leq C / \ln^{1/4} T$ and hence $\tilde{\mathfrak{m}}_T(B(t, K \ln \ln T)) \leq \frac{C(K) \ln \ln T}{\ln^{1/4} T}$. Thus

$$\begin{aligned} \int \tilde{\mathfrak{m}}_T^{r-1}(B(t, K \ln \ln T)) d\tilde{\mathfrak{m}}_T(t) &\leq \frac{C^{r-1}(K) (\ln \ln T)^{r-1}}{\ln^{(r-1)/4} T} \|\tilde{\mathfrak{m}}_T\|_\infty \\ &\leq \frac{C^{r-1}(K) (\ln \ln T)^{r-1}}{\ln^{\frac{r-2}{4}} T} \rightarrow 0 \end{aligned}$$

since $r > 2$. This implies property (b) of Proposition 4.1.

⁸ Estimates such as (6.2) are often called *anticoncentration inequalities* since (6.2) shows that the probability that $\tau_{(\cdot)}$ belongs to a unit interval is small no matter where this interval is located.

To establish property (c) of Proposition 4.1 it suffices to show that there is a number \mathbf{v} such that

$$\sup_{x \in X} \left| \lim_{T \rightarrow \infty} \frac{\sqrt{\ln T} V_T(x)}{T^2} - \mathbf{v} \right| = 0, \quad (6.3)$$

where

$$V_T(x) = \int \left(\int_0^T H(F_T(x, y)) dt \right)^2 d\nu(y).$$

We have

$$V_T(x) = \sum_{k_1, k_2 \in \mathbb{Z}} \mathcal{I}_{k_1, k_2}(x), \quad (6.4)$$

where

$$\begin{aligned} \mathcal{I}_{k_1, k_2}(x) \\ = \int_0^T \int_0^T 1_{\xi_{t_1} = k_1} 1_{\xi_{t_2} = k_2} \rho(h_{t_1}x, h_{t_2}x, k_2 - k_1 + \beta(q_{t_2}) - \beta(q_{t_1})) dt_1 dt_2, \end{aligned} \quad (6.5)$$

q_t is the configurational component of $h_t(x)$ and

$$\rho(x', x'', s) = \int H(x', y) H(x'', G_s y) d\nu(y). \quad (6.6)$$

Fix a large R and partition the sum in (6.4) into three parts. Let I be sum of the terms where

$$|k_2 - k_1| \leq R, \quad |k_1 - s_T(x)| \leq R\sqrt{\ln T}; \quad (6.7)$$

II be sum of the terms where $|k_2 - k_1| > R$; and III be sum of the terms where

$$|k_2 - k_1| \leq R \quad \text{but} \quad |k_1 - s_T(x)| > R\sqrt{\ln T}.$$

We split our analysis in two parts. Lemma 6.2 says that for large R the main contribution to the variance comes from I , while Lemma 6.3 obtains the asymptotics of the main contribution.

Lemma 6.2 *For each $\delta > 0$ there is $R_0 > 0$ such that for $R \geq R_0$ there exists $T_0 = T_0(R)$ such that for $T \geq T_0$*

$$|II| \leq \frac{\delta T^2}{3\sqrt{\ln T}} \quad \text{and} \quad |III| \leq \frac{\delta T^2}{3\sqrt{\ln T}}.$$

Lemma 6.3 For each $\delta > 0$ there is $R_0 > 0$ such that for $R \geq R_0$ there exists $T_0 = T_0(R)$ such that for $T \geq T_0$

$$\left| \frac{I}{T^2/\sqrt{\ln T}} - \frac{\mathfrak{p}(0)}{\sqrt{2}} \Lambda(H) \right| < \frac{\delta}{3},$$

where⁹

$$\Lambda(H) = \sum_{k \in \mathbb{Z}} \iint \rho(x', x'', \beta(q'') - \beta(q') + k) d\mu(x') d\mu(x'').$$

Since δ is arbitrary, combining Lemmas 6.2 and 6.3, we obtain that

$$\lim_{T \rightarrow \infty} \frac{\sqrt{\ln T} V_T}{T^2} = \frac{\mathfrak{p}(0)}{\sqrt{2}} \Lambda(H)$$

completing the proof (6.3) (with $\mathbf{v} = \frac{\mathfrak{p}(0)}{\sqrt{2}} \Lambda(H)$) and thus verifying **h2**. It remains to prove Lemmas 6.2 and 6.3.

Proof of Lemma 6.2 Since G_t is exponentially mixing and H satisfies (5.1), there are constants C_1, c_1 such that $|\rho(x', x'', t)| \leq C_1 e^{-c_1 t}$ uniformly in x', x'' . Hence using Proposition 6.1(c), we obtain

$$\begin{aligned} |\mathcal{I}| &\leq \sum_{k_1 \in \mathbb{Z}} \sum_{k_2: |k_2 - k_1| \geq R} \\ &\text{mes}(t_1 \in [0, T] : \xi_{t_1}(x) = k_1) \text{mes}(t_2 \in [0, T] : \xi_{t_2}(x) = k_2) \times C_1 e^{-c_1 |k_1 - k_2|} \\ &\leq \frac{C' T}{\sqrt{\ln T}} \sum_{k_1} \text{mes}(t_1 \leq T : \xi_{t_1}(x) = k_1) e^{-c_1 R} \leq \frac{C'' T^2}{\sqrt{\ln T}} e^{-c_1 R}, \end{aligned}$$

where the second inequality is obtained by summing over k_2 for fixed k_1 . Taking R_0 so large that $C'' e^{-c_1 R_0} \leq \frac{\delta}{3}$ we obtain the required estimate on \mathcal{I} .

Similarly, after summing over k_2 we obtain

$$\begin{aligned} |\mathcal{III}| &\leq \frac{C' R T}{\sqrt{\ln T}} \sum_{k_1: |k_1 - s_T(x)| > R \sqrt{\ln T}} \text{mes}(t_1 \leq T : \xi_{t_1}(x) = k_1) \\ &= \frac{C' R T}{\sqrt{\ln T}} \text{mes}(t_1 \leq T : |\xi_{t_1}(x) - s_T(x)| > R \sqrt{\ln T}). \end{aligned} \quad (6.8)$$

⁹ Note that Λ depends on H since ρ depends on H , see (6.6).

By Proposition 6.1(a)

$$\lim_{T \rightarrow \infty} \frac{\text{mes} \left(t_1 \leq T : |\xi_{t_1}(x) - s_T(x)| > R\sqrt{\ln T} \right)}{T} = \mathbb{P}(|\mathcal{N}| > R),$$

where \mathcal{N} is the normal random variable with density \mathfrak{p} . Therefore for large T , (6.8) is smaller than $\left(\frac{T^2}{\sqrt{\ln T}} \right) \times (2C'R \cdot \mathbb{P}(|\mathcal{N}| \geq R))$. Since $\lim_{R \rightarrow \infty} R \cdot \mathbb{P}(|\mathcal{N}| \geq R) = 0$ we can make (6.8) smaller than $\frac{\delta T^2}{3\sqrt{\ln T}}$ by taking R sufficiently large. This completes the proof of the lemma. \square

Proof of Lemma 6.3 The plan of the proof is the following. We need to estimate $\sum_{k_1, k_2} \mathcal{I}_{k_1, k_2}$ where \mathcal{I}_{k_1, k_2} is defined in (6.5) and k_1, k_2 are as in (6.7). First, we choose $R_0 = R_0(\delta)$ so that for all $R > R_0$, an integral \mathfrak{J}_R over the interval $[-R, R]$ is $\delta/100$ close to the limiting integral over \mathbb{R} (see (6.14)). Then, we replace $h_{t_1}x = (q_{t_1}, v_{t_1})$ and $h_{t_2}x = (q_{t_2}, v_{t_2})$ in (6.5) by elements of a fixed ε -net on X . This allows us to decouple the integrals over t_1 and t_2 in (6.5) at the expense of introducing a small relative error of size $\eta = \eta(\delta, R)$ assuming that the diameter of our net $\varepsilon = \varepsilon(\delta, R, \eta)$ is small enough. Once the two integrals in the definition of \mathcal{I}_{k_1, k_2} are decoupled, we can use the mixing local limit theorem for both integrals. Specifically, we will use Proposition 6.1(b) with A being any element of a fixed partition $\{\mathcal{C}_l\}$ of X , where each partition element \mathcal{C}_l contains exactly one element of the ε -net. To quantify Proposition 6.1(b), we choose another small parameter $\bar{\eta} = \bar{\eta}(\delta, R, \eta, \varepsilon)$ and then a large $T_0 = T_0(\delta, R, \eta, \varepsilon, \bar{\eta})$ so that the term $o_{T \rightarrow \infty}(1)$ in Proposition 6.1(b) is smaller than $\bar{\eta}$. We will obtain a Riemann sum over k_1 which we will replace by a Riemann integral to complete the proof.

Now we give the details of the proof. For an $\varepsilon = \varepsilon(\delta, R, \eta)$, divide X into cubes $\{\mathcal{C}_l\}$ with diameter smaller than ε . Let $x_l = (q_l, v_l)$ be the center of \mathcal{C}_l . Next decompose $\mathcal{I}_{k_1, k_2} = \sum_{l_1, l_2} \mathcal{I}_{k_1, k_2, l_1, l_2}$, where

$$\begin{aligned} \mathcal{I}_{k_1, k_2, l_1, l_2} &= \int_0^T \int_0^T 1_{\xi_{t_1}(x) = k_1} 1_{\mathcal{C}_{l_1}}(h_{t_1}x) 1_{\xi_{t_2}(x) = k_2} 1_{\mathcal{C}_{l_2}}(h_{t_2}x) \\ &\quad \times \rho(h_{t_1}x, h_{t_2}x, k_2 - k_1 + \beta(q_{t_2}) - \beta(q_{t_1})) dt_1 dt_2. \end{aligned}$$

Using uniform continuity of $\rho(x', x'', t)$ on the set $|t| \leq 2R$, we see that for any $\eta > 0$ we can take ε so small that for all $(x', x'') \in \mathcal{C}_{l_1} \times \mathcal{C}_{l_2}$ and for all k_1, k_2 satisfying $|k_1 - k_2| \leq R$, we have

$$|\rho(x', x'', k_2 - k_1 + \beta(q_{t_2}) - \beta(q_{t_1})) - \rho_{k_1, k_2, l_1, l_2}| \leq \eta,$$

where

$$\rho_{k_1, k_2, l_1, l_2} = \rho(x_{l_1}, x_{l_2}, k_2 - k_1 + \beta(q_{l_2}) - \beta(q_{l_1})).$$

Therefore, we have

$$\begin{aligned} & \left| \mathcal{I}_{k_1, k_2, l_1, l_2} - \int_0^T \int_0^T 1_{\xi_{l_1}(x)=k_1} 1_{\mathcal{C}_{l_1}}(h_{l_1}x) 1_{\xi_{l_2}(x)=k_2} 1_{\mathcal{C}_{l_2}}(h_{l_2}x) \rho_{k_1, k_2, l_1, l_2} dt_1 dt_2 \right| \\ &= |\mathcal{I}_{k_1, k_2, l_1, l_2} - J_{k_1, l_1} J_{k_2, l_2} \rho_{k_1, k_2, l_1, l_2}| \leq \eta J_{k_1, l_1} J_{k_2, l_2} =: \kappa_{k_1, l_1, k_2, l_2} \end{aligned} \quad (6.9)$$

where

$$J_{k, l} = \text{mes}(t \in [0, T] : \xi_t(x) = k, h_t x \in \mathcal{C}_l).$$

Next, Proposition 6.1(b) implies that for any fixed $\bar{\eta}$, R , and partition $\{\mathcal{C}_l\}$, we can take T so large so that for all k with $\left| \frac{k - s_T(x)}{\sqrt{\ln T}} \right| \leq R$,

$$|J_{k, l} - J_{k, l, 1}| < J_{k, l, 2}, \quad (6.10)$$

where

$$J_{k, l, 1} = \frac{T}{\sqrt{\ln T}} \mu(\mathcal{C}_l) \mathfrak{p} \left(\frac{k - s_T(x)}{\sqrt{\ln T}} \right), \quad J_{k, l, 2} = \frac{T}{\sqrt{\ln T}} \bar{\eta}. \quad (6.11)$$

That is, $J_{k, l, 1}$ is the leading term and $J_{k, l, 2}$ is an error term. Then by (6.10) and by the triangle inequality, we have

$$|J_{k_1, l_1} J_{k_2, l_2} - J_{k_1, l_1, 1} J_{k_2, l_2, 1}| \leq J_{k_1, l_1, 2} J_{k_2, l_2} + J_{k_1, l_1, 1} J_{k_2, l_2, 2} \leq C \frac{T^2 \bar{\eta}}{\ln T}.$$

Multiplying by $\rho_{k_1, k_2, l_1, l_2}$ and then summing for (k_1, k_2) as in (6.7) and for l_1, l_2 , we obtain

$$\left| \sum_{k_1, l_1, k_2, l_2} J_{k_1, l_1} J_{k_2, l_2} \rho_{k_1, k_2, l_1, l_2} - \sum_{k_1, l_1, k_2, l_2} J_{k_1, l_1, 1} J_{k_2, l_2, 1} \rho_{k_1, k_2, l_1, l_2} \right| \leq (6.12)$$

$$\|\rho\|_\infty \text{Card}(\{(k_1, k_2)\}) \text{Card}^2(\{\mathcal{C}_l\}) C \frac{T^2 \bar{\eta}}{\ln T} \leq \bar{C} R^2 \text{Card}^2(\{\mathcal{C}_l\}) \frac{T^2}{\sqrt{\ln T}} \bar{\eta}. \quad (6.13)$$

We choose $\bar{\eta}$ small so that the right hand side of (6.13) is smaller than $\frac{\delta T^2}{10\sqrt{\ln T}}$.

To study the leading term, i.e. the second sum in (6.12), let us note that by (6.11),

$$\mathfrak{R} := \frac{\ln T}{T^2} \sum_{k_1, l_1, k_2, l_2} J_{k_1, l_1, 1} J_{k_2, l_2, 1} \rho_{k_1, k_2, l_1, l_2}$$

is expressed by

$$\mathfrak{R} = \sum_{k_1, l_1, k_2, l_2} \mu(C_{l_1}) \mu(C_{l_2}) \mathfrak{p} \left(\frac{k_1 - s_T(x)}{\sqrt{\ln T}} \right) \mathfrak{p} \left(\frac{k_2 - s_T(x)}{\sqrt{\ln T}} \right) \rho_{k_1, k_2, l_1, l_2}.$$

Next, decreasing ε if necessary¹⁰ and then taking T sufficiently large, we can approximate the Riemann sum \mathfrak{R} by the corresponding integral. Thus we can achieve that

$$\left| \frac{\mathfrak{R}}{\sqrt{\ln T}} - \mathfrak{I}_R \right| \leq \frac{\delta}{10}$$

where

$$\begin{aligned} \mathfrak{I}_R = & \left(\int_{-R}^R \mathfrak{p}^2(z) dz \right) \sum_{|k| < R} \iint \rho(x', x'', k + \beta(q(x'')) \right. \\ & \left. - \beta(q(x')) d\mu(x') d\mu(x'') \right). \end{aligned}$$

Here, as before, $\sqrt{\ln T}$ appears in the denominator to account for the summation with respect to k_1 , the sum with respect to k_2 is rewritten as a sum over $k = k_2 - k_1$, and we have used the fact that due the first constraint in (6.7),

$$\mathfrak{p} \left(\frac{k_2 - s_T(x)}{\sqrt{\ln T}} \right) = \mathfrak{p} \left(\frac{k_1 - s_T(x)}{\sqrt{\ln T}} \right) + O \left(\frac{R}{\sqrt{\ln T}} \right).$$

Next, we verify that the error terms $\kappa_{k_1, l_1, k_2, l_2}$ defined in (6.9) are negligible. To this end, we claim that

$$\sum_{k_1, l_1, k_2, l_2} |\kappa_{k_1, l_1, k_2, l_2}| \leq \frac{CR^2 T^2}{\sqrt{\ln T}} \eta$$

uniformly in ε . Indeed, each term in the sum satisfies

$$|\kappa_{k_1, l_1, k_2, l_2}| \leq \frac{CT^2 \eta}{\ln T} \mu(C_{l_1}) \mu(C_{l_2}) \leq \frac{\bar{C} \varepsilon^6 T^2 \eta}{\ln T}$$

¹⁰ Recall that ε is the diameter of $\{C_l\}$.

and the number of terms is of order

$$O\left(R^2\sqrt{\ln T} \operatorname{Card}^2(\{\mathcal{C}_l\})\right) = O\left(\frac{R^2\sqrt{\ln T}}{\varepsilon^6}\right).$$

Therefore choosing η sufficiently small, we get

$$|\kappa_{k_1, l_1, k_2, l_2}| \leq \frac{T^2\delta}{10\sqrt{\ln T}}.$$

Finally, using the fact that for Gaussian densities

$$\int_{-\infty}^{\infty} \mathfrak{p}^2(z) dz = \frac{\mathfrak{p}(0)}{\sqrt{2}},$$

we obtain $\lim_{R \rightarrow \infty} \mathfrak{I}_R = \frac{\mathfrak{p}(0)}{\sqrt{2}} \Lambda(H)$. Thus choosing R_0 so large that for all $R > R_0$,

$$\left| \mathfrak{I}_R - \frac{\mathfrak{p}(0)}{\sqrt{2}} \Lambda(H) \right| \leq \frac{\delta}{100} \quad (6.14)$$

completes the proof of the lemma. \square

We have finished the proofs of Lemmas 6.2 and 6.3. The proof of **h2** is complete. \square

6.2 Variance

Proof of h3. Recalling the definition of ρ we can rewrite

$$\begin{aligned} \Lambda(H) &= \sum_k \iiint H(x', y) H(x'', G_{k+\beta(q(x''))-\beta(q(x'))} y) d\mu(x') d\mu(x'') d\nu(y) \\ &= \sum_k \iiint H(x', G_{\beta(q(x'))} y) H(x'', G_{k+\beta(q(x''))} y) d\mu(x') d\mu(x'') d\nu(y) \\ &= \rho(\mathbf{H}) \end{aligned}$$

where

$$\mathbf{H}(y) = \int_X H(x, G_{\beta(q(x))} y) d\mu(x) \quad (6.15)$$

and

$$\rho(\mathbf{H}) = \sum_k \int \mathbf{H}(y) \mathbf{H}(G_k y) d\nu(y). \quad (6.16)$$

Observe that for each $\mathbf{H} \in C^r(Y)$ there is $H \in C^r(X \times Y)$ such that

$$\mathbf{H}(y) = \int_X H(x, G_{\beta(q(x))}y) d\mu(x).$$

Indeed we can just take $H(x, y) = \psi(q(x))\mathbf{H}(G_{-\beta(q(x))}y)$ where ψ is a probability density supported on a small ball centered at q_0 . (Note that $\beta(q(x))$ is smooth if $d(q(x), q_0)$ is smaller than the injectivity radius of our surface Q .)

Therefore **h3** follows from the result below. \square

Theorem 6.4 *Let G be a diffeomorphism of a compact manifold Y which preserves a smooth measure ν . Assume that (G, ν) is exponentially mixing (of order 2) on $C^s(Y)$. Then $\exists \mathbf{H} \in C^s$ such that $\rho(\mathbf{H}) \neq 0$ where¹¹*

$$\rho(\mathbf{H}) = \sum_{k=-\infty}^{\infty} [\nu(\mathbf{H}(\mathbf{H} \circ G_k)) - (\nu(\mathbf{H}))^2]. \quad (6.17)$$

Proof Call a point $y_0 \in Y$ slowly recurrent if for each A, K there exists $r_0 = r_0(A, K)$ such that for each $r \leq r_0$ we have

$$\nu(B(y_0, r) \cap G_{-k}B(y_0, r)) \leq \frac{\nu(B(y_0, r))}{|\ln r|^A}$$

for $1 \leq k \leq K|\ln r|$. By [36, Lemma 4.13] for exponentially mixing systems almost every y_0 is slowly recurrent. Take such a point y_0 and let $r \leq \frac{r_0(2, K)}{2}$ where K is large enough (see (6.18) below). With these parameters fixed, choose a function ψ such that

- (i) $\text{supp}(\psi) \in B(y_0, r)$;
- (ii) $\nu(\psi) = 0$;
- (iii) $\|\psi\|_{C^0} \leq 1$;
- (iv) $\nu(\psi^2) \geq c_1 \nu(B(y_0, r))$;
- (v) $\|\psi\|_{C^s} \leq c_2 r^{-s}$.

By (ii), $\rho(\psi) = \nu(\psi^2) + 2 \sum_{k=1}^{\infty} \nu(\psi(\psi \circ G_k))$. By (iv) the first term is at least $c_1 \nu(B(y_0, r))$. We will show that the remaining sum is of the lower order if r is small enough. Indeed by exponential mixing and (v), $|\nu(\psi(\psi \circ G_k))| \leq c_3 r^{-2s} \theta^k$. Hence

$$\sum_{k=K|\ln r|}^{\infty} |\nu(\psi(\psi \circ G_k))| \leq \frac{c_1}{10} \nu(B(y_0, r)) \quad (6.18)$$

¹¹ Note that if \mathbf{H} is given by (6.15) where H satisfies (5.1) then $\nu(\mathbf{H}) = 0$ so (6.17) reduces to (6.16).

if K is large enough. Note that $|\psi(y)\psi(G_k y)| \leq 1$ and moreover, this product is zero unless $y \in B(y_0, r) \cap G_{-k}B(y_0, r)$. Since y_0 is slowly recurrent it follows that for $1 \leq k \leq K|\ln r|$

$$|\nu(\psi(\psi \circ G_k))| \leq \nu(B(y_0, r) \cap G_{-k}B(y_0, r)) \leq \frac{\nu(B(y_0, r))}{|\ln r|^2}.$$

Hence by further decreasing r if necessary we get that

$$\sum_{k=1}^{K|\ln r|} |\nu(\psi(\psi \circ G_k))| \leq \frac{c_1}{10} \nu(B(y_0, r)). \quad (6.19)$$

Combining (ii), (6.18) and (6.19) we obtain the result. \square

Remark 6.5 While Theorem 6.4 appears to be new for general exponentially mixing systems, it seems to be well known for all currently known examples of exponentially mixing systems.

We also note that Theorem 6.4 is much stronger than what is needed to prove Theorem 1.3(a). Namely, to prove Theorem 1.3(a) it suffices to produce one flow G which is exponentially mixing of all orders and one function \mathbf{H} such that $\rho(\mathbf{H}) \neq 0$. In particular, one can take $(Y, \nu) = (X, \mu)$ (recall that X is the unit tangent bundle of a compact hyperbolic surface with constant negative curvature) and $G = \mathbf{G}$ —the geodesic flow. Next, take

$$\mathbf{H}(y) = \int_0^1 J(\mathbf{G}_s y) dy \quad \text{where} \quad J(y) = \omega(q(y))(v(y))$$

and ω is a harmonic one form. Then

$$\rho(\mathbf{H}) = \int_Y \int_{-\infty}^{\infty} J(y) J(\mathbf{G}_s y) ds d\nu(y) = 4 \int_Y J^2(y) d\nu(y)$$

where the first identity is obtained by direct computation and the second one is proven in [81, Theorem 2]. In fact, in the case $G = \mathbf{G}$, a much stronger result than Theorem 6.4 is known, namely, the set of \mathbf{H} such that $\rho(\mathbf{H}) = 0$ is a linear subspace of infinite codimension.

Indeed, if $\rho(\mathbf{H}) = 0$ then L^2 Gottschalk–Hedlund Theorem ([23]), implies that \mathbf{H} is an L^2 coboundary, that is there is an L^2 function \mathbf{A} such that $\mathbf{H} = \mathbf{A} - \mathbf{A} \circ \mathbf{G}_1$. Then the Livšic theorem for partially hyperbolic systems ([108, Theorem A]) implies that \mathbf{A} has a continuous (in fact smooth) version. It then follows that the ergodic sums of \mathbf{H} are uniformly bounded, which implies that \mathbf{H} has zero mean with respect to any \mathbf{G} invariant ergodic measure. Since there are uncountably many such measures ([14]), the condition $\rho(\mathbf{H}) = 0$ holds on a subspace of infinite codimension.

6.3 Mixing local limit theorem for geodesic flow

Proof (Proof of Proposition 6.1) Part (a) is [42, Theorem 5.1] but we review the proof as it will be needed for parts (b) and (c). The key idea is to rewrite the temporal limit theorem for the horocycle flow as a central limit theorem for the geodesic flow. To be more precise, let $\mathfrak{h}(x, t)$ and $\mathfrak{g}(x, t)$ denote the configurational component of the horocycle $\mathcal{H}(x, t)$ and the geodesic of length t starting from x . Consider the quadrilateral $\Pi(x, t, T)$ formed by

$$\mathfrak{h}(x, t), -\mathfrak{g}(h_t(x), T), -\mathfrak{h}(\mathbf{G}_{-\ln T}x, t/T), \mathfrak{g}(x, T),$$

where $-$ indicates that the curve is run in the opposite direction. This curve $\Pi(x, t, T)$ is contractible as can be seen by shrinking t and T to zero. Therefore the Stokes Theorem gives

$$\xi_t(x) = \left(\int_0^{\ln T} \tau^*(\mathbf{G}_r h_u \bar{x}) dr \right) + \beta(h_u \bar{x}) - \beta(\bar{x}), \quad (6.20)$$

where $\bar{x} = \mathbf{G}_{-\ln T}x$, $u = t/T$ and $\tau^*(q, v) = \lambda(v)$. Note that if t is uniformly distributed on $[0, T]$ then $u = t/T$ is uniformly distributed on $[0, 1]$. Since the curvature is constant, it follows that $h_u \bar{x}$ is uniformly distributed on $\mathcal{H}(\bar{x}, 1)$. Now part (a) follows from the central limit theorem for the geodesic flow \mathbf{G} .

To prove part (b), write

$$\hat{\tau}_s(y) = \int_0^s \tau^*(\mathbf{G}_r y) dr + \beta(y) - \beta(\mathbf{G}_s y). \quad (6.21)$$

Then by (6.20), we have

$$\frac{\sqrt{\ln T}}{T} \int_0^T \mathbf{1}_{\xi_t(x)=k} \mathbf{1}_{h_t(x) \in A} dt = \sqrt{\ln T} \int_0^1 \mathbf{1}_{\hat{\tau}_{\ln T}(h_u(\bar{x}))=k} \mathbf{1}_{\mathbf{G}_{\ln T}(h_u(\bar{x})) \in A} du.$$

Denote $\mathfrak{t} = \ln T$. Then the claim of part (b) is reduced to showing that

$$\sqrt{\mathfrak{t}} \int_0^1 \mathbf{1}_{\hat{\tau}_{\mathfrak{t}}(h_u \bar{x})=k} \mathbf{1}_{\mathbf{G}_{\mathfrak{t}}(h_u \bar{x}) \in A} du = \mathfrak{p}(k/\sqrt{\mathfrak{t}}) \mu(A) + o_{\mathfrak{t} \rightarrow \infty}(1). \quad (6.22)$$

Next, we claim that if \mathfrak{n} is a smooth measure on X , then

$$\sqrt{\mathfrak{t}} \int_X \mathbf{1}_{\hat{\tau}_{\mathfrak{t}}(\bar{x})=k} \mathbf{1}_{\mathbf{G}_{\mathfrak{t}}(\bar{x}) \in A} d\mathfrak{n} = \mathfrak{p}(k/\sqrt{\mathfrak{t}}) \mu(A) + o_{\mathfrak{t} \rightarrow \infty}(1). \quad (6.23)$$

Note that (6.23) is a special case of the *mixing local limit theorem* proven in [40]. Specifically, with the notation of [40, Definition 3.1] (Mixing local limit

theorem for flows), we choose $u_{\mathcal{V}}$ to be the counting measure on $\mathcal{V} = \mathbb{Z}$ and define the functions $\mathfrak{X} : X \rightarrow \mathbb{R}$, $\mathfrak{Y} : X \rightarrow \mathbb{R}$, $\mathcal{F} = \mathcal{F}_t : X \rightarrow \mathbb{R}$, $\mathfrak{Z} : \mathbb{R} \rightarrow \mathbb{R}$ by

$$\mathfrak{X}(\tilde{x}) = \frac{d\mathfrak{n}}{d\mu}(\tilde{x}), \quad \mathfrak{Y}(\tilde{x}) = 1_{\tilde{x} \in A}, \quad \mathcal{F}(\tilde{x}) = \hat{\tau}_{\mathfrak{t}}(\tilde{x}) - W(t), \quad \mathfrak{Z}(z) = 1_{z=0},$$

where $W = k/\sqrt{t}$ and $W(t) = k$. Noting that by (6.21) and by choosing $\varphi = \tau^*$, \mathcal{F} is indeed of the required form, we conclude that [40, Definition 3.1] gives (6.23). (Minor remark: \mathfrak{Y} is required to be continuous in [40, Definition 3.1]. Our \mathfrak{Y} is not continuous, but it is almost everywhere continuous due to our assumption on A which is sufficient for the mixing local limit theorem to hold, cf. Remarks 2.1 and 3.3 in [40].)

Mixing Local Limit Theorem for Anosov flows is proven in [40]. However, we need to verify that we have $\mathcal{V} = \mathbb{Z}$ in [40, Definition 3.1]. To this end, we need to verify the technical assumption of [40, Proposition 6.1], namely that the observable is “minimal and its linearized group falls into case (B)”.

To explain the exact meaning of this condition, we need some definitions. Let us represent the geodesic flow \mathbf{G} as a suspension flow over a Poincaré section M such that the first return map $\mathcal{T} : M \rightarrow M$ is Markov ([13]). Let $r : M \rightarrow \mathbb{R}_+$ be the first return time, that is \mathbf{G} is a suspension over the base (M, \mathcal{T}) with roof function r . Denote $\bar{\tau}(x) = \hat{\tau}_{r(x)}(x)$ for $x \in M$. Define

$$r_n(x) = \sum_{j=0}^{n-1} r(\mathcal{T}^j x), \quad \bar{\tau}_n(x) = \sum_{j=0}^{n-1} \bar{\tau}(\mathcal{T}^j x). \quad (6.24)$$

For a function $\mathfrak{f} : M \rightarrow \mathbb{R}^d$, denote by $\mathfrak{S}(\mathfrak{f})$ the smallest closed subgroup of \mathbb{R}^d , a translate of which contains the range of \mathfrak{f} and let $\mathfrak{M}(\mathfrak{f}) = \bigcap_{\mathfrak{h} \sim \mathfrak{f}} \mathfrak{S}(\mathfrak{h})$, where \sim stand for the coboundary relation. Now let $\mathfrak{f} : M \rightarrow \mathbb{R}^2$ be defined by $\mathfrak{f}(x) := (\bar{\tau}(x), r(x))$. It is clear that $\mathfrak{S}(\mathfrak{f}) = \mathbb{Z} \times \mathbb{R}$. The condition quoted at the end of the previous paragraph states that $\mathfrak{M}(\mathfrak{f}) = \mathbb{Z} \times \mathbb{R}$. This is proven for a different system in [38, Lemma A.3]. However the same proof works for any system assuming that the following two conditions hold: first, for any positive constant b , the suspension over (M, \mathcal{T}) with roof function $b + r$ is weakly mixing and second, $\bar{\tau}$ is minimal. The first condition follows from the contact property of the geodesic flow (see the proof of Lemma A.2 in [38] which only relies on the contact structure). To verify the second condition assume the contrary, that is, $\bar{\tau}$ is not minimal. In this case there would exist $\mathfrak{h} \sim \bar{\tau}$ and $\ell > 1$ such that $\mathfrak{S}(\mathfrak{h}) = \ell\mathbb{Z}$. As a result the integral of λ from (2.2) over all

periodic orbits would be a multiple of ℓ . This however would contradict the main result of [67]. The contradiction shows that $\bar{\tau}$ is in fact, minimal.¹²

Note that the LHS of (6.22) also has the form of (6.23), however, in the case of (6.22) the initial measure is not smooth, in fact, it is a uniform measure on an unstable curve of the geodesic flow. However, one can deduce (6.22) from (6.23) by the standard argument going back to Margulis' thesis [84] approximating the measures supported on unstable curve by smooth measures. We sketch the argument here for completeness. To simplify the notation we assume that β is continuous (and hence smooth) on $\mathfrak{h}(\bar{x}, 1)$. If it is not the case, we break¹³ $\mathfrak{h}(\bar{x}, 1)$ into several pieces and apply the argument below to each piece. Note that it suffices to show that for each compactly supported Lipschitz probability density ψ on \mathbb{R} and each Lipschitz function ϕ on X we have

$$I_t := \sqrt{t} \int_{\mathbb{R}} \psi(u) 1_{\hat{\tau}_t(h_u \bar{x})=k} \phi(\mathbf{G}_t h_u \bar{x}) du = \mathfrak{p}(k/\sqrt{t}) \mu(\phi) + o_{t \rightarrow \infty}(1) \quad (6.25)$$

since (6.22) follows from (6.25) by approximating $1_{[0,1]}$ and 1_A from above and below by Lipschitz functions. To prove (6.25), take a small ε and consider

$$I_{t,\varepsilon} = \varepsilon^{-2} \sqrt{t} \iint_{\mathbb{R}^3} \psi(u) \psi(s/\varepsilon) \psi(t/\varepsilon) 1_{\hat{\tau}_t(x(u,s,t))=k} \phi(\mathbf{G}_t x(u, s, t)) du ds dt,$$

where $x(u, s, t) = \mathbf{G}_t \tilde{h}_s h_u \bar{x}$ and \tilde{h} is the unstable horocycle flow. On one hand, for each fixed ε the distribution of $x(u, s, t)$ is smooth, whence (6.23) implies

$$I_{t,\varepsilon} = \mathfrak{p}(k/\sqrt{t}) \mu(\phi) + o_{t \rightarrow \infty}(1). \quad (6.26)$$

On the other hand since $x(u, s, t)$ belongs to the weak stable manifold of $h_u \bar{x}$, we conclude that $\mathbf{G}_t h_u \bar{x}$ and $\mathbf{G}_t x(u, s, t)$ are $O(\varepsilon)$ close. Hence there exists a constant $C = C(\psi, \phi)$ such that

$$|1_{\hat{\tau}_t(x(u,s,t))=k} \phi(\mathbf{G}_t x(u, s, t)) - 1_{\hat{\tau}_t(h_u \bar{x})=k} \phi(\mathbf{G}_t h_u \bar{x})| \leq C\varepsilon$$

¹² The minimality of $\bar{\tau}$ is not essential for our argument. If $\bar{\tau}$ was not minimal the argument of Sect. 6.1 would still go through but the summations over $k \in \mathbb{Z}$ would need to be replaced by the summations over $k \in \ell\mathbb{Z}$ for some $\ell > 1$. What is important is that the results of [40] allow us to describe the local distribution of τ_t in all the cases.

¹³ Note that the discontinuity set of β on Q is a finite number of geodesic arcs. Namely let $Q = \mathbb{H}^2/\Gamma$. If q is a discontinuity point of β , then there is $\bar{\gamma} \in \Gamma \setminus \{Id\}$ such that $d(q, q_0) = d(q, \bar{\gamma} q_0) = \min_{\gamma \in \Gamma} d(q, \gamma q_0)$. Since the diameter of Q is finite, the discontinuity set of the map $x \mapsto \beta(q(x))$ on X is contained in a finite number of analytic surfaces transverse to the orbits of h_u .

unless $\mathbf{G}_t x(u, s, t) \in D_\varepsilon$ where D_ε denotes the $C\varepsilon$ neighborhood of the discontinuity set of β . Accordingly, denoting by \mathbf{n}_ε the initial distribution of $x(u, s, t)$ defined by

$$\mathbf{n}_\varepsilon(A) = \iiint_{\mathbb{R}^3} \psi(u) \psi(s/\varepsilon) \psi(t/\varepsilon) A(\mathbf{G}_t \tilde{h}_s h_u \bar{x}) du ds dt,$$

we obtain

$$\begin{aligned} & |I_{t,\varepsilon} - I_t| \\ & \leq \bar{C} \left[\sqrt{t} \mathbf{n}_\varepsilon(x : |\hat{\tau}_t(x) - k| \leq 1) \varepsilon + \sqrt{t} \mathbf{n}_\varepsilon(x : |\hat{\tau}_t(x) - k| \leq 1, \mathbf{G}_t x \in D_\varepsilon) \right]. \end{aligned} \quad (6.27)$$

Since \mathbf{n}_ε is smooth, (6.23) gives

$$\sqrt{t} \mathbf{n}_\varepsilon(x : |\hat{\tau}_t(x) - k| \leq 1) = O(1). \quad (6.28)$$

Next, approximating 1_{D_ε} from above by a Lipschitz function and arguing as before we get that

$$\sqrt{t} \mathbf{n}_\varepsilon(x : |\hat{\tau}_t(x) - k| \leq 1, \mathbf{G}_t x \in D_\varepsilon) = O(\varepsilon). \quad (6.29)$$

Combining (6.27), (6.28) and (6.29) we see that $I_{t,\varepsilon} = I_t + O(\varepsilon)$ where the implied constant depends on the Lipschitz norms of ψ and ϕ . Since ε is arbitrary, (6.25) follows from (6.26). Part (b) of Proposition 6.1 follows.

To prove part (c), we again use the approach of [40] to lift the anticoncentration inequality from discrete to continuous time. Let us represent the geodesic flow \mathbf{G} as a suspension flow over a Poincaré section M such that the first return map $\mathcal{T} : M \rightarrow M$ is Markov ([13]). The approximation arguments of the proof of part (b) also show that it suffices to prove that for any smooth measure \mathbf{n} on M ,

$$\mathbf{n}(x : \hat{\tau}_t(x) = k) \leq C/\sqrt{t} \quad (6.30)$$

holds uniformly in k . Recall the notation $r_n, \bar{\tau}_n$ from (6.24). Let $R = \max(\|r\|_\infty, \|\bar{\tau}\|_\infty)$. Focusing on the last time before time t then the orbit crosses M we see that it suffices to show that there exists \bar{C} so that for all t and m ,

$$\mathbf{n}(\{x \in M : \exists n : t - R < r_n(x) \leq t, \bar{\tau}_n(x) = m\}) \leq \bar{C}/\sqrt{t} \quad (6.31)$$

(indeed, (6.30) follows from (6.31) by summing over m such that $|m - k| \leq R$).

To prove (6.31), we use a *discrete anticoncentration* estimate. Namely, we use the following fact: there exists a constant \hat{C} and a two dimensional Gaussian

density \mathfrak{g} such that for each m, n

$$P_n \leq \frac{\hat{C}}{n} \left[\mathfrak{g} \left(\frac{t - \mathbf{r}n}{\sqrt{n}}, \frac{m}{\sqrt{n}} \right) + \frac{1}{\sqrt{n}} \right], \quad (6.32)$$

where

$$P_n = \mathfrak{n}(x \in M : t - R < r_n(x) \leq t, \bar{\tau}_n(x) = m)$$

and \mathbf{r} is the mean free path (the average time between the crossings of M).

The estimate (6.32) was proven by Pène ([94]). Specifically, the last displayed formula on page 834 of [94] in the special case $k = 0$, $A = B = \bar{M}$ (with the notation of [94]) reduces to (6.32). Now to prove (6.31), it is sufficient to show that $\sum_{n=c_1 t}^{c_2 t} P_n \leq \bar{C}/\sqrt{t}$, where $c_1 = 1/R$ and $c_2 = 1/\min r$. Indeed, the number of returns to M before time t is always between $c_1 t$ and $c_2 t$. Since $\sum_{n=c_1 t}^{c_2 t} \frac{1}{n^{3/2}} \leq C/\sqrt{t}$, it is enough to prove that

$$\sum_{n=c_1 t}^{c_2 t} P'_n \leq \bar{C}/\sqrt{t},$$

where $P'_n = \frac{1}{n} \mathfrak{g} \left(\frac{t - \mathbf{r}n}{\sqrt{n}}, \frac{m}{\sqrt{n}} \right)$. Now let $I_k = [t/\mathbf{r} - 2^k \sqrt{t}, t/\mathbf{r} - 2^{k-1} \sqrt{t}]$, where k ranges over positive integers such that $t/\mathbf{r} - 2^{k-1} \sqrt{t} > c_1 t$. Then

$$\sum_{n \in I_k} P'_n \leq |I_k| \frac{1}{c_1 t} \left[\exp \left(-\frac{\mathbf{r}^2 2^{2k-2} t}{c_2 t} \right) \right] \leq C \frac{1}{\sqrt{t}} 2^k \exp(-c_3 2^{2k}).$$

Summing over k , we obtain $\sum_{n=c_1 t}^{t/\mathbf{r}} P'_n \leq C/\sqrt{t}$. A similar argument gives $\sum_{t/\mathbf{r}}^{n=c_2 t} P'_n \leq C/\sqrt{t}$. This completes the proof of (6.31). \square

7 Theorem 2.3

In this section we prove Theorem 2.3. Recall that μ denotes the Lebesgue measure on \mathbb{T}^m and $R_\alpha : \mathbb{T}^m \rightarrow \mathbb{T}^m$ is defined by $R_\alpha(x) = x + \alpha$.

The main result used in the proof is the following proposition which gives bounds on ergodic averages of the base rotation by α :

Proposition 7.1 *For every $\kappa/2 < r < m$, there exists $d \in \mathbb{N}$ such that for every $\alpha \in \mathbb{D}(\kappa)$, we have:*

D1. for every $\phi \in C^r(\mathbb{T}^{\mathbf{m}}, \mathbb{R})$ with $\mu(\phi) = 0$,

$$\frac{1}{\sqrt{n}} \left\| \sum_{0 \leq k < n} \phi(\cdot + k\alpha) \right\|_2 \rightarrow 0,$$

as $n \rightarrow \infty$.

D2. there is a function $\tau := \tau^{(\alpha)} \in C^r(\mathbb{T}^{\mathbf{m}}, \mathbb{R}^d)$ such that $\mu(\tau) = \mathbf{0}$ and

$$n^{2\kappa\mathbf{m}} \mu \left(\{x \in \mathbb{T}^{\mathbf{m}} : \left| \sum_{0 \leq k < n} \tau(x + k\alpha) \right| < \log^2 n \} \right) \rightarrow 0,$$

as $n \rightarrow \infty$.

Let us show how to prove Theorem 2.3 using the above proposition:

Proof of Theorem 2.3 For $r \in (\kappa/2, \mathbf{m})$ let $d \in \mathbb{N}$ be from Proposition 7.1 and fix $\alpha \in D(\kappa)$. Let $\tau = \tau^{(\alpha)}$ be from D2 and consider $F(x, y) = (x + \alpha, G_{\tau(x)} y)$ where (G_t, Y, ν) is smooth \mathbb{R}^d action that is exponentially mixing of all orders. F has zero entropy by Lemma 2.1 and so **r1** holds.

Property **r2** follows from Theorem 3.1. Namely, by D1 it follows that (3.1) holds for $f = R_\alpha$ and for all $A \in C^r(\mathbb{T}^{\mathbf{m}}, \mathbb{R})$ with $\sigma^2(\cdot)$ identically equal to 0. Moreover property (3.2) follows from D2 by taking \mathbf{m} so that $2\kappa\mathbf{m} \geq 5$. Next, we show **r3**, i.e., that the variance is non identically zero. Let $\tau_k(x) = \sum_{i < k} \tau(x + i\alpha)$. Recall that the last part of Theorem 3.1, namely (3.3) states that for functions H satisfying (5.1), the asymptotic variance is given by

$$\sigma^2(H) = \sum_{k=-\infty}^{\infty} \int_{T^{\mathbf{m}}} \int_Y H(x, y) H(x + k\alpha, G_{\tau_k(x)} y) d\nu(y) d\mu(x). \quad (7.1)$$

We shall use that the map $f(x) = x + \alpha$ satisfies the following: for every $\delta > 0$, and every $x_0 \in \mathbb{T}^{\mathbf{m}}$, if $p = \left(\frac{D(\alpha)}{2\delta} \right)^{1/\kappa}$, then

$$f^j B(x_0, \delta) \cap B(x_0, \delta) = \emptyset \quad \text{for every } |j| \leq p. \quad (7.2)$$

Indeed, if the intersection is non-empty, then by $\alpha \in D(\kappa)$

$$2\delta > ||x_0 - (x_0 + j\alpha)|| = ||j\alpha|| \geq \frac{D(\alpha)}{|j|^\kappa}.$$

Let $\phi \in C^\infty(\mathbb{R})$ be a non-negative function supported on the unit interval. Set $H(x, y) = \phi \left(\frac{|x - x_0|}{\delta} \right) D(y)$, where $D \in C^\infty(Y)$, $\nu(D) = 0$. Note that

H satisfies (5.1) and so its asymptotic variance is given by (7.1). Next, the term in (7.1) corresponding to $k = 0$ equals

$$\begin{aligned} & \int_{\mathbb{T}^m} \int_Y \left[\phi \left(\frac{|x - x_0|}{\delta} \right) D(y) \right]^2 d\nu(y) d\mu(x) \\ &= \int_{\mathbb{T}^m} \phi^2 \left(\frac{|x - x_0|}{\delta} \right) d\mu(x) \cdot \int_Y D(y)^2 d\nu(y). \end{aligned}$$

By a change of variables, the above term is equal to to

$$\|D\|_2^2 \cdot \delta^m \int_{\mathbb{R}^m} \phi^2(|\mathbf{x}|) d\mathbf{x}.$$

Notice that by (7.2), the terms in (7.1) with $0 < |k| \leq p$ are equal to zero since for such k , the function $\phi \left(\frac{|x - x_0|}{\delta} \right) \phi \left(\frac{|f^k x - x_0|}{\delta} \right)$ is identically equal to 0.

For $|k| > p$, notice that for every $x \in \mathbb{T}^m$ by exponential mixing of G

$$\int_Y H(x, y) H(x + k\alpha, G_{\tau_k(x)} y) d\nu(y) \leq C \|D\|_r^2 \cdot e^{-\eta|\tau_k(x)|}. \quad (7.3)$$

If $|\tau_k(x)| \geq \log^2 k$, then the above integral is upper bounded by

$$C' \|D\|_r^2 \cdot k^{-2}.$$

By D2, $\mu(|\tau_k(x)| < \log^2 k) \leq C'' k^{-2\kappa m}$. It follows that

$$\begin{aligned} & \int_{\mathbb{T}^m} \int_Y H(x, y) H(x + k\alpha, G_{\tau_k(x)} y) d\nu(y) d\mu(x) \\ & \leq C' \|D\|_r^2 \cdot k^{-2} \delta^m + C'' k^{-2\kappa m} \|D\|_0^2, \end{aligned} \quad (7.4)$$

where the first term reflects the contribution of points x where $|\tau_k(x)| \geq \log^2 k$ (and the factor δ^m appears since the LHS of (7.3) vanishes unless $x \in \text{supp}(\phi) \subset B(0, \delta)$) and the second term reflects the contribution of points x where $|\tau_k(x)| < \log^2 k$. Therefore (7.1) is equal to

$$\|D\|_2^2 \cdot \delta^m \int_{\mathbb{R}^m} \phi^2(|\mathbf{x}|) d\mathbf{x} + Err,$$

where the error term comes from the terms with $k \neq 0$ and so it satisfies

$$|Err| \leq C' \delta^m \|D\|_r^2 \sum_{|k| \geq p} k^{-2} + C'' \|D\|_0^2 \sum_{|k| \geq p} k^{-2\kappa m}$$

$$\leq C''' \left[\|D\|_r^2 \delta^{\mathbf{m}} p^{-1} + \|D\|_0^2 p^{-2\kappa m + 1} \right]$$

(recall that the correlations vanish for $0 < |k| < p$.)

Since $p = \left(\frac{D(\alpha)}{2\delta}\right)^{1/\kappa}$, by taking δ small enough, we can guarantee that

$$C''' \|D\|_r^2 \delta^{\mathbf{m}} p^{-1} < \frac{1}{3} \cdot \|D\|_2^2 \cdot \delta^{\mathbf{m}} \int_{\mathbb{R}^{\mathbf{m}}} \phi^2(|\mathbf{x}|) d\mathbf{x},$$

and

$$C''' \|D\|_0 p^{-2\kappa m + 1} < \frac{1}{3} \cdot \|D\|_2^2 \cdot \delta^{\mathbf{m}} \int_{\mathbb{R}^{\mathbf{m}}} \phi^2(|\mathbf{x}|) d\mathbf{x}.$$

Therefore the LHS of (7.1) is positive. This finishes the proof. \square

It remains to prove Proposition 7.1:

Proof of Proposition 7.1 We start with property *D1*, which is much simpler. Note that if $\phi(x) = \sum_{k \neq 0} a_k e^{2\pi i \langle k, x \rangle}$ then

$$\phi_N(x) = \sum_{k \neq 0} a_k e^{2\pi i \langle k, x \rangle} \frac{1 - e^{2\pi i N \langle k, \alpha \rangle}}{1 - e^{2\pi i \langle k, \alpha \rangle}}.$$

Therefore

$$\|\phi_N\|_2^2 = \sum_{k \neq 0} |a_k|^2 |A_k(N)|^2 \quad (7.5)$$

where $A_k(N) = \frac{1 - e^{2\pi i N \langle k, \alpha \rangle}}{1 - e^{2\pi i \langle k, \alpha \rangle}}$. A simple calculation gives

$$|A_k(N)| = \left| \frac{1 - e^{2\pi i N \langle k, \alpha \rangle}}{1 - e^{2\pi i \langle k, \alpha \rangle}} \right| = \frac{|\sin(\pi N \langle k, \alpha \rangle)|}{|\sin(\pi \langle k, \alpha \rangle)|}. \quad (7.6)$$

Since $r > \kappa/2$, **Property D1** is a direct consequence of the following:

Lemma 7.2 *For every $\alpha \in \mathbb{D}(\kappa)$ there exists $C > 0$ such that for every $r < \kappa$ and $\phi \in C^r(\mathbb{T}^{\mathbf{m}})$, we have*

$$\|\phi_N\|_2 \leq C N^{1-(r/\kappa)}.$$

Proof Note that

$$|A_k(N)| \leq N \cdot \frac{|\sin(\pi N \langle k, \alpha \rangle)| |\pi \langle k, \alpha \rangle|}{|\sin(\pi \langle k, \alpha \rangle)| |\pi N \langle k, \alpha \rangle|} \leq C_0 \cdot N.$$

Also if $|\langle k, \alpha \rangle| \leq 1$, then $|A_k(N)| \leq \frac{1}{|\sin(\pi \langle k, \alpha \rangle)|} \leq C_0(\pi |\langle k, \alpha \rangle|)^{-1}$. Therefore

$$|A_k(N)|^2 \leq C \min \{ \langle k, \alpha \rangle^{-2}, N^2 \}.$$

Since $\alpha \in \mathbb{D}(\kappa)$, using the above estimate on $|A_k(N)|^2$, we get

$$\|\phi_N\|_2^2 \leq CD(\alpha) \sum_{|k| \leq N^{1/\kappa}} |k|^{2\kappa} |a_k|^2 + C \sum_{|k| \geq N^{1/\kappa}} N^2 |a_k|^2 = C'[I + \mathcal{I}]$$

where

$$\begin{aligned} I &\leq \sum_{|k| \leq N^{1/\kappa}} (|k|^{2r} |a_k|^2) k^{2(\kappa-r)} \leq \sum_{|k| \leq N^{1/\kappa}} (|k|^{2r} |a_k|^2) N^{2(1-\frac{r}{\kappa})} \\ &\leq C \|\phi\|_{C^r}^2 (N^{1-(r/\kappa)})^2, \end{aligned}$$

$$\text{and } \mathcal{I} \leq \sum_{|k| \geq N^{1/\kappa}} (|k|^{2r} |a_k|^2) (N^{1-(r/\kappa)})^2 \leq C \|\phi\|_{C^r}^2 (N^{1-(r/\kappa)})^2. \quad \square$$

So it remains to establish property **D2**. We start with the following lemma:

Lemma 7.3 *Let $\alpha \in \mathbb{D}(\kappa)$. There exists $R_{\mathbf{m}} > 0$ such that for every $N \in \mathbb{N}$ there exists $k_N \in \mathbb{Z}^{\mathbf{m}}$ satisfying:*

$$|\langle k_N, \alpha \rangle| < \frac{1}{4N}, \quad |k_N| \leq R_{\mathbf{m}} N^{1/\mathbf{m}}.$$

Proof Let $\alpha = (\alpha_1, \dots, \alpha_m)$. For $N \in \mathbb{N}$, consider the lattice

$$\mathcal{L}(\alpha, N) = \begin{pmatrix} N^{-1/\mathbf{m}} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & N^{-1/\mathbf{m}} & 0 \\ 0 & \dots & 0 & N \end{pmatrix} \begin{pmatrix} 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & 0 \\ \alpha_1 & \dots & \alpha_{\mathbf{m}} & 1 \end{pmatrix} \mathbb{Z}^{\mathbf{m}+1} \subset \mathbb{R}^{\mathbf{m}+1}$$

The points in this lattice are of the form

$$\begin{aligned} e &= (x, z) \in \mathbb{R}^{\mathbf{m}} \times \mathbb{R} \\ \text{where } x &= \frac{k}{N^{1/\mathbf{m}}}, \quad z = N \cdot (\langle k, \alpha \rangle + m) \text{ and } (k, m) \in \mathbb{Z}^{\mathbf{m}} \times \mathbb{Z}. \end{aligned}$$

Let $R_{\mathbf{m}}$ be such that a ball \mathcal{B} of radius $R_{\mathbf{m}}$ in $\mathbb{R}^{\mathbf{m}}$ has volume $2^{\mathbf{m}+3}$. Then

$$\text{vol}(\mathcal{B} \times [-1/4, 1/4]) \geq 2^{\mathbf{m}+2}.$$

So by Minkowski's Theorem, $\mathcal{L}(\alpha, N)$ contains a non-zero vector (x, z) in $\mathcal{B} \times [-1/4, 1/4]$. Let $x = \frac{k_N}{N^{1/\mathbf{m}}}$ and $z = N(\langle k_N, \alpha \rangle + m)$. Then $|z| \leq 1/4$ and $x \in \mathcal{B}$. This finishes the proof. \square

The following lemma gives a lower bound on the $A_k(N)$ for k of the form k_{2^ℓ} , $\ell \in \mathbb{N}$.

Lemma 7.4 *For every $\alpha \in \mathbb{D}(\kappa)$ let $(k_N)_{N \in \mathbb{N}}$ be the sequence from Lemma 7.3. There exists $c > 0$ such that for every $l \in \mathbb{N}$ and every $N \in [2^l, 2^{l+1}]$, we have*

$$\frac{|A_{k_{2^l}}(N)|}{|k_{2^l}|^r} \geq c \cdot N^{1-r/\mathbf{m}}.$$

Proof By the bound on k_{2^l} from Lemma 7.3 it suffices to show that

$$|A_{k_{2^l}}(N)| \geq c' \cdot N.$$

By Lemma 7.3, $|N \langle k_{2^l}, \alpha \rangle| < 1/2$. Now using the estimate $C^{-1} < \frac{\sin z}{z} < C$ for $z = N \langle k_{2^l}, \alpha \rangle$ and $z = \langle k_{2^l}, \alpha \rangle$ in (7.6), we obtain the result. \square

For $\alpha \in \mathbb{D}(\kappa)$, let $(k_{2^l})_{l \in \mathbb{N}}$ be the sequence from Lemma 7.4. For a real sequence $\{a_l\}_{l \in \mathbb{N}} \subset [-1, 1]$, let $\tau^\alpha((a_l)) = \tau(a_l) : \mathbb{T}^{\mathbf{m}} \rightarrow \mathbb{C}$ be given by

$$(\tau(a_l))(x) = \sum_{l>0} \frac{a_l e^{2\pi i \langle k_{2^l}, x \rangle}}{|k_{2^l}|^r l^2}. \quad (7.7)$$

For $d \in \mathbb{N}$ let $\tau(a_l^{(1)}, \dots, a_l^{(d)}) : \mathbb{T}^{\mathbf{m}} \rightarrow \mathbb{C}^d$ be defined by $(\tau(x))_j = (\tau(a_l^{(j)}))(x)$. Let $\{a_l^{(j)}\}_{j \leq d, l \in \mathbb{N}}$ be random variables uniformly distributed on the unit interval in \mathbb{R}^d and the corresponding probability measure is denoted by $\mathbb{P}_{\bar{a}}$, i.e.

$$\begin{aligned} \mathbb{P}_{\bar{a}}(a_l^{(j)} \in A_{i,\ell}, \text{ for } j \leq d, \ell \in \mathbb{N}) \\ = \prod_{j \leq d, \ell \in \mathbb{N}} \text{Leb}(\{x \in [-1, 1] : a_l^{(j)}(x) \in A_{j,\ell}\}) \end{aligned}$$

Lemma 7.5 *For every $\varepsilon > 0$ there exists $C > 0$ such that for every $x \in \mathbb{T}^{\mathbf{m}}$ and every $N \in \mathbb{N}$,*

$$\mathbb{P}_{\bar{a}}(\|(\tau(\bar{a}))_N(x)\| \leq N^\varepsilon) < \left(\frac{C}{N^{1-r/\mathbf{m}-2\varepsilon}} \right)^d.$$

Proof Since for a fixed x different components of τ are independent, it suffices to consider the case $d = 1$. In this case, τ is given by (7.7). Let l be such that $N \in [2^l, 2^{l+1}]$. We now fix all the a_j for $j \neq l$. Then, since N, x and all frequencies 2^j except 2^l are fixed, we can write (with some $\mathfrak{c} \in \mathbb{C}$ depending on a_j , $j \neq l$ and N),

$$\tau_N(x) = \mathfrak{c} + \frac{a_l A_{k_{2^l}}(N) e^{2\pi i \langle k_{2^l}, x \rangle}}{|k_{2^l}|^r l^2}. \quad (7.8)$$

Let $M = (M_1, M_2) := \frac{1}{|k_{2^l}|^r l^2} \left(\Re \left(A_{k_{2^l}}(N) e^{2\pi i \langle k_{2^l}, x \rangle} \right), \Im \left(A_{k_{2^l}}(N) e^{2\pi i \langle k_{2^l}, x \rangle} \right) \right)$. By Lemma 7.4,

$$|M| = \frac{|A_{k_{2^l}}(N)|}{|k_{2^l}|^r l^2} \geq c \cdot N^{1-r/\mathbf{m}-\epsilon}.$$

Let us WLOG assume that $|M_1| \geq c/2 \cdot N^{1-r/\mathbf{m}-\epsilon}$ (if $|M_2| \geq c/2 \cdot N^{1-r/\mathbf{m}-\epsilon}$ the proof is analogous). It then follows that the measure of $z \in [-1, 1]$ for which $|M_1 \cdot z - \Re(\mathfrak{c})| < N^\epsilon$, is bounded above by $\frac{2}{cN^{1-r/\mathbf{m}-2\epsilon}}$. Since a_l is uniformly distributed on $[-1, 1]$, (7.8) finishes the proof. \square

Now we are ready to define the map τ and hence also finish the proof of **D2**.

Take $d \in \mathbb{N}$ such that $d(1-r/\mathbf{m}-2\epsilon) > 10\kappa m$. Summing the estimates of Lemma 7.5 over N , we obtain that for some $C' > 0$ and every fixed $x \in \mathbb{T}^{\mathbf{m}}$,

$$\mathbb{P}_{\bar{a}} \left(\{ \text{there exists } N \geq n : \|(\tau(\bar{a}))_N(x)\| \leq N^\epsilon \} \right) < \frac{C'}{n^{10\kappa m - 1}}.$$

It follows by Fubini's theorem that

$$(\mathbb{P}_{\bar{a}} \times \mu) \left(\{(\bar{a}, x) : \text{for all } N \geq n, \|(\tau(\bar{a}))_N(x)\| \geq N^\epsilon \} \right) \geq 1 - \frac{C'}{n^{10\kappa m - 1}}.$$

Using Fubini's theorem again, we get that there exists \mathfrak{A}_n with $\mathbb{P}(\mathfrak{A}_n) \geq 1 - \frac{C'}{n^{4\kappa m}}$, such that for every $\bar{a} \in \mathfrak{A}_n$,

$$\mu(\{x : \text{for all } N \geq n, \|(\tau(\bar{a}))_N(x)\| \geq N^\epsilon\}) \geq 1 - \frac{C'}{n^{4\kappa m}}.$$

It is then enough to take $\bar{a} \in \bigcap_{n \geq N_0} \mathfrak{A}_n$ for any fixed N_0 (notice that $\bigcap_{n \geq N_0} \mathfrak{A}_n$ is non-empty if N_0 is large enough). Then the corresponding

$\tau(\bar{a}) : \mathbb{T}^m \rightarrow \mathbb{C}^d = \mathbb{R}^{2d}$ satisfies **D2** (with $2d$ instead of d). This finishes the proof of the proposition. \square

Part IV. Flows with intermediate mixing properties satisfying CLT.

8 Surface flows in the base

8.1 Proofs of Theorems 2.4 and 2.5

In this section we prove Theorems 2.4 and 2.5. The proofs rely on two auxiliary results, Proposition 8.1 and Lemma 8.2, which will be proven later.

Proposition 8.1 *There exists a set $\mathcal{P} \subset \mathbb{T}$ with $\text{Leb}(\mathcal{P}) = 1$ such that for every $\alpha \in \mathcal{P}$ if $(\varphi_t) \in \mathcal{K}(\alpha, \text{logsym}) \cup \mathcal{K}(\alpha, \gamma)$, with $\gamma \in B_{\text{sing}}$, then there exists $\epsilon, \delta > 0$ such that for every $A \in C^3(M)$,*

$$T^\delta \left| \mu \left(\left\{ x \in M : \left| \int_0^T A(\varphi_t x) dt - T\mu(A) \right| = O(T^{1/2-\epsilon}) \right\} \right) - 1 \right| \rightarrow 0, \quad \text{as } T \rightarrow \infty.$$

Lemma 8.2 *Let F_T be a $C^\infty(T, T^{-1})$ flow on $M \times Y$:*

$$F_T(x, y) = (\varphi_T(x), G_{\tau_T}(y)).$$

Suppose that G -action on (Y, v) is exponentially mixing of all orders and that the base flow on M preserves a measure μ and satisfies the following: there exists $C, m > 0$ such that for every $\delta > 0$, we have

$$\varphi_s B(x_0, \delta) \cap B(x_0, \delta) = \emptyset \quad (8.1)$$

for $|s| \in (C\delta, p)$, with $p = \left(\frac{1}{C\delta}\right)^{\frac{1}{m}}$. Then there exists $H \in C^\infty(M \times Y)$ satisfying (5.1) such that $\Sigma^2(H) > 0$ where

$$\Sigma^2(H) = \int_{-\infty}^{\infty} \int_M \int_Y H(x, y) H(f_t x, G_{\tau_t(x)} y) d\nu(y) d\mu(x) dt. \quad (8.2)$$

The proof of Lemma 8.2 is relatively short and will be given in Sect. 8.2. The proof of Proposition 8.1 is longer and will be given in Sect. 8.3.

Proof of Theorem 2.4. By [49] there exists a full measure set \mathcal{P}' such that for every $\alpha \in \mathcal{P}'$, every $(\varphi_t) \in \mathcal{K}(\alpha, \text{logsym})$ is weakly mixing and not mixing.

Let us take $(\varphi_t) \in \mathcal{K}(\alpha, \text{logsym})$, with $\alpha \in \mathcal{P}' \cap \mathcal{P} \cap \mathcal{D}$, where \mathcal{P} is from Proposition 8.1, and

$$\mathcal{D} := \left\{ \alpha \in \mathbb{T} : \exists C, m > 0 \text{ such that } \|k\alpha\| \geq \frac{C}{k^m}, \text{ for } k \in \mathbb{Z} \setminus \{0\} \right\}$$

is the set of Diophantine numbers (the set \mathcal{D} will be used in proving that the variance is not identically zero).

Then (φ_t) is weakly mixing but not mixing and so (F_T) is also not mixing (since the base is not mixing). To see that F_T is weakly mixing we note that if $H \in C^\infty(M \times Y)$, then analogously to (5.11)

$$H(x, y) = \tilde{H}(x, y) + \bar{H}(x),$$

where $\bar{H}(x) = \int_Y H(x, y) d\nu$, and for every $x \in M$, $\int_Y \tilde{H}(x, y) d\nu = 0$. We can WLOG assume that $\int_M \bar{H}(x) d\mu = \int_{M \times Y} H(x, y) d\mu d\nu = 0$.

Then

$$\begin{aligned} & \int_M \int_Y H(x, y) H(F_T(x, y)) d\nu d\mu \\ &= \int_M \int_Y \tilde{H}(x, y) \tilde{H}(F_T(x, y)) d\nu d\mu + \int_M \int_Y \tilde{H}(x, y) \bar{H}(\varphi_T x) d\nu d\mu \\ &+ \int_M \int_Y \bar{H}(x) \tilde{H}(F_T(x, y)) d\nu d\mu + \int_M \int_Y \bar{H}(x) \bar{H}(\varphi_T(x)) d\nu d\mu. \end{aligned}$$

The mixed terms (involving \bar{H} and \tilde{H}) are 0, since for every $x \in M$, $\int_Y \tilde{H}(x, y) d\nu = 0$ and \bar{H} only depends on x . Moreover, the term involving only \tilde{H} goes to 0 as T goes to ∞ by exponential mixing of (G_t) and positivity of τ . Finally, by weak mixing of (φ_t)

$$\frac{1}{R} \int_0^R \left| \int_M \int_Y \bar{H}(x) \bar{H}(\varphi_T(x)) d\nu d\mu \right| dT \rightarrow 0$$

as $R \rightarrow \infty$.

Putting this together, we get

$$\frac{1}{R} \int_0^R \left| \int_M \int_Y H(x, y) H(F_T(x, y)) d\nu d\mu \right| dT \rightarrow 0$$

as $R \rightarrow \infty$.

So (F_T) is weakly mixing. This gives **w1**. Next **w2** follows from Theorem 3.2, since Proposition 8.1 implies that (3.4) is satisfied with $\sigma^2(\cdot)$ identically equal

to 0, and by Remark 3.3 we know that (3.5) holds. Moreover since $\sigma^2(A) = 0$ for all functions which depend only on the base variables, the limiting variance for functions satisfying (5.1) is given by (8.2). Hence the fact that the limiting variance is non-zero follows from Lemma 8.2 once we check that the base flow satisfies (8.1). To check (8.1) we consider the representation of φ as a special flow over a rotation. Thus

$$\varphi_s(\theta, u) = (\theta + n\alpha, u + s - f_n(\theta))$$

for some $|n| \leq \bar{C}|s|$ where f_n is the ergodic sum of f . If $n = 0$ then (8.1) holds since the second coordinates differ by at least $C\delta$. If $n \neq 0$ then the first coordinates differ by at least δ since α is Diophantine. This shows that (8.1) holds and completes the proof of Theorem 2.4. \square

Proof of Theorem 2.5 By [47] there exists a full measure set \mathcal{P}' such that for every $\alpha \in \mathcal{P}'$, every $(\varphi_t) \in \mathcal{K}(\alpha, \gamma)$ is polynomially mixing. Let us take $(\varphi_t) \in \mathcal{K}(\alpha, \gamma)$, with $\alpha \in \mathcal{P}' \cap \mathcal{P} \cap \mathcal{D}$. Then (φ_t) is polynomially mixing. Moreover by Proposition 8.1 for τ , it follows that τ satisfies *polynomial large deviation bounds*:

$$\mu\left(\{x \in M : |\tau_T(x) - T\mu(\tau)| < \epsilon\}\right) \leq C \cdot T^{-\delta}.$$

Therefore by Theorem 4.1(b) in [35]¹⁴, (F_T) is also polynomially mixing. Moreover, the entropy of (φ_t) is zero and so (F_T) is not K . This gives **n1**. Next, **n2** follows from Theorem 3.2 since by Proposition 8.1, (3.4) holds with $\sigma^2(\cdot)$ identically equal to 0 and (3.5) holds by Remark 3.3. Now **n3** follows by Lemma 8.2 similarly to the proof of Theorem 2.4. \square

8.2 Variance

Proof of Lemma 8.2 Let ϕ be a non negative function supported on the unit interval, with $\phi(t) \equiv 1$ for $t \in [1/4, 3/4]$. Set $H(x, y) = \phi\left(\frac{d(x, x_0)}{\delta}\right)D(y)$, where D is a C^∞ observable on Y with $\nu(D) = 0$. Note that $\nu(D) = 0$ implies that H satisfies (5.1).

We split the integral over $(-\infty, \infty)$ into three summands $\Sigma^2(H) = I_1 + I_2 + I_3$, where

$$I_1 = \int_{-C\delta}^{C\delta} \int_M \int_Y H(x, y) H(\varphi_s x, G_{\tau_s(x)} y) d\nu(y) d\mu(x) ds,$$

¹⁴ Although Theorem 4.1(b) in [35] only covers the discrete case, the proof is the same for continuous time, see Remark 4.11 in [35].

$$I_2 = \int_{(-p, -C\delta) \cup (C\delta, p)} \int_M \int_Y H(x, y) H(\varphi_s x, G_{\tau_s(x)} y) d\nu(y) d\mu(x) ds$$

and

$$I_3 = \int_{(-\infty, -p) \cup (p, \infty)} \int_M \int_Y H(x, y) H(\varphi_s x, G_{\tau_s(x)} y) d\nu(y) d\mu(x) ds.$$

Notice that I_2 equals zero as for every $|s| \in (C\delta, p)$,

$$\phi\left(\frac{d(x, x_0)}{\delta}\right) \cdot \phi\left(\frac{d(\varphi_s x, x_0)}{\delta}\right),$$

is identically equal to zero by (8.1). Moreover, since τ is positive, and (G_t) is exponentially mixing, it follows that for any $u \geq p$, $\tau_u(x) \geq c_\tau \cdot u$, and so for some global $C' > 0$,

$$\begin{aligned} |I_3| &\leq C' \cdot \left| \int_{|u|>p} \int_Y D(y) D(G_{\tau_u(x)}(y)) d\nu du \right| \\ &\leq C'' \|D\|_r^2 \cdot \int_{|u|>p} e^{-\eta c_\tau u} du \\ &\leq C''' \|D\|_r^2 \cdot e^{-\eta c_\tau p}. \end{aligned}$$

Finally,

$$\begin{aligned} I_1 &= \int_{-C\delta}^{C\delta} \int_M \int_Y H(x, y) H(\varphi_s x, G_{\tau_s(x)} y) d\nu(y) d\mu(x) ds \\ &= \int_{-C\delta}^{C\delta} \int_M \int_Y H(x, y) H(\varphi_s x, y) d\nu(y) d\mu(x) ds + O(\delta^4) \\ &= \|D\|_2^2 \cdot \int_{-C\delta}^{C\delta} \int_M \phi\left(\frac{d(x, x_0)}{\delta}\right) \cdot \phi\left(\frac{d(\varphi_s x, x_0)}{\delta}\right) d\mu ds + O(\delta^4). \end{aligned}$$

Next, there exists $c'_\phi > 0$

$$\int_{-C\delta}^{C\delta} \int_M \phi\left(\frac{d(x, x_0)}{\delta}\right) \cdot \phi\left(\frac{d(\varphi_s x, x_0)}{\delta}\right) d\mu ds \geq c'_\phi \cdot C^2 \delta \cdot \delta^2 = c'_\phi C^2 \delta^3.$$

If we take δ sufficiently small we can then guarantee that $|I_1| > 2|I_3|$ and $|I_1| > 0$ (the first inequality since we have $p = \left(\frac{1}{C\delta}\right)^{\frac{1}{m}}$). Summarizing $\Sigma^2(H) > 0$. \square

8.3 Proof of Proposition 8.1

We start with some results on deviation of ergodic averages for functions with logarithmic singularities and with power singularities.

For $N \in \mathbb{N}$, let $\theta_{\min, N} := \min_{j < N} \|\theta + j\alpha\|$, where $\theta \in \mathbb{T}$ and $\|z\| = \min\{z, 1-z\}$. In the lemmas below we want to cover the cases of logarithmic and power singularities simultaneously. For roof functions with logarithmic singularities one can get much better bounds (with deviations being a power of \log) but we do not pursue the optimal bounds here since the bounds of the present section are sufficient for our purposes. Let $J \in C^2(\mathbb{T} \setminus \{0\})$ be any function satisfying

$$\lim_{\theta \rightarrow 0^+} \frac{J(\theta)}{\theta^{-\gamma}} = P \quad \text{and} \quad \lim_{\theta \rightarrow 1^-} \frac{J(\theta)}{(1-\theta)^{-\gamma}} = Q, \quad (8.3)$$

for some constants P, Q . Notice that by l'Hopital's rule it follows that any f as in (2.8) satisfies (8.3) (with $P = Q = 0$ if f has logarithmic singularities). Recall that $\gamma \leq 2/5$.

In what follows, let (a_n) denote the continued fraction expansion and (q_n) denote the sequence of denominators of α , i.e. $q_0 = q_1 = 1$ and

$$q_{n+1} = a_{n+1}q_n + q_{n-1}. \quad (8.4)$$

Set

$$J_m(x) := \sum_{0 \leq j < m} J(x + j\alpha).$$

Lemma 8.3 *For every $x \in \mathbb{T}$ and every $n \in \mathbb{N}$,*

$$\left| J_{q_n}(\theta) - q_n \int_{\mathbb{T}} J(\vartheta) d\vartheta \right| = O\left(\theta_{\min, q_n}^{-\gamma}\right).$$

Proof Let $\bar{J}(\theta) = (1 - \chi_{[-\frac{1}{10q_n}, \frac{1}{10q_n}]}(\theta)) \cdot J(\theta)$. Then \bar{J} is of bounded variation. Notice that for $i \neq j$, $0 \leq i, j < q_n$,

$$\|(\theta + i\alpha) - (\theta + j\alpha)\| = \|(i - j)\alpha\| \geq \sup_{0 < s < q_n} \|s\alpha\| \geq \frac{1}{2q_n}.$$

Therefore the cardinality of the set $\{\theta + j\alpha\}_{j < q_n} \cap \left[-\frac{1}{10q_n}, \frac{1}{10q_n}\right]$, is either zero or one. It follows that

$$|\bar{J}_{q_n}(\theta) - J_{q_n}(\theta)| = O\left(\theta_{\min, q_n}^{-\gamma}\right),$$

by the definition of θ_{\min, q_n} . By the Denjoy–Koksma inequality,

$$\left| \bar{J}_{q_n}(\theta) - q_n \int_{\mathbb{T}} \bar{J}(\vartheta) d\vartheta \right| \leq \text{Var}(\bar{J}) = O(q_n^\gamma).$$

Moreover, since $\left| \{\theta + j\alpha\}_{j < q_n} \cap \left[-\frac{10}{q_n}, \frac{10}{q_n} \right] \right| \geq 1$ it follows that $\theta_{\min, q_n} \leq \frac{10}{q_n}$, and so $q_n^\gamma = O(\theta_{\min, q_n}^{-\gamma})$. It remains to notice that

$$\left| \int_{\mathbb{T}} \bar{J} d\vartheta - \int_{\mathbb{T}} J d\vartheta \right| = \int_0^{\frac{1}{10q_n}} J d\vartheta + \int_{1-\frac{1}{10q_n}}^1 J d\vartheta = O(q_n^\gamma / q_n),$$

by the definition of \bar{J} . Since $\gamma < \frac{1}{2}$, the result follows. \square

Lemma 8.4 *Fix $\zeta, C > 0$ and assume that α is such that $\sup_{n \in \mathbb{N}} \frac{q_{n+1}}{q_n^{1+\zeta}} \leq C$ for some $\zeta, C > 0$. Then for every $N \in \mathbb{N}$*

$$\left| J_N(\theta) - N \int_{\mathbb{T}} J(\vartheta) d\vartheta \right| = O\left(N^\zeta \log N \cdot \theta_{\min, N}^{-\gamma}\right).$$

Proof Let $N = \sum_{k \leq M} b_k q_k$, with $b_k \leq a_k$, $b_M \neq 0$, $M = O(\log N)$ be the Ostrovska expansion of N . For every point $\bar{\theta} = \theta + j\alpha$, $j < N$ with $j + q_k < N$, we have that $\bar{\theta}_{\min, q_k} \geq \theta_{\min, N}$. Hence for each such point Lemma 8.3 gives

$$|J_{q_k}(\bar{\theta}) - q_k \int_{\mathbb{T}} J(\vartheta) d\vartheta| = O\left(\theta_{\min, N}^{-\gamma}\right). \quad (8.5)$$

For $k \leq M$ and $0 \leq j < b_k$ let $\theta_{j,k} := \theta + (\sum_{s \leq k-1} b_s q_s + j q_k) \alpha$. Since $N = \sum_{k \leq M} b_k q_k$ it follows that for every $k \leq M$ we have $(\theta_{j,k})_{\min, q_k} \geq \theta_{\min, N}$. Using cocycle identity we can split $J_N(\theta) = \sum_{k \leq M} \sum_{j < b_k} J_{q_k}(\theta_{j,k})$. Then using (8.5) for each $\theta_{j,k}$ (and remembering that $(\theta_{j,k})_{\min, q_k} \geq \theta_{\min, N}$), we get

$$\begin{aligned} \left| J_N(\theta) - N \int_{\mathbb{T}} J(\vartheta) d\vartheta \right| &= O\left(M \cdot \sup_k b_k \cdot \theta_{\min, N}^{-\gamma}\right) \\ &= O\left(\log N \cdot N^\zeta \theta_{\min, N}^{-\gamma}\right), \end{aligned}$$

where we use that $M = O(\log N)$ and (using (8.4))

$$\sup_k b_k \leq \sup_k a_k = O(q_k^\zeta) = O(N^\zeta).$$

This finishes the proof. \square

We now define the full measure set \mathcal{P} from Proposition 8.1. Let $0 < \zeta < 1/1000$ and

$$\mathcal{P} := \{\alpha \in \mathbb{T} : \exists C > 0 \text{ such that } q_{n+1} < Cq_n^{1+\zeta} \text{ for every } n \in \mathbb{N}\}. \quad (8.6)$$

The set \mathcal{P} has full measure by Khintchine's theorem, [70]. Assume now that we fix $(\varphi_t) \in \mathcal{K}(\alpha, \text{logsym}) \cup \mathcal{K}(\alpha, \gamma)$, with $\gamma \in B_{\text{sing}}$ (in particular $\gamma \leq 2/5$). By definition $R_\alpha : \mathbb{T} \rightarrow \mathbb{T}$ is the first return map and f is the first return time. In particular for every $x \in M$ (except the singularity), $x = \varphi_s(\theta)$, where $\theta \in \mathbb{T}$ and $s < f(\theta)$. We will denote this by $x = (\theta, s)$.

Let $c = \inf_{\mathbb{T}} f > 0$. For $T > 0$, we say that $\theta \in \mathbb{T}$ is T -good if the orbit $\{\theta + j\alpha\}_{j \leq \frac{T}{c}}$ does not visit the interval $\left[-\frac{1}{T^{1+1/100}}, \frac{1}{T^{1+1/100}}\right]$. We have the following

Lemma 8.5 *Let $(\varphi_t) \in \mathcal{K}(\alpha, \text{logsym}) \cup \mathcal{K}(\alpha, \gamma)$ be a flow on M . Let*

$$W(T) := \{x = (\theta, s) \in M : \theta \text{ is } T\text{-good}\}.$$

Then there exists $\eta > 0$ such that $T^\eta \mu(W(T)^c) \rightarrow 0$ as $T \rightarrow \infty$.

Proof For an interval $I \subset \mathbb{T}$, let $I^f := \{(\theta, s) : s < f(\theta), \theta \in I\}$. Note that

$$(W(T))^c = \bigcup_{j \leq \frac{T}{c}} I_j^f,$$

where $I_j = \left[-j\alpha - \frac{1}{T^{1+1/100}}, -j\alpha + \frac{1}{T^{1+1/100}}\right]$. Moreover, by the diophantine assumptions on α , all the intervals I_j are pairwise disjoint. Therefore, for $j \neq 0$,

$$\sup_{\theta \in I_j} f(\theta) \leq C \cdot T^{(1+1/100)\gamma}.$$

Hence

$$\mu \left(\bigcup_{0 \neq j \leq \frac{T}{c}} I_j^f \right) \leq CT^{(1+1/100)(\gamma-1)}. \quad (8.7)$$

Moreover, since f satisfies (2.8), for some $\eta > 0$,

$$\mu(I_0^f) = \int_{\left[-\frac{1}{T^{1+1/100}}, \frac{1}{T^{1+1/100}}\right]} f d\text{Leb} \leq T^{-2\eta}. \quad (8.8)$$

for T sufficiently large. Combining (8.7) and (8.8) gives the result. \square

Using the two lemmas above we can prove Proposition 8.1.

Proof of Proposition 8.1 Let \mathcal{P} be as in (8.6) and let $(\varphi_t) \in \mathcal{K}(\alpha, \text{logsym}) \cup \mathcal{K}(\alpha, \gamma)$, with $\gamma \in B_{\text{sing}}$ (in particular $\gamma \leq 2/5$). Let $A \in C^3(M)$ and denote $A_T(x) := \int_0^T A(\varphi_t x) dt$. We will show that there exists $C > 0$ such that for every T , and every $x \in W(T)$, we have

$$|A_T(x) - T\mu(A)| \leq CT^{1/2-1/1000}.$$

This by Lemma 8.5 will finish the proof of the proposition, as $\mu(W(T)) \rightarrow 1$ as $T \rightarrow \infty$. Let $x = (\theta, s) \in W(T)$, i.e. θ is T -good. Then we have in particular that

$$s < f(\theta) \leq CT^{(1+1/100)\gamma} \leq CT^{1/2-1/1000}$$

and

$$|A_T(\theta, s) - A_T(\theta, 0)| < C's \leq \bar{C}'T^{1/2-1/1000}.$$

Therefore, it is enough to show that if $(\theta, 0) \in W(T)$, then

$$|A_T(\theta, 0) - T\mu(A)| \leq C''T^{1/2-1/1000}. \quad (8.9)$$

for some constant $C'' > 0$. For $r > 0$ let $N(\theta, 0, r)$ be such that

$$\varphi_r(\theta, 0) = (\theta + N(\theta, 0, r)\alpha, \bar{r}),$$

i.e. $N(\theta, 0, r)$ is equal to the number of returns to the transversal \mathbb{T} up to time r .

Note that since $c = \min_{\mathbb{T}} f > 0$, we have that the minimal return time is c and so

$$cN(\theta, 0, T) \leq T. \quad (8.10)$$

Therefore $\|\theta + N(\theta, 0, T)\alpha\| \geq \min_{j \leq \frac{T}{c}} \|\theta + j\alpha\| \geq T^{-1-1/100}$, since θ is T -good. In particular by (2.8)

$$f(\theta + N(\theta, 0, T)\alpha) \leq C'''T^{(1+1/100)\gamma}. \quad (8.11)$$

So

$$\begin{aligned}
& \int_0^T A(\varphi_t(\theta, 0))dt - T\mu(A) \\
&= O\left(T^{(1+1/100)\gamma}\right) + \left(\int_0^{N(\theta, 0, T)} A(\varphi_t(\theta, 0))dt - N(\theta, 0, T)\mu(A)\right) \\
&\quad + (T - N(\theta, 0, T))\mu(A).
\end{aligned}$$

Since $\gamma \leq 2/5$, it is enough to bound the second and last term above. It is therefore enough to prove the following: for every θ which is T -good,

$$|T - N(\theta, 0, T)| = O\left(T^{1/2-1/1000}\right), \quad (8.12)$$

and

$$\left|\int_0^{N(\theta, 0, T)} A(\varphi_t(\theta, 0))dt - N(\theta, 0, T)\mu(A)\right| = O\left(T^{1/2-1/1000}\right). \quad (8.13)$$

Since $N(\theta, 0, T)$ is the number of returns up to time T , we have

$$f_{N(\theta, 0, T)}(\theta) \leq T \leq f_{N(\theta, 0, T+1)}(\theta) \leq f_{N(\theta, 0, T)}(\theta) + C'''T^{(1+1/100)\gamma},$$

the last inequality by (8.11). Hence up to an additional negligible error of size $T^{(1+1/100)\gamma}$, it is enough to control

$$|f_{N(\theta, 0, T)}(\theta) - N(\theta, 0, T)|.$$

By (8.10) it follows that $\theta_{\min, N(\theta, 0, T)} \geq T^{-1-1/100}$. Hence Lemma 8.4, the estimate $N(\theta, 0, T) \leq T/c$ and the fact that $\int_{\mathbb{T}} f dLeb = 1$ imply that

$$|f_{N(\theta, 0, T)}(\theta) - N(\theta, 0, T)| \leq O\left(T^{\zeta+(1+1/100)\gamma} \log T\right).$$

Since $\zeta + (1+1/100)\gamma \leq 1/1000 + (1+1/100)2/5 \leq 1/2 - 1/1000$, (8.12) follows.

To prove (8.13) we can WLOG assume that $\mu(A) = 0$. Note that

$$\begin{aligned}
\int_0^{N(\theta, 0, T)} A(\varphi_t(\theta, 0))dt &= \sum_{i=0}^{N(\theta, 0, T)-1} \int_0^{f(\theta+i\alpha)} A(\varphi_s(\theta + i\alpha, 0))ds \\
&= \sum_{i=0}^{N(\theta, 0, T)-1} F(\theta + i\alpha)
\end{aligned}$$

where $F(\theta) = \int_0^{f(\theta)} A(\varphi_s(\theta, 0))ds$. Moreover, $Leb(F) = \mu(A) = 0$ and F is smooth except at 0.

Let \mathbf{p} be the fixed point. We claim that for every $\epsilon > 0$ there exists $\delta > 0$ such that

$$(A(\mathbf{p}) - \epsilon)f(\theta) - 2\|A\|_0\delta^{-1} \leq F(\theta) \leq (A(\mathbf{p}) + \epsilon)f(\theta) + 2\|A\|_0\delta^{-1}. \quad (8.14)$$

Indeed, we write

$$F(\theta) = \int_0^{\delta^{-1}} A(\varphi_s(\theta, 0))ds + \int_{\delta^{-1}}^{f(\theta)} A(\varphi_s(\theta, 0)).$$

The first integral is estimated trivially by $\|A\|_0\delta^{-1}$. If the second integral is non trivial, i.e. $f(\theta) > \delta^{-1}$, this means that for sufficiently small $\delta > 0$, $\varphi_s(\theta, 0)$ is in ϵ^2 neighborhood of the fixed point \mathbf{p} for every $s \in [\delta^{-1}, f(\theta))$. Therefore, $|A(\varphi_s(\theta, 0)) - A(\mathbf{p})| < \epsilon$. In particular,

$$(A(\mathbf{p}) - \epsilon)[f(\theta) - \delta^{-1}] \leq \int_{\delta^{-1}}^{f(\theta)} A(\varphi_s(\theta, 0)) \leq (A(\mathbf{p}) + \epsilon)[f(\theta) - \delta^{-1}].$$

Putting the above together, we get (8.14). Since f satisfies (8.3) and $A \in \mathcal{C}^3$, it follows by (8.14) that

$$\lim_{\theta \rightarrow 0^+} \frac{F(\theta)}{\theta^{-\gamma}} = P' \text{ and } \lim_{\theta \rightarrow 1^-} \frac{F(\theta)}{(1 - \theta)^{-\gamma}} = Q'$$

where $P' = PA(\mathbf{p})$, $Q' = QA(\mathbf{p})$. Thus $F(\cdot)$ also satisfies the assumptions (8.3). So by Lemma 8.4, the fact that θ is T -good and the bound $N(\theta, 0, T) \leq \frac{T}{c}$,

$$\left| \sum_{i=0}^{N(\theta, 0, T)-1} F(\theta + i\alpha) \right| = O\left(T^{\zeta + (1+1/100)\gamma} \log T\right) = O\left(T^{1/2 - 1/1000}\right).$$

This finishes the proof of (8.13) and completes the proof of the proposition. \square

Part V. Non Bernoullicity of T, T^{-1} transformations.

This part is devoted to the proof of Theorem 2.7. Our approach is motivated by [60, 99]. In particular, the statement of the key Proposition 15.2 is similar to the corresponding statements of [60, 99]. However, its proof in our case is different, since the other authors rely on fine properties of the ergodic sums of the cocycle τ while our approach uses exponential mixing in the fiber. We note that in dimension $d \geq 3$ if the dynamics of the fiber is the full \mathbb{Z}^d shift then the corresponding skew product is Bernoulli ([32]). Therefore using properties of the fiber dynamics is essential. We exploit them mainly

by establishing that the relative atoms (on the fiber) of the past partition are points (see Proposition 11.1). The proof uses the geometry of Weyl chambers, see Sect. 11. We emphasize that Proposition 11.1 does not hold if the fiber dynamics is the full \mathbb{Z}^d shift with $d \geq 3$. Another place where the fiber mixing plays a key role is Sect. 13.

Another important ingredient in our approach is the use of Bowen-Hamming distance (see Proposition 12.1) which allows us to handle continuous higher rank actions in the fiber, and so it plays a crucial role in constructing the example of Theorem 1.5. We also emphasize that the systems considered in [60, 99] were shown by the authors not to be *loosely Bernoulli*. We believe that our methods would work also to show non loose Bernoullicity at a cost of rather technical combinatorial considerations as one needs to consider the \bar{f} metric instead of the Hamming metric. To keep the presentation relatively simple and since our goal was to establish smooth K but non Bernoulli examples satisfying CLT, we restrict our attention to only dealing with non Bernoullicity.

We note that the assumption that τ has zero mean in Theorem 2.7 is essential. Indeed, if τ has non-zero mean, then by [35, Theorem 4.1(a)], F is exponentially mixing, and then one can show using the argument of [61] that F is Bernoulli. The details are given in a separate paper [37].

9 Background on symbolic dynamics

Symbolic dynamics provides a powerful tool for studying hyperbolic systems. In this section we briefly recall the facts from symbolic dynamics needed in our proof.

Let $\Omega = \{1, \dots, p\}$ be a finite set with p elements and $A = (A_{ij})$ be an $p \times p$ matrix whose entries are zeroes and ones. The subshift of finite type is the set

$$\Sigma_A = \left\{ \{\omega_j\}_{j=-\infty}^{\infty} \in \Omega^{\mathbb{Z}} : A_{\omega_j \omega_{j+1}} = 1 \text{ for all } j \in \mathbb{Z} \right\}.$$

The shift on Σ_A is defined by $\sigma(\omega)_j = \omega_{j+1}$. We endow Σ_A with the distance

$$d(\omega', \omega'') = 2^{-k} \text{ where } k = \max(\bar{k} \geq 0 : \omega'_j = \omega''_j \forall j : |j| < \bar{k}).$$

Σ_A is topologically mixing iff there is $q > 0$ such that all entries of A^q are positive. We shall assume henceforth that A is such that Σ_A is topologically mixing. Given a Hölder function, $\phi : \Sigma_A \rightarrow \mathbb{R}$ we define its *pressure* by $P(\phi) = \inf[ent_{\mu} + \mu(\phi)]$, where the minimum is taken over all shift invariant measures and ent_{μ} is the entropy of μ . An invariant measure μ is called the *equilibrium measure for ϕ* if $P(\phi) = ent_{\mu} + \mu(\phi)$.

A *word* is a finite sequence $\bar{\omega}_0, \dots, \bar{\omega}_{n-1}$ such that $A_{\bar{\omega}_j \bar{\omega}_{j+1}} = 1$ for all $0 \leq j \leq n-1$. The set

$$\mathcal{D}(\bar{\omega}_0, \dots, \bar{\omega}_{n-1}) = \{\omega \in \Sigma_A : \omega_j = \bar{\omega}_j \ \forall j \in [0, n-1]\}$$

is called a *cylinder of length n*. An invariant measure μ is called *Gibbs with potential ϕ* if there is a constant $K > 0$ such that for each $n \in \mathbb{Z}$ for each cylinder \mathcal{D} of length n for each $\omega \in \mathcal{D}$

$$\frac{1}{K} \leq \frac{\mu(\mathcal{D})}{e^{\phi_n(\omega) - nP(\phi)}} \leq K. \quad (9.1)$$

It is known (see e.g. [92, Chapter 3]) that ϕ is a Hölder function on Σ_A then it has unique equilibrium state μ_ϕ which is also a Gibbs measure with potential ϕ .

The Gibbs property (9.1) implies the following important *quasi independence* estimate. If μ is a Gibbs measure with a Hölder potential, then there is a constant \bar{K} such that if \mathcal{D}_1 and \mathcal{D}_2 are cylinders of lengths n_1 and n_2 respectively and if $n \geq n_1$ is such that $\mathcal{D}_1 \cap \sigma^{-n}\mathcal{D}_2 \neq \emptyset$ then

$$\frac{1}{\bar{K}} \leq \frac{\mu(\mathcal{D}_1 \cap \sigma^{-n}\mathcal{D}_2)}{\mu(\mathcal{D}_1)\mu(\mathcal{D}_2)} \leq \bar{K}. \quad (9.2)$$

We note the following consequence of (9.2). Let $\mathcal{F}_{a,b}$ denote the σ -algebra generated by $\{\omega_j\}_{a \leq j \leq b}$. Then there is a constant \hat{K} such that for each set $B \subset \mathcal{F}_{k,\infty}$

$$\frac{1}{\hat{K}} \leq \frac{\mu(B|\mathcal{F}_{-\infty,k})}{\mu(B|\mathcal{F}_{k,k})} \leq \hat{K}. \quad (9.3)$$

Indeed since μ is shift invariant it suffices to analyze the the case $k = 0$. Consider two cylinders $\bar{\mathcal{D}} = \mathcal{D}(\bar{\omega}_0, \dots, \bar{\omega}_n)$ and $\tilde{\mathcal{D}} = \mathcal{D}(\tilde{\omega}_0, \dots, \tilde{\omega}_m)$ with $\bar{\omega}_n = \tilde{\omega}_0$. Then

$$\begin{aligned} \mu(\omega \in \tilde{\mathcal{D}} | \sigma^{1-n}\omega \in \bar{\mathcal{D}}) &= \frac{\mu(\mathcal{D}(\bar{\omega}_0, \dots, \bar{\omega}_n, \tilde{\omega}_1, \dots, \tilde{\omega}_m))}{\mu(\mathcal{D}(\bar{\omega}_0, \dots, \bar{\omega}_n))} \\ &\geq \frac{\mu(\mathcal{D}(\tilde{\omega}_1, \dots, \tilde{\omega}_m))}{\bar{K}} \geq \frac{\mu(\tilde{\mathcal{D}})}{\bar{K}^2 \mu(\mathcal{D}(\tilde{\omega}_0))} = \frac{1}{\bar{K}^2} \mu(\omega \in \tilde{\mathcal{D}} | \mathcal{F}_{0,0})(\tilde{\omega}_0) \end{aligned}$$

proving the lower bound in (9.3). The upper bound is similar.

Let $f : \mathbb{T}^m \rightarrow \mathbb{T}^m$ be an Anosov diffeomorphism preserving a smooth measure $\bar{\mu}$. Using Markov partitions one constructs a measure preserving (Hölder) isomorphism \mathbf{j} between $(\Sigma_A, \sigma, \mu_\phi)$ and $(\mathbb{T}^m, f, \bar{\mu})$ where Σ_A is a topologically transitive subshift of finite type (SFT), μ_ϕ is the Gibbs measure

with Holder potential

$$\phi(\omega) = \ln |\det(df|E^u)(\mathbf{j}(\omega))| \quad (9.4)$$

and E^u is the unstable distribution of f ([14]). (Note that the fact that $\bar{\mu}$ is the equilibrium measure for the potential $\bar{\phi}(x) = \ln |\det(df|E^u)(x)|$ follows from Pesin entropy formula). Therefore, Theorem 2.7 follows from:

Theorem 9.1 *Let $d \geq 1$, Σ_A be a topologically mixing subshift of finite type and μ be a Gibbs measure with a Hölder potential. Let (G_t) be a geodesic flow on $SL(2, \mathbb{R})/\Gamma$ (if $d = 1$), or a Weyl chamber flow on $SL(d+1, \mathbb{R})/\Gamma$ (when $d \geq 2$). Let $\tau : \Sigma_A \rightarrow \mathbb{R}^d$ be a mean zero Hölder cocycle which is not cohomologous to a cocycle taking value in a proper linear subspace of \mathbb{R}^d . Then the homeomorphism on $\Sigma_A \times SL(d+1, \mathbb{R})/\Gamma$ defined by*

$$F(x, y) = (\sigma \omega, G_{\tau(\omega)} y)$$

with the invariant measure $\zeta = \mu \times \text{Haar}$ is non-Bernoulli.

10 Weyl chamber flow

Let $d \geq 1$. Let $H := SL(d+1, \mathbb{R})$, Γ be a co-compact lattice in H and $Y := H/\Gamma$. Let $D_+ \subset H$ be the subgroup of diagonal matrices in H with positive elements. It is easy to see that D_+ is isomorphic to \mathbb{R}^d . The group D_+ acts on Y by left translation. When $d = 1$, this one parameter flow is called *geodesic flow*. When $d \geq 2$, it is a \mathbb{R}^d action, which is called *Weyl Chamber flow*. Let $\mathfrak{h} = \mathfrak{sl}(d+1, \mathbb{R})$ be the Lie algebra of H and let d_H denote the right-invariant metric on H and d_Y the induced metric on Y . For $1 \leq i, j \leq d+1$, let $v_{i,j}$ be the elementary $(d+1) \times (d+1)$ matrix with only one nonzero entry equal to one in the row i and the column j . If $i \neq j$ let $\mathfrak{h}_{i,j} \subset \mathfrak{h}$ be the subalgebra generated by $v_{i,j}$. Let $\mathfrak{o} \subset \mathfrak{h}$ be the subalgebra of diagonal matrices with zero trace. Then

$$\mathfrak{h} = \mathfrak{o} \oplus \left(\bigoplus_{i \neq j} \mathfrak{h}_{i,j} \right). \quad (10.1)$$

For each pair (i, j) define $\chi_{i,j} : \mathbb{R}^d \rightarrow \mathbb{R}$ by $\chi_{i,j}(\mathbf{t}) = \chi_{i,j}(t_1, \dots, t_d) = t_i - t_j$. Then for $v \in \mathfrak{h}_{i,j}$

$$G_{\mathbf{t}} \cdot \exp(v) = \exp(e^{\chi_{i,j}(\mathbf{t})} v) \cdot G_{\mathbf{t}}.$$

The $\chi_{i,j}$ are exactly the Lyapunov functionals of G in classical Lyapunov theory. For every $i \neq j$, the equation $t_i = t_j$ defines a hyperplane $H_{i,j}$ in

\mathbb{R}^d , where the functional $\chi_{i,j}$ vanishes (notice that for $i = j$, $\chi_{i,i} \equiv 0$). The connected components of

$$\mathbb{R}^d \setminus \bigcup_{i \neq j} H_{i,j}$$

are called Weyl chambers of the action G . By continuity each Lyapunov functional has constant sign in a Weyl chamber. For any Weyl chamber \mathcal{C} we denote

$$\mathfrak{h}_{\mathcal{C}}^+ := \bigoplus_{\substack{\chi_{i,j} > 0 \text{ on } \mathcal{C}}} \mathfrak{h}_{i,j}$$

with an analogous notation for $\mathfrak{h}_{\mathcal{C}}^-$. The above distributions define foliations on G : for $y \in H$ let $\mathbb{W}_{\mathcal{C}}^+(y) = \exp(\mathfrak{h}_{\mathcal{C}}^+)y$ and $\mathbb{W}_{\mathcal{C}}^-(y) = \exp(\mathfrak{h}_{\mathcal{C}}^-)y$ respectively.

To simplify notations, we enumerate the Lyapunov functionals as $\{\chi_i\}_{1 \leq i \leq m}$ and the corresponding splitting (10.1) as

$$\mathfrak{h} = \bigoplus_{i \leq m} \mathfrak{h}_i.$$

Using the above splitting and the exponential map we can introduce the system of local coordinates on Y : there exists a constant ζ_0 such that if $d_H(y, y') \leq \zeta_0$, then

$$y = \exp(Z)y', \text{ where } Z = \sum_i Z_i, \text{ and } Z_i \in \mathfrak{h}_i. \quad (10.2)$$

By [66], any Weyl chamber flow is exponentially mixing.¹⁵

Moreover, we say that G is *exponentially mixing on balls* if there exist $C, \eta', \eta > 0$ such that for every $\mathbf{v} \in \mathbb{R}^d$, every $B(y, r), B(y', r') \subset Y$ with $y, y' \in Y$ and $r, r' \in (e^{-\eta' \|\mathbf{v}\|}, 1)$ the following holds:

$$|\nu(B(y, r) \cap G_{\mathbf{v}} B(y', r')) - \nu(B(y, r))\nu(B(y', r'))| \leq C e^{-\eta \|\mathbf{v}\|}. \quad (10.3)$$

A standard approximation argument (see eg. [50]) shows that exponential mixing for sufficiently smooth functions implies that G is exponentially mixing on balls. So we have:

Lemma 10.1 *Any Weyl chamber flow is exponentially mixing on balls.*

¹⁵ In fact, by [9] G is exponentially mixing of all orders. The multiple exponential mixing plays important role in verifying that F satisfies the CLT if $d \geq 3$ (see §B.1), but it is not needed in the proof of Theorem 9.1.

11 Relative atoms of the past partition

Recall that $F : (\Sigma_A \times Y, \zeta) \rightarrow (\Sigma_A \times Y, \zeta)$ is given by $F(\omega, y) = (\sigma\omega, G_{\tau(\omega)}y)$. Let \mathcal{P}_ϵ be a partition of Σ_A given by cylinders on coordinates $[-\epsilon^{-\frac{1}{\beta}}, 0]$, where β is the Hölder exponent of τ . Let \mathcal{Q}_ϵ be a partition of Y into sets with piecewise smooth boundaries and of diameter $\leq \epsilon$.

Recall that Ω denotes the alphabet of the shift space. For $\omega^- = (\dots, \omega_{-1}, \omega_0) \in \Omega^{\mathbb{Z}_{\leq 0}}$, let

$$\Sigma_A^+(\omega^-) = \{\omega^+ = (\omega_1, \omega_2, \dots) \in \Omega^{\mathbb{Z}^+} : (\dots, \omega_{-1}, \omega_0, \omega_1, \dots) \in \Sigma_A\}.$$

Note that $\Sigma_A^+(\omega^-)$ only depends on ω_0 . We will also use the notation $\omega = (\omega^-, \omega^+)$ and $\Sigma_A^+(\omega) = \Sigma_A^+(\omega^-)$. For $\omega = (\omega^-, \omega^+)$ and $S^+ \subset \Sigma_A^+(\omega)$, we write

$$\mu_\omega^+(S^+) = \mu(\{(\omega^-, \bar{\omega}^+) : \bar{\omega}^+ \in S^+\}).$$

With a slight abuse of notation, we also denote by μ_ω^+ a measure on Σ_A defined by $\mu_\omega^+(S) = \mu_\omega^+(\{\bar{\omega}^+ : (\omega^-, \bar{\omega}^+) \in S\})$. Notice that, for any measurable subset $S \subset \Sigma_A$,

$$\mu(S) = \int_{\Sigma_A} \mu_\omega^+(S) d\mu(\omega).$$

We can assume that τ only depends on the past. Indeed, if this is not the case, then ([92, Proposition 1.2]) τ is cohomologous to another Hölder function $\bar{\tau}$ depending only on the past: $\tau(\omega) = \bar{\tau}(\omega^-) + h(\omega) - h(\sigma\omega)$. If \bar{F} is the (T, T^{-1}) transformation constructed using $\bar{\tau}$ and $L(\omega, y) = (\omega, G_{h(\omega)}y)$, then $L \circ F = \bar{F} \circ L$. Since F and \bar{F} are conjugate, we can indeed assume that τ only depends on the past.

The main result of this section is:

Proposition 11.1 *There exists $\epsilon_0 > 0$ and a full measure set $V \subset \Sigma_A \times Y$ such that for every $(\omega, y) \in V$, the atoms of*

$$\bigvee_{i=0}^{\infty} F^i(\mathcal{P}_{\epsilon_0} \times \mathcal{Q}_{\epsilon_0})$$

are of the form $\{\omega^- \times \Sigma_A^+(\omega^-)\} \times \{y\}$, i.e. the past of ω and the Y -coordinate are fixed.

Before we prove the above proposition, we need some lemmas. For a non-zero χ_i , let $\mathcal{C}_i \subset \mathbb{R}^d$ be a cone

$$\mathcal{C}_i = \{\mathbf{a} \in \mathbb{R}^d : \chi_i(\mathbf{a}) \geq c' \|\mathbf{a}\|\}, \quad \text{where } c' = \min_{i: \chi_i \neq 0} \|\chi_i\|/2.$$

We start with the following lemma:

Lemma 11.2 *Let (G, Y, v) be a Weyl chamber flow. Choose cones $\hat{\mathcal{C}}_i$ properly contained in \mathcal{C}_i . Then for each $a \in \mathbb{R}_+$ there exists $\kappa = \kappa(G, a) > 0$ such that the following holds. Let $\{\mathbf{a}_j\}_{j \in \mathbb{N}} \subset \mathbb{R}^d$, be a sequence such that $\mathbf{a}_1 = 0$ and*

- A. $\sup_j \|\mathbf{a}_{j+1} - \mathbf{a}_j\| < a$;
- B. for every i we have $\sup_{j: \mathbf{a}_j \in \hat{\mathcal{C}}_i} \|\mathbf{a}_j\| = \infty$.

Then for any $y, y' \in Y$ with $y' \notin \mathcal{O}(y)$, where $\mathcal{O}(y)$ denotes the G -orbit of y , there is $j \in \mathbb{N}$ such that $d_Y(G_{\mathbf{a}_j}y, G_{\mathbf{a}_j}y') \geq \frac{\kappa}{4}$.

In order to prove the lemma, we first establish the divergence on the universal cover.

Lemma 11.3 *There exists $\bar{\kappa} > 0$ such that for any $y, y' \in H$ with $y' \notin \mathcal{O}(y)$ and any $\{\mathbf{a}_j\}$ satisfying A., B., there exists j_0 such that*

$$d_H(G_{\mathbf{a}_{j_0}}y, G_{\mathbf{a}_{j_0}}y') > \bar{\kappa}.$$

Proof To simplify notation we denote $G_{\mathbf{t}}y$ simply by $\mathbf{t}y$.

Fix $y, y' \in H$. WLOG, assume $d_H(y, y') < \zeta_0$ where ζ_0 is defined above (10.2). We can write $y = \exp(Z)y'$, where $Z \in \mathfrak{h}$, and $Z = \bigoplus_i Z_i$ with $Z_i \in \mathfrak{h}_i$. Since $y' \notin \mathcal{O}(y)$, there exists i such that $\chi_i \neq 0$ and $Z_i \neq 0$. Accordingly there is a Weyl chamber \mathcal{C} such that splitting $Z = Z^+ + Z^-$ with $Z^\pm \in \mathfrak{h}_{\mathcal{C}}^\pm$ we have $Z^+ \neq 0$. Let $y'' = \mathbb{W}_{\mathcal{C}}^-(y) \cap \mathbb{W}_{\mathcal{C}}^+(y')$. Then $y'' \neq y'$ since $Z \notin \mathfrak{h}_{\mathcal{C}}^-$.

Let $\hat{\mathcal{C}}$ be a cone which is strictly contained inside \mathcal{C} . Note that by the definition of y'' , there is a global constant $K > 0$ such that for each $\mathbf{a}_j \in \mathcal{C}$ we have $d_H(\mathbf{a}_j y, \mathbf{a}_j y'') \leq K\zeta_0$. By triangle inequality, $d_H(\mathbf{a}_j y, \mathbf{a}_j y') \geq d_H(\mathbf{a}_j y', \mathbf{a}_j y'') - d_H(\mathbf{a}_j y, \mathbf{a}_j y'')$. To complete the proof it is enough to note that since the vectors in $\mathfrak{h}_{\mathcal{C}}^+$ are expanded by $\hat{\mathcal{C}}$ at a uniform rate and $\sup_{j: \mathbf{a}_j \in \hat{\mathcal{C}}} \|\mathbf{a}_j\| = \infty$, there exists j such that $d_H(\mathbf{a}_j y', \mathbf{a}_j y'') \geq K\zeta_0 + \bar{\kappa}$, for some $\bar{\kappa} > 0$. \square

Proof of Lemma 11.2 To simplify notation we denote $G_{\mathbf{t}}y$ simply by $\mathbf{t}y$. Since $\Gamma \subset H$ is co-compact, it follows that there exists $c > 0$ such that

$$\inf_{y \in H} \inf_{\gamma \neq e} d_H(y, y\gamma) > c > 0. \quad (11.1)$$

Let $C_1 := \sup_j \|\mathbf{a}_{j+1} - \mathbf{a}_j\| < \infty$ and let $C = C(\alpha) > 0$ be such that

$$\sup_{0 < d_H(y, y') \leq 1} \sup_{\|\mathbf{b}\| < C_1} \frac{d_H(\mathbf{b}y, \mathbf{b}y')}{d_Y(y, y')} \leq C, \quad (11.2)$$

Let $0 < \kappa < \bar{\kappa}$ be such that $c \geq (C + 1/4)\kappa$ (recall that $\bar{\kappa}$ is the constant from Lemma 11.3). Let $y, y' \in Y$, with $y' \notin \mathcal{O}(y)$, with $d_Y(y, y') \leq \kappa/4$. By taking appropriate lifts of y and y' to H , we can assume that $d_H(y, y') \leq \kappa/4$. By Lemma 11.3, there exists $j_0 \in \mathbb{N}$ such that $d_H(\mathbf{a}_{j_0}y, \mathbf{a}_{j_0}y') > \kappa/4$. Let us take the smallest j_0 with this property. Then, $d_H(\mathbf{a}_{j_0-1}y, \mathbf{a}_{j_0-1}y') \leq \kappa/4$. Therefore by the bound in (11.2)

$$d_H(\mathbf{a}_{j_0}y, \mathbf{a}_{j_0}y') = d_H\left((\mathbf{a}_{j_0} - \mathbf{a}_{j_0-1})(\mathbf{a}_{j_0-1}y), (\mathbf{a}_{j_0} - \mathbf{a}_{j_0-1})(\mathbf{a}_{j_0-1}y')\right) \leq C\kappa.$$

Take $\gamma \in H$ such that $d_Y(\mathbf{a}_{j_0}y, \mathbf{a}_{j_0}y') = d_H(\mathbf{a}_{j_0}y, \mathbf{a}_{j_0}y'\gamma)$. By (11.1) we get

$$\begin{aligned} d_H(\mathbf{a}_{j_0}y, \mathbf{a}_{j_0}y'\gamma) &\geq d_H(\mathbf{a}_{j_0}y', \mathbf{a}_{j_0}y'\gamma) - d_H(\mathbf{a}_{j_0}y, \mathbf{a}_{j_0}y') \\ &\geq c - C\kappa \geq \kappa/4. \end{aligned}$$

This finishes the proof. \square

Recall that for $\tau : \Sigma_A \rightarrow \mathbb{R}^d$ and $n \in \mathbb{N}$, we denote $\tau_n(\omega) := \sum_{j=0}^{n-1} \tau(\sigma^j \omega)$, and $\tau_{-n}(\omega) = -\tau_n(\sigma^{-n} \omega)$. The next result, proven in §B.2, helps verifying condition (B) of Lemma 11.2.

Lemma 11.4 *Let $\tau : \Sigma_A \rightarrow \mathbb{R}^d$ be a zero mean Hölder function that is not cohomologous to a function taking values in a linear subspace of \mathbb{R}^d of dimension $< d$. Then for any cone $\mathcal{C} \subset \mathbb{R}^d$, for μ a.e. $\omega \in \Sigma_A$*

$$\sup_{v \in (\{\tau_n(\omega)\}_{n \geq 0}) \cap \mathcal{C}} \|v\| = \infty \quad \text{and} \quad \sup_{v \in (\{\tau_n(\omega)\}_{n < 0}) \cap \mathcal{C}} \|v\| = \infty. \quad (11.3)$$

Proof of Proposition 11.1 Take $\epsilon_0 := \kappa(G, \|\tau\|_{C^0})/5$ where κ is from Lemma 11.2. By Corollary 2 in [57], the skew product F is ergodic. Let Λ be the set of points (ω, y) whose forward and backward orbits are dense and such that (11.3) holds for ω and every cone $\{\hat{\mathcal{C}}_i\}$ from Lemma 11.2. By ergodicity of F and Lemma 11.4 $\zeta(\Lambda) = 1$.

Notice that if $(\omega, y) \in \Lambda$, and $(\omega, y), (\bar{\omega}, y')$ lie in the same atom of $\bigvee_{i=0}^{\infty} F^i(\mathcal{P} \times \mathcal{Q}_{\epsilon_0})$, then $\omega^- = \bar{\omega}^-$. Since τ depends only on the past, $\tau_{-j}(\omega) = \tau_{-j}(\bar{\omega})$ for $j \in \mathbb{N}$. We will show that $y' = y$.

Assume first that $y' \in \mathcal{O}(y)$ and let $y' = G_w \cdot y$, for some $w \in \mathbb{R}^d$. Let Q be an atom of \mathcal{Q}_{ϵ_0} . Note that there exists $q \in Q$ and $\epsilon = \epsilon(w) > 0$ such that $B(q, \epsilon) \subset Q$ and $G_w \cdot B(\epsilon, q) \cap Q = \emptyset$. Indeed, if not then Q would be invariant under the translation G_w which is impossible if $G_w \neq Id$ by Moore ergodicity theorem [87].

This contradiction shows that such q and ϵ exist. Since the F orbit of (ω, y) is dense, there exists n , such that $F^{-n}(\omega, y) \in \Sigma_A \times B(\epsilon, q) \subset \Sigma_A \times Q$. Let $u = \tau_{-n}(\omega)$. Then $G_u y' = G_w G_u y \notin Q$. So $F^{-n}(\omega, y)$ and $F^{-n}(\omega', y')$ are not in the same atom of $\mathcal{P} \times \mathcal{Q}$. If $y' \notin \mathcal{O}(y)$ then we use Lemmas 11.4 and 11.2 with $\mathbf{a}_j = \tau_j(\omega)$ to finish the proof. \square

Remark 11.5 We believe that ALL partially hyperbolic algebraic abelian actions satisfy the assertion of Proposition 11.1. However, the proof is more complicated if there is a polynomial growth in the center. We plan to deal with the general situation in a forthcoming paper.

12 Non Bernoullicity under zero drift. Proof of Theorem 9.1

12.1 The main reduction

We introduce the notion of (ϵ, n) -closeness which is an averaged version of Bowen closeness. Let d denote the product metric on $\Sigma \times Y$. Two points $(\omega, y), (\omega', y') \in \Sigma_A \times Y$ are called (ϵ, n) -close if

$$\# \left\{ i \in [1, n] : d(F^i(\omega, y), F^i(\omega', y')) < \epsilon \right\} \geq (1 - \epsilon)n.$$

We will now state two propositions that imply Theorem 9.1.

Proposition 12.1 *If F is Bernoulli then for every $\epsilon, \delta > 0$ there exists n_0 such that for every $n \geq n_0$ there exists a measurable set $W \subset \Sigma_A \times Y$ with $\xi(W) > 1 - \delta$ such that if $(\omega, y), (\bar{\omega}, \bar{y}) \in W$, then there exists a map $\Phi_{(\omega^-, y)(\bar{\omega}^-, \bar{y})} : \Sigma_A^+(\omega) \rightarrow \Sigma_A^+(\bar{\omega})$ with $(\Phi_{\omega^-, \bar{\omega}^-})_*(\mu_\omega^+) = \mu_{\bar{\omega}}^+$ and a set $U_{\omega^-} \subset \Sigma_A^+(\omega)$ such that:*

- (1) $\mu_\omega^+(U_{\omega^-}) > 1 - \delta$;
- (2) if $z \in U_{\omega^-}$ then $((\omega^-, z), y)$ and $((\bar{\omega}^-, \Phi_{(\omega^-, y)(\bar{\omega}^-, \bar{y})}z), \bar{y})$ are (ϵ, n) -close.

We will also need another result. For $\epsilon > 0, n \in \mathbb{N}, \omega \in \Sigma_A, y' \in Y$, let

$$\begin{aligned} D(\omega, y', \epsilon, n) \\ := \left\{ y \in Y : \exists \omega' \in \Sigma_A \text{ s.t. } (\omega, y) \text{ and } (\omega', y') \text{ are } (\epsilon, n)\text{-close} \right\}. \end{aligned}$$

Proposition 12.2 *There exists $\epsilon' > 0$, an increasing sequence $\{n_k\}$, and a family of sets $\{\Omega_k\}$, $\Omega_k \subset \Sigma_A$, $\mu(\Omega_k) \rightarrow 1$, such that*

$$\lim_{k \rightarrow \infty} \sup_{\substack{\omega \in \Omega_k \\ y' \in Y}} \nu(D(\omega, y', \epsilon', n_k)) = 0.$$

We will prove Proposition 12.1 in a Sect. 12.2 and Proposition 12.2 in Sect. 12.3. Now we show how these two propositions imply Theorem 9.1:

Proof of Theorem 9.1 We argue by contradiction. Fix $\epsilon = \epsilon'/100$, $\delta = \epsilon$, and let $n = n_k$ (for some sufficiently large k , specified below). Let $W \subset \Sigma_A \times Y$ be the set from Proposition 12.1. Let

$$W^y := \{\omega \in \Sigma_A : (\omega, y) \in W\} \quad \text{and} \quad W_\omega := \{y \in M : (\omega, y) \in W\}.$$

By Fubini's theorem, there exists $Z \subset \Sigma_A$, $\mu(Z) \geq 1 - 2\epsilon$ such that for every $\omega \in Z$, $\nu(W_\omega) > 1/2$. Let k be large enough (in terms of ϵ) such that $\mu(Z \cap \Omega_k) \geq 1 - 4\epsilon$. By Fubini's theorem, it follows that there exists $Z' \subset Z \cap \Omega_k$, $\mu(Z') > 1 - 4\epsilon$ such that for $\omega \in Z'$, $\mu_\omega^+(Z \cap \Omega_k) > 1 - 8\epsilon$. In particular, it follows that

$$\mu_\omega^+(\{\bar{\omega}^+ \in U_{\omega^-} : (\omega^-, \bar{\omega}^+) \in Z \cap \Omega_k\}) > 1 - 16\epsilon.$$

Let $\omega = (\omega^-, \omega^+) \in Z \cap \Omega_k \cap (\{\omega^-\} \times U_{\omega^-})$ and let $(\bar{\omega}, y') \in W$. Since $\omega \in Z$ it follows that $\nu(W_\omega) > 1/2$. Since $\omega \in \Omega_k$, it follows that for k large enough there exists

$$y \in W_\omega \setminus D(\omega, y', \epsilon', n_k). \quad (12.1)$$

Since $\omega^+ \in U_{\omega^-}$, by (2) we get that (ω^-, ω^+, y) and $(\bar{\omega}^-, \Phi_{\omega^-, \bar{\omega}^-}(\omega^+), y')$ are (ϵ, n_k) -close. This by the definition of $D(\omega, y', \epsilon', n_k)$ implies that $y \in D(\omega, y', \epsilon', n_k)$. This however contradicts (12.1). This contradiction finishes the proof. \square

12.2 Hamming–Bowen closeness

We start with introducing the notion of VWB (very weak Bernoulli) partitions in the setting of skew-product for which the assertion of Proposition 11.1 holds (see e.g. [22] or [62]). Let \mathcal{R} be a partition of $\Sigma_A \times Y$. Two points $(\omega, y), (\omega', y') \in \Sigma_A \times Y$ are called $(\epsilon, n, \mathcal{R})$ -matchable if

$$\#\{i \in [1, n] : F^i(\omega, y) \text{ and } F^i(\omega', y') \text{ are in the same } \mathcal{R} \text{ atom}\} \geq (1 - \epsilon)n.$$

Definition 12.3 F is very weak Bernoulli with respect to \mathcal{R} if and only if for every $\epsilon' > 0$, there exists n' such that for every $n \geq n'$ there exists

a measurable set $W' \subset \Sigma_A \times M$ with $\mu \times \nu(W') > 1 - \epsilon'$ such that if $(\omega, y), (\bar{\omega}, \bar{y}) \in W'$, then there exists a map $\Phi_{(\omega^-, y)(\bar{\omega}^-, \bar{y})} : \Sigma_A^+(\omega) \rightarrow \Sigma_A^+(\bar{\omega})$ with $(\Phi_{\omega^-, \bar{\omega}^-})_*(\mu_\omega^+) = \mu_{\bar{\omega}}^+$ and a set $U'_{\omega^-} \subset \Sigma_A^+(\omega)$ such that:

- (1) $\mu_\omega^+(U'_{\omega^-}) > 1 - \epsilon'$;
- (2) if $z \in U'_{\omega^-}$ then $((\omega^-, z), y)$ and $((\bar{\omega}^-, \Phi_{(\omega^-, y)(\bar{\omega}^-, \bar{y})}z), \bar{y})$ are $(\epsilon', n, \mathcal{R})$ -matchable.

Proof of Proposition 12.1 Recall that by [88] if F is Bernoulli then it is VWB with respect to every non-trivial partition.

Let $(\mathcal{P} \times Q)_n$ be the sequence of partitions defined above, where the atoms have diameter that goes to 0 as $n \rightarrow \infty$. Let \bar{n} be such that the atoms of $(\mathcal{P} \times Q)_{\bar{n}}$ have diameter $\leq \epsilon$. This then implies that if two points (ω, y) and (ω', y') are (ϵ, n) matchable, then they are (ϵ, n) -close. It is then enough to use VWB definition for $(\mathcal{P} \times Q)_{\bar{n}}$ with $\epsilon' = \min\{\delta, \epsilon\}$. This finishes the proof. \square

Remark 12.4 Now we explain why it is easier to work with closeness rather than matchability, in the case $G = \mathbb{R}^d$. Notice that if (ω, y) and (ω', y') are (ϵ, n) -close, and $\|u\| < \delta < \epsilon$, then (ω, y) and $(\omega', G_u y')$ are $(\epsilon + \delta, n)$ close.¹⁶ This is not necessarily true for matchability (if the orbit of y' is always close to the boundary of the partition). This property of closeness crucially simplifies our consideration as it allows us to obtain a crucial inclusion (15.4).

12.3 Proof of Proposition 12.2

Given Ω_k, n_k denote

$$a_k(\epsilon') := \sup_{\substack{\omega \in \Omega_k \\ y' \in Y}} \nu(D(\omega, y', \epsilon', n_k)).$$

Proposition 12.5 *There exists $n_1 \in \mathbb{N}$ and a family of sets $\{\Omega_k\}$ such that if $\epsilon_k := \left(1 - \frac{1}{50k^2}\right) \epsilon_{k-1}$, $\epsilon_1 := \frac{1}{10n_1}$ and $n_{k+1} = (10k)^{100} \cdot n_k$, then we have*

$$a_k(\epsilon_k) \rightarrow 0, \text{ as } k \rightarrow \infty.$$

¹⁶ Notice that for any $i \in \mathbb{N}$ the points $F^i(\omega', y')$ and $F^i(\omega', G_u y')$ are δ close. Indeed, they have the same first coordinate and the second one is $G_{\tau_i(\omega)} y'$ vs $G_{u+\tau_i(\omega)} y'$ which are δ close since $\|u\| < \delta$.

We remark that the recursive relations in Proposition 12.5 imply that

$$\epsilon_k = \epsilon_1 \prod_{j=2}^k \left(1 - \frac{1}{50j^2}\right), \quad (12.2)$$

$$n_{k+1} = n_1 (10^k k!)^{100}. \quad (12.3)$$

Proposition 12.5 which is proven in Sect. 15 immediately implies Proposition 12.2:

Proof of Proposition 12.2 We define $\epsilon' := \inf_{k \geq 1} \epsilon_k = \prod_{j=2}^{\infty} \epsilon_1 \left(1 - \frac{1}{50j^2}\right)$. Then by the definition of $\{\epsilon_k\}$, $\epsilon' > 0$ and monotonicity, we have

$$0 \leq a_k(\epsilon') \leq a_k(\epsilon_k) \rightarrow 0,$$

as $k \rightarrow \infty$. This finishes the proof. \square

13 Consequence of exponential mixing

We have the following quantitative estimates on independence of the sets $D(\omega, y', \epsilon', n_k)$ under the action G_t . This is the only place in the proof where we use exponential mixing of G_t .

In this section we shall denote $\ell_k = 2k^{20} \sqrt{n_{k-1}}$.

Lemma 13.1 *For $k \in \mathbb{N}$ let $\omega_1, \omega_2 \in \Sigma_A$ be such that*

$$\sup_{r \leq n_{k-1}} \|\tau_r(\omega_i)\| \leq \ell_k \quad (13.1)$$

for $i = 1, 2$. Then, for any $y_1, y_2 \in Y$, any $v \in \mathbb{Z}^d$, $\|v\| \geq k^{25} n_{k-1}^{1/2}$, and any $\epsilon > 0$.

$$\begin{aligned} & v \left(G_v(D(\omega_1, y_1, \epsilon, n_{k-1})) \cap D(\omega_2, y_2, \epsilon, n_{k-1}) \right) \\ & \leq C_{\#} \cdot \prod_{i=1,2} v \left(D(\omega_i, y_i, \epsilon + 2^{-n_{k-1}^{1/2}}, n_{k-1}) \right). \end{aligned}$$

Proof Let $L := \max\{\sup_{\|v\|=1} \|G_v\|_{C^1}, 100\}$. Then if $d(y, y') \leq (2L)^{-\ell_k}$, then

$$d(G_{uy}, G_{uy'}) \leq L^{\ell_k} \cdot (2L)^{-\ell_k} \leq 2^{-\ell_k} \leq 2^{-n_{k-1}^{1/2}}$$

for all $u \in \mathbb{A}$ with $\|u\| \leq \ell_k$. Using this for $u = \tau_r(\omega_i)$, $r < n_{k-1}$, (13.1) implies that if $d(y, y') \leq (2L)^{-\ell_k}$, then

$$d(G_{\tau_j(\omega_i)}(y), G_{\tau_j(\omega_i)}(y')) \leq 2^{-n_{k-1}^{1/2}}, \quad \text{for all } j < n_{k-1}. \quad (13.2)$$

Therefore for every $y \in D(\omega_i, y_i, \epsilon, n_{k-1})$,

$$B\left(y, (2L)^{-\ell_k}\right) \subset D(\omega_i, y_i, \epsilon + 2^{-n_{k-1}^{1/2}}, n_{k-1}). \quad (13.3)$$

Using Besicovitch theorem for the cover $\{B\left(y, (2L)^{-\ell_k}\right)\}$, where

$$y \in D(\omega_i, y_i, \epsilon, n_{k-1}),$$

we get a finite cover by a family of balls $\{B_s^{j,i}\}_{j \leq C', s \leq m_j}$ $i = 1, 2$, such that for every $i \in \{1, 2\}$, $j \leq C'$, the balls $\{B_s^{j,i}\}_{s \leq m_j}$ are pairwise disjoint. Therefore

$$\begin{aligned} & \nu\left(G_v(D(\omega_1, y_1, \epsilon, n_{k-1})) \cap D(\omega_2, y_2, \epsilon, n_{k-1})\right) \\ & \leq \sum_{j, j'} \sum_{s, s'} \nu(G_v(B_s^{j,1}) \cap B_{s'}^{j',2}). \end{aligned}$$

Using that G is exponentially mixing on balls in the sense of (10.3), and the fact that $e^{-\eta' \|v\|} \leq (\frac{1}{2L})^{\ell_k}$ (since $\|v\| \geq k^{25} n_{k-1}^{1/2}$) we get that the above term is upper bounded by

$$C \cdot \sum_{j, j'} \sum_{s, s'} \nu(B_s^{j,1}) \nu(B_{s'}^{j',2}) = C \left[\sum_j \sum_s \nu(B_s^{j,1}) \right] \cdot \left[\sum_{j'} \sum_{s'} \nu(B_{s'}^{j',2}) \right]. \quad (13.4)$$

Since the balls are disjoint for fixed i and j , we have

$$\sum_s \nu(B_s^{j,i}) = \nu\left(\bigcup_s B_s^{j,i}\right) \leq \nu(D(\omega_i, y_i, \epsilon + 2^{-n_{k-1}^{1/2}}, n_{k-1})),$$

where the last inequality follows from (13.3). Since the cardinality of j 's is globally bounded (only depending on the manifold Y), (13.4) is upper bounded by

$$C \cdot C_d \cdot \prod_i \nu(D(\omega_i, y_i, \epsilon + 2^{-n_{k-1}^{1/2}}, n_{k-1})).$$

This finishes the proof. \square

We also have the following lemma.

Lemma 13.2 *For any constant $C_2 > 1$ the following is true. If $n_1 > C_2$ and b_k is a sequence of real numbers satisfying*

$$b_1 \leq \left(\frac{1}{100n_1} \right)^{300d} \text{ and } b_k \leq C_2 \cdot n_k^{2d+1} \cdot b_{k-1}^2,$$

then $b_k \rightarrow 0$.

Proof By induction, we see that

$$\ln b_k \leq (2^{k-1} - 1) \ln C_2 + (2d + 1) \left[\sum_{l=2}^k 2^{k-l} \ln n_l \right] + 2^{k-1} \ln b_1$$

Now using (12.3), we obtain

$$\begin{aligned} \ln b_k &\leq (2^{k-1} - 1) \ln C_2 + (2d + 1) \left[\sum_{l=2}^k 2^{k-l} 100l(\ln 10 + \ln l) \right] \\ &\quad + 2^{k+2}d \ln n_1 + 2^{k-1} \ln b_1. \end{aligned}$$

Using the condition on b_1 , the result follows. \square

14 Construction of Ω_k

Let n_1 be a number specified below and n_k be defined by (12.3). For $k \geq 2$ define

$$\begin{aligned} A_k &:= \left\{ \omega \in \Sigma_A : \#\{(i, j) \in [0, (10k)^{100}] \times [0, (10k)^{100}] \mid \right. \\ &\quad \left. i \neq j : \frac{1}{(|j - i|n_{k-1})^{1/2}} \|\tau_{(j-i)n_{k-1}}(\sigma^{in_{k-1}}\omega)\| \geq k^{-20} \right. \\ &\quad \left. > (10k)^{200}(1 - k^{-9}) \right\}, \\ B_k &:= \left\{ \omega \in \Sigma_A : \#\{i < (10k)^{100} : \sup_{r \leq n_{k-1}} \frac{1}{n_{k-1}^{1/2}} \|\tau_r(\sigma^{in_{k-1}}\omega)\| \leq k^{20} \right. \\ &\quad \left. > (10k)^{100}(1 - k^{-9}) \right\}. \end{aligned}$$

For $\omega \in \Sigma_A$, let $\omega_{[0, n-1]}$ denote the cylinder in coordinates $[0, \dots, n-1]$ determined by ω and let

$$\tilde{A}_k = \bigcup_{\omega \in A_k} \omega_{[0, n_k-1]} \quad \text{and} \quad \tilde{B}_k = \bigcup_{\omega \in B_k} \omega_{[0, n_k-1]}.$$

This way, \tilde{A}_k and \tilde{B}_k are unions of cylinders of length n_k .

The next lemma is proven in §B.3.

Lemma 14.1 *For any $C_0 > 0$, there exists an n_0 , such that if $n_1 \geq n_0$, we have:*

m1. *for every $k \geq 1$, $\min(\mu(\tilde{A}_k), \mu(\tilde{B}_k)) \geq 1 - C_0 k^{-8}$.*

m2. *for every $\omega \in \tilde{A}_k$,*

$$\begin{aligned} \# \left\{ (i, j) \in [0, (10k)^{100}] \times [0, (10k)^{100}], i \neq j : \right. \\ \left. \frac{1}{(|j - i| n_{k-1})^{1/2}} \|\tau_{(j-i)n_{k-1}}(\sigma^{in_{k-1}} \omega)\| \geq k^{-20}/2 \right\} > (10k)^{200} (1 - k^{-9}) \end{aligned} \quad (14.1)$$

and for every $\omega \in \tilde{B}_k$,

$$\# \left\{ i < (10k)^{100} : \sup_{r \leq n_{k-1}} \frac{1}{n_{k-1}^{1/2}} \|\tau_r(\sigma^{in_{k-1}} \omega)\| \leq 2k^{20} \right\} > (10k)^{100} (1 - k^{-9}). \quad (14.2)$$

Define

$$\bar{\Omega}_1 := \left\{ \omega : \|\tau_{n_1}(\omega)\| \geq n_1^{1/2-1/11} \right\} \quad \text{and} \quad \Omega_1 := \bigcup_{\omega \in \bar{\Omega}_1} \omega_{[0, n_1-1]}. \quad (14.3)$$

Notice that by Hölder continuity of τ it follows that for every $\omega \in \Omega_1$, we have $\|\tau_{n_1}(\omega)\| \geq n_1^{1/2-1/10}$, if n is large enough.

We suppose that n_1 is large enough, see below. For $k \geq 2$ we define:

$$\begin{aligned} \Omega_k := \tilde{A}_k \cap \tilde{B}_k \cap \left\{ \omega \in \Sigma_A : \# \{ i < (10k)^{100} : \sigma^{in_{k-1}}(\omega) \in \Omega_{k-1} \} \right. \\ \left. > (10k)^{100} (1 - k^{-5}) \right\}. \end{aligned}$$

Lemma 14.2 *For every k , the set Ω_k is a union of cylinders of length n_k .*

Proof For $k = 1$, this follows from the definition of Ω_1 . Also by definition the sets \tilde{A}_k and \tilde{B}_k are unions of cylinders of length n_k . Now inductively, if Ω_{k-1} is a union of cylinders of length n_{k-1} , then for every $i < (10k)^{100}$, the event

$\sigma^{in_{k-1}}(\omega) \in \Omega_{k-1}$, depends only on the $[in_{k-1}, (i+1)n_{k-1}]$ coordinates of ω . Since $i < (10k)^{100}$, the union of these events depends only on the first n_k coordinates of ω . \square

Let $\mathbf{C}_k = \{\mathcal{C} : \mathcal{C} \text{ is a union of cylinders of length } n_{k-1}\}$. Since μ is Gibbs, by (9.2) there exists a constant $C_1 \geq 1$ independent of the cylinders \mathcal{C} and of k such that for any cylinders $\mathcal{C}_1, \mathcal{C}_2 \in \mathbf{C}_k$, for any $m \geq n_{k-1}$

$$\mu(\mathcal{C}_1 \cap \sigma^m \mathcal{C}_2) \leq C_1 \mu(\mathcal{C}_1) \mu(\mathcal{C}_2).$$

We obtain by induction that for any $\mathcal{C}_1, \dots, \mathcal{C}_\ell \in \mathbf{C}_k$, any $j_1 < \dots < j_\ell$,

$$\mu\left(\bigcap_{i=1}^{\ell} \sigma^{j_i n_{k-1}} \mathcal{C}_i\right) \leq C_1^\ell \prod_{i=1}^{\ell} \mu(\mathcal{C}_i). \quad (14.4)$$

We assume that n_1 is so large that $\mu(\Omega_1) \geq 1 - C_1^{-2} 2^{-200}$.

Proposition 14.3 *There exists a constant $C_0 > 0$, such that for any $k \geq 1$,*

$$\mu(\Omega_k) \geq 1 - C_0 k^{-7}. \quad (14.5)$$

Proof of Proposition 14.3: Set $C_0 = \frac{1}{C_1^2 20^{200}}$. We prove (14.5) by induction.

By the choice of n_1 and C_0 , (14.5) holds for $k = 1$. Now assume it holds for $k-1 \geq 1$. We are going to show it holds for k .

We claim that $\mu(D_k) \leq C_0 k^{-7}/3$, where

$$D_k = \left\{ \omega \in \Sigma_A : \#\{i < (10k)^{100} : \sigma^{in_{k-1}}(\omega) \in \Omega_{k-1}\} < (10k)^{100} - (10k)^{95} \right\}.$$

By Lemma 14.2, the set Ω_{k-1} is a union of cylinders of length n_{k-1} . So is the complement Ω_{k-1}^c .

Divide the interval $[0, (10k)^{100}]$ into $10(10k)^{94}$ intervals of length $10^5 k^6$. If $\omega \in D_k$, one of those intervals I should contain at least k visits to Ω_{k-1}^c . Let i_1, \dots, i_k be the times of the first k visits inside I . By (14.4), for each tuple i_1, \dots, i_k

$$\mu\left(\sigma^{i_j n_{k-1}} \omega \in \Omega_{k-1}^c \text{ for } j = 1, \dots, k\right) \leq (C_1 \mu(\Omega_{k-1}^c))^k.$$

Since the number of tuples inside I is less than $|I|^k = 10^5 k^{6k}$,

$$\mu\left(\#\{i \in I : \sigma^i \omega \in \Omega_{k-1}^c\} \geq k\right) \leq (10k)^{6k} C_1^k \mu(\Omega_{k-1}^c)^k.$$

Since there are $10(10k)^{94}$ intervals, we have

$$\mu(D_k) \leq 10(10k)^{94}(10k)^{6k}C_1^k\mu(\Omega_{k-1}^c)^k \leq \frac{1}{C_1^k 2^{100k}k^k} \leq C_0 k^{-7}/3.$$

By **m1** in Lemma 14.1 and the definition of Ω_k , we obtain $\mu(\Omega_k) \geq 1 - C_0 k^{-7}$.

□

Definition 14.4 We say that a pair $(i, j) \in [0, (10k)^{100}]^2$ is n_k -good for ω if for $v \in \{i, j\}$ $\sigma^{vn_{k-1}}\omega \in \Omega_{k-1}$,

$$\frac{1}{(|j - i|n_{k-1})^{1/2}} \|\tau_{(j-i)n_{k-1}}(\sigma^{in_{k-1}}\omega)\| \geq k^{-20}/2, \quad (14.6)$$

and

$$\sup_{r \leq n_{k-1}} \frac{1}{n_{k-1}^{1/2}} \|\tau_r(\sigma^{vn_{k-1}}\omega)\| \leq 2k^{20}. \quad (14.7)$$

By definition of Ω_k , there are at least $(10k)^{200}(1 - 5k^{-5})$ n_k -good pairs (i, j) , for every $\omega \in \Omega_k$.

15 Proof of Proposition 12.5

We will show that Proposition 12.5 holds for sets Ω_k and n_1 from Sect. 14. Let $C_2 = 10^{200} \cdot C_\# \cdot d^d \cdot 100^d (\sup \|\tau\|)^d$, where $C_\#$ is from Lemma 13.1.

We start with the following lemma:

Lemma 15.1 *Let $n_1 > C_2$ be sufficiently large. Then*

$$a_1(\epsilon_1) \leq \left(\frac{1}{100n_1}\right)^{300d}.$$

Proof Let $\omega \in \Omega_1$ and $y \in D(\omega, y', \epsilon_1, n_1)$. Thus there is some ω' so that (ω, y) and (ω', y') are (ϵ_1, n_1) -close. Since $\epsilon_1 = \frac{1}{10n_1}$ it follows that for every $0 \leq i \leq n_1 - 1$,

$$d\left(F^i(\omega, y), F^i(\omega', y')\right) < \epsilon_1.$$

Since τ depends only on the past and is Hölder continuous with exponent β , this implies in particular that

$$\|\tau_i(\omega) - \tau_i(\omega')\| \leq C\epsilon_1^\beta \quad \text{for } i \leq n_1.$$

Let $\epsilon_0 = \epsilon_1^\beta$. Using closeness of $F^i(\omega, y)$ and $F^i(\omega', y')$ on the second coordinate, we get

$$d\left(G_{\tau_i(\omega)}y, G_{\tau_i(\omega)}y'\right) < 2C\epsilon_0 \quad \text{for } i \leq n_1. \quad (15.1)$$

We claim that (15.1) implies that

$$d_H\left(G_{\tau_i(\omega)}y, G_{\tau_i(\omega)}y'\right) < 2C\epsilon_0 \quad \text{for } i \leq n_1. \quad (15.2)$$

Indeed, if not let $i_0 \leq n_1$ be the smallest index i for which (15.2) doesn't hold. This means that

$$d_H\left(G_{\tau_{i_0-1}(\omega)}y, G_{\tau_{i_0-1}(\omega)}y'\right) < 2C\epsilon_0.$$

Note that by (15.1) there is some γ so that

$$d_H\left(G_{\tau_{i_0}(\omega)}y, G_{\tau_{i_0}(\omega)}y'\gamma\right) < 2C\epsilon_0,$$

and by the definition of i_0 , $\gamma \neq e$. The last two displayed inequalities imply that for some global constant $C'' > 0$,

$$d_H\left(G_{\tau_{i_0}(\omega)}y', G_{\tau_{i_0}(\omega)}y'\gamma\right) < C''\epsilon_0.$$

If ϵ_0 is small enough, this gives a contradiction with the systole bound (11.1). So (15.2) indeed holds.

Since $\omega \in \Omega_1$ (see (14.3)), it follows that

$$\|\tau_{n_1}(\omega)\| \geq n_1^{1/2-1/10}. \quad (15.3)$$

It follows that $G_{\tau_{n_1}(\omega)}$ expands the leaves of one of the Lyapunov foliations by at least $e^{cn_1^{2/5}}$. Hence each leaf intersects the set of y' satisfying (15.2) in a set of measure $O\left(e^{-cn_1^{2/5}}\right)$.

Therefore $v(D(\omega, y', \epsilon_1, n_1)) \leq C' \cdot e^{-cn_1^{2/5}}$, whence $a_1(\epsilon_1) \leq C \cdot e^{-cn_1^{2/5}} \leq \left(\frac{1}{100n_1}\right)^{300d}$ if n_1 is sufficiently large. The proof is finished. \square

The next result constitutes a key step in the proof.

Lemma 15.2 *For any $k \in \mathbb{N}$, any $\omega \in \Omega_k$, any $y' \in M$ and any $y \in D(\omega, y', \epsilon_k, n_k)$, there exists $(i_{k-1}, j_{k-1}) \in [1, (10k)^{100}]^2$, such that*

$|i_{k-1} - j_{k-1}| \geq (10k)^{95}$, (i_{k-1}, j_{k-1}) is n_k good for ω (see Definition 14.4) and there are u_k, v_k such that $\|u_k\| \leq (\sup |\tau|)n_k$, $\|v_k\| \leq (\sup |\tau|)n_k$, and

$$G_{\tau_{i_{k-1}n_{k-1}}(\omega)}y \in D\left(\sigma^{i_{k-1}n_{k-1}}\omega, G_{u_k}y', \left(1 - \frac{1}{100k^4}\right)\epsilon_{k-1}, n_{k-1}\right),$$

$$G_{\tau_{j_{k-1}n_{k-1}}(\omega)}y \in D\left(\sigma^{j_{k-1}n_{k-1}}\omega, G_{v_k}y', \left(1 - \frac{1}{100k^4}\right)\epsilon_{k-1}, n_{k-1}\right).$$

Before we prove the above lemma, let us show how it implies Proposition 12.5.

Proof of Proposition 12.5 Let $\Lambda_k = \{u : \|u\| \leq (\sup |\tau|)n_k, 100dn_ku \in \mathbb{Z}^d\}$. It is easy to see that $\#\Lambda_k \leq (100d(\sup |\tau|)n_k^2)^d$. Notice that for any ℓ_k with $\|\ell_k\| \leq n_k$ there exists $\ell \in \Lambda_k$ such that $\|\ell_k - \ell\| \leq n_k^{-1}$. Therefore, for any $\bar{\omega} \in \Sigma_A$

$$D\left(\bar{\omega}, G_{\ell_k}y', \left(1 - \frac{1}{100k^4}\right)\epsilon_{k-1}, n_{k-1}\right) \subset D\left(\bar{\omega}, G_{\ell}y', \delta_{k-1}, n_{k-1}\right) \quad (15.4)$$

where $\delta_{k-1} := \left(1 - \frac{1}{100k^4}\right)\epsilon_{k-1} + \frac{1}{n_k}$. Now combining Lemma 15.2 and (15.4) with the choice $\ell_k \in \{u_k, v_k\}$ where u_k, v_k are from Lemma 15.2, we deduce

$$D(\omega, y', \epsilon_k, n_k) \subset \bigcup_{(i_{k-1}, j_{k-1}) \in [1, (10k)^{100}]^2} \bigcup_{u, v \in \Lambda_k} \bigcap_{(w, z) \in \{(i_{k-1}, u), (j_{k-1}, v)\}} G_{-\tau_{w n_{k-1}}(\omega)} \\ D\left(\sigma^{w n_{k-1}}\omega, G_z y', \delta_{k-1}, n_{k-1}\right). \quad (15.5)$$

Fix u, v and $(i, j) = (i_{k-1}, j_{k-1})$. Then by invariance of the measure,

$$v\left(G_{-\tau_{i n_{k-1}}(\omega)}D(\sigma^{i n_{k-1}}\omega, G_u y', \delta_{k-1}, n_{k-1}) \cap G_{-\tau_{j n_{k-1}}(\omega)}D(\sigma^{j n_{k-1}}\omega, G_v y', \delta_{k-1}, n_{k-1})\right) \\ = v\left(G_{\tau_{j n_{k-1}}(\omega) - \tau_{i n_{k-1}}(\omega)}D(\sigma^{i n_{k-1}}\omega, G_u y', \delta_{k-1}, n_{k-1}) \cap D(\sigma^{j n_{k-1}}\omega, G_v y', \delta_{k-1}, n_{k-1})\right). \quad (15.6)$$

Since i, j are n_k good and $|i - j| \geq (10k)^{95}$, it follows by (14.6) that

$$\|\tau_{j n_{k-1}}(\omega) - \tau_{i n_{k-1}}(\omega)\| \geq k^{25}n_{k-1}^{1/2}.$$

Moreover, since i, j are n_k good, by (14.7), for $w \in \{i, j\}$,

$$\sup_{r < n_{k-1}} \|\tau_r(\sigma^{wn_{k-1}}\omega)\| \leq 2k^{20}n_{k-1}^{1/2}.$$

Therefore, by Lemma 13.1 (with $\omega_w = \sigma^{wn_{k-1}}\omega$), it follows that (15.6) is bounded from above by

$$C_{\#} \prod_{w \in \{i, j\}} \nu(D(\sigma^{wn_{k-1}}\omega, G_u y', \delta_{k-1} + 2^{-n_{k-1}^{1/2}}, n_{k-1})). \quad (15.7)$$

Moreover, since i, j are good, $\sigma^{wn_{k-1}}(\omega) \in \Omega_{k-1}$. Also by (12.3), $n_k \leq (1 + 1/100) \cdot 2^{n_k^{1/2}}$. Since $\inf \epsilon_k > 0$ and n_k grows exponentially, using (12.3) again, we have

$$\delta_{k-1} + 2^{-n_{k-1}^{1/2}} = \left(1 - \frac{1}{100k^4}\right)\epsilon_{k-1} + \frac{1}{n_k} + 2^{-n_{k-1}^{1/2}} \leq \epsilon_{k-1}.$$

Using this, we obtain that (15.7) is bounded by $C_{\#}(a_{k-1}(\epsilon_{k-1}))^2$. Using (15.5) and summing over all $u, u' \in \Lambda_k$ and $(i_{k-1}, j_{k-1}) \in [1, (10k)^{100}]^2$ (using that $k^{200} \leq n_k$), we have

$$\begin{aligned} a_k(\epsilon_k) &\leq C_{\#} \cdot [100d(\sup |\tau|)n_k^2]^d \cdot (10k)^{200} \cdot a_{k-1}(\epsilon_{k-1})^2 \\ &\leq \left(10^{200} \cdot C_{\#} \cdot (100d(\sup |\tau|))^d\right) \cdot n_k^{2d+1} a_{k-1}(\epsilon_{k-1})^2. \end{aligned}$$

This by Lemmas 15.1 and 13.2 (with $C_2 = 10^{200} \cdot C_{\#} \cdot (100d(\sup |\tau|))^d$ and $b_k = a_k(\epsilon_k)$) implies that $a_k(\epsilon_k) \rightarrow 0$ which finishes the proof. \square

It remains to prove Lemma 15.2.

Proof of Lemma 15.2 We consider the intervals $[rn_{k-1}, (r+1)n_{k-1}]$. Since $y \in D(\omega, y', \epsilon_k, n_k)$, it follows from the definition of $\{\epsilon_k\}$ that for at least $(10k)^{98}$ of $r < (10k)^{100}$, the points

$$F^{rn_{k-1}}(\omega, y) \text{ and } F^{rn_{k-1}}(\omega', y') \text{ are } \left(\left(1 - \frac{1}{100k^4}\right)\epsilon_{k-1}, n_{k-1}\right)\text{-close.} \quad (15.8)$$

Otherwise the cardinality of $i \leq n_k$ such that $d(F^i(\omega, y), F^i(\omega', y')) < \epsilon_k$ would be bounded above by

$$(10k)^{98}n_{k-1} + ((10k)^{100} - (10k)^{98})n_{k-1} \left(1 - \left(1 - \frac{1}{100k^4}\right)\epsilon_{k-1}\right)$$

$$< (10k)^{100} n_{k-1} \left(1 - \left(1 - \frac{1}{50k^2} \right) \epsilon_{k-1} \right) = n_k (1 - \epsilon_k).$$

This however contradicts the fact that (ω, y) and (ω', y') are (ϵ_k, n_k) -close. So there exists at least $(10k)^{196}$ pairs $(i, j) \in [0, (10k)^{100}]^2$ which satisfy (15.8). Note that

$$\#\{(i, j) \in [0, (10k)^{100}]^2 : |i - j| < (10k)^{95}\} \leq (10k)^{100+95}.$$

Therefore

$$\begin{aligned} \#\{(i, j) \in [0, (10k)^{100}]^2 : (i, j) \text{ satisfies (15.8) and } |i - j| \geq (10k)^{95}\} \\ \geq (10k)^{196} - (10k)^{195}. \end{aligned}$$

Moreover, since $\omega \in \Omega_k$, the cardinality of n_k -good pairs (i, j) (see Definition 14.4) is at least $(10k)^{200} - 5(10k)^{195}$. Since $(10k)^{196} - (10k)^{195} > 5(10k)^{195}$, it follows that there exists (i, j) such that (15.8) holds for $r = i$ and $r = j$, and (i, j) is n_k -good. This means that for $r = i, j$,

$$(\sigma^{rn_{k-1}}\omega', G_{\tau_{rn_{k-1}}(\omega')}y') \quad \text{and} \quad (\sigma^{rn_{k-1}}\omega, G_{\tau_{rn_{k-1}}(\omega)}y) \quad (15.9)$$

are $\left(\left(1 - \frac{1}{100k^4} \right) \epsilon_{k-1}, n_{k-1} \right)$ -close. Hence we find that for some $\|u_k\| \leq (\sup |\tau|)n_k$,

$$G_{\tau_{in_{k-1}}(\omega)}y \in D(\sigma^{in_{k-1}}\omega, G_{u_i}y', (1 - 1/(100k^4))\epsilon_{k-1}, n_{k-1}),$$

and the same holds for j with some v_k . This finishes the proof. \square

Acknowledgements We thank Bassam Fayad and Jean-Paul Thouvenot for useful discussions. D. D. was partially supported by the NSF grant DMS-1956049, C. D. was partially supported by AMS Simons travel grant and Nankai Zhide Foundation, 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.

Part VI. Appendices

Appendix A. Entropy of skew products

Proof of Lemma 2.1. We prove the statement for (T, T^{-1}) diffeomorphisms, the result for flows then follows by considering the time 1 map.

By Ruelle inequality it suffices to show that all Lyapunov exponents of F are non positive.¹⁷ Differentiating (1.3) we get that for each $(x, y) \in (X \times Y)$, $u \in T_x X$, $v \in T_y Y$

$$DF^N(x, y) \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} Df^N u \\ \sum_{j=1}^d d(\tau_{(j)})_N(u) \mathcal{Y}_j + D(G_{\tau_N})(v) \end{pmatrix} (x, y),$$

where $\tau_{(j)}$ denotes the j -th component of τ , $\mathcal{Y}_j = \frac{d}{ds}|_{s=0} G_{s e_j}$ and $\{e_j\}$ is the standard basis in \mathbb{R}^d .

Since f has zero entropy, the Pesin formula shows that the Lyapunov exponents of f are zero. Hence $\lim_{N \rightarrow \infty} \frac{\ln \|Df^N(x)\|}{N} = 0$ for a.e. x . Also since

$$(d(\tau_{(j)})_N(v))(x) = \sum_{n=0}^{N-1} d\tau_{(j)}(f^n x) (Df^n v)$$

it follows that for a.e. x and all $j \in \{1, \dots, d\}$, $\limsup_{N \rightarrow \infty} \frac{\ln \|d(\tau_{(j)})_N(x)\|}{N} \leq 0$.
Also for a.e. (x, y)

$$\limsup_{N \rightarrow \infty} \frac{\ln \|DG_{\tau_N}(x)(y)\|}{N} \leq C \lim_{N \rightarrow \infty} \frac{\|\tau_N(x)\|}{N} = 0$$

where the last step follows since f is ergodic and τ has zero mean.

The foregoing discussion shows that for a.e. (x, y) , $\limsup_{N \rightarrow \infty} \frac{\ln \|DF^N(x, y)\|}{N} \leq 0$. Therefore all Lyapunov exponents of F indeed non positive, and so $ent_\zeta(F) = 0$. \square

Appendix B. Ergodic sums over subshifts of finite type

B.1 CLT for (T, T^{-1}) transformations with SFT in the base

Theorem B.1 Consider a generalized (T, T^{-1}) transformation (1.2) with (X, f) being a subshift of finite type, μ is a Gibbs measure with a Hölder potential, and G_t is an \mathbb{R}^d action which is exponentially mixing of all orders.

¹⁷ Applying this result to F^{-1} gives that all exponents of F are in fact zero, but we do not need this fact for the proof of Lemma 2.1.

Suppose that $d \geq 3$ and $\tau : X \rightarrow \mathbb{R}^d$ is an irreducible Hölder cocycle. Then F satisfies the CLT on the space of Hölder functions.

Remark B.2 As it was mentioned in Sect. 2.4, this result is a special case of Theorem 5.1 in [35]. We include the proof here to make this paper more self contained and to demonstrate the power of Theorem 3.1. We also note that in contrast to [35] the present proof does not rely on the exponential mixing of f , it just uses the the properties of the local distribution of τ such as the anticoncentration inequality (B.2) below.

Proof By Lemma 5.6 it suffices to show that F satisfies the quenched CLT in the sense of Definition 5.1.

We define \mathfrak{m}_N by (5.2) and check the conditions of Proposition 4.1.

(a) is evident.

To prove property (b), let $\ell(x, t, N) = \text{Card}\{n \leq N : |\tau_n(x) - t| \leq 1\}$. We claim that for each p , there is a constant C_p such that for each $t \in \mathbb{R}^d$ for each n

$$\mu(\ell^p(\cdot, t, n)) \leq C_p. \quad (\text{B.1})$$

Indeed,

$$\begin{aligned} \mu(\ell^p(\cdot, t, n)) &\leq \sum_{q=1}^p \hat{C}_p \sum_{n_1 \leq n_2 \leq \dots \leq n_q \leq n} \mu\left(\prod_{j=1}^q 1_{\|\tau_{n_j}(x) - t\| \leq 1}\right) \\ &\leq \sum_{q=1}^p \hat{C}_p \sum_{n_1 < n_2 \leq \dots < n_q \leq n} \mu\left(1_{\|\tau_{n_1}(x) - t\| \leq 1} \left[\prod_{j=2}^q 1_{\|\tau_{n_j - n_{j-1}}(f^{n_{j-1}}x)\| \leq 2} \right]\right). \end{aligned}$$

The multiple anticoncentration inequality of [35, Lemma A.4] tells us that there is a constant \bar{C} such that for each tuple (n_1, \dots, n_q) we have

$$\begin{aligned} \mu\left(1_{\|\tau_{n_1}(x) - t\| \leq 1} \left[\prod_{j=2}^q 1_{\|\tau_{n_j - n_{j-1}}(f^{n_{j-1}}x)\| \leq 2} \right]\right) \\ \leq \bar{C}(n_1)^{-d/2} \left[\prod_{j=2}^q (n_j - n_{j-1})^{-d/2} \right]. \quad (\text{B.2}) \end{aligned}$$

Summing over n_1, \dots, n_q , we obtain (B.1).

With (B.1) proven, the Markov inequality implies that for each ε, t, p we have

$$\mu\left(x : \ell(x, t, N) \geq N^{(1/5)-\varepsilon}\right) \leq \frac{C_p}{N^{[(1/5)-\varepsilon]p}}.$$

It follows that

$$\mu \left(x : \exists t : \|t\| \leq \|\tau\|N \text{ and } \ell(x, t, N) \geq N^{(1/5)-\varepsilon} \right) \leq \frac{C_p^* N^d}{N^{[(1/5)-\varepsilon]p}}.$$

Taking $p = 6d$, $\varepsilon = 0.01$, property (b) follows.

Recall (5.5). In view of Lemma 5.5, to prove property (c) it suffices to check that (5.8) holds for some $\beta > 1$. Using (5.7) we get

$$\begin{aligned} \int_M |\sigma_{0,k}(x)| d\mu(x) &\leq C \sum_{m=0}^{\infty} [\mu(\|\tau_k\| \in [m, m+1]) e^{-cm}] \\ &\leq C \sum_{m=0}^{\infty} \left[\frac{m^{d-1}}{k^{d/2}} e^{-cm} \right] \leq \frac{C}{k^{d/2}}, \end{aligned}$$

where the second inequality relies on (B.2) with $q = 1$ (noting that we can cover the set $\{z \in \mathbb{R}^d : \|z\| \in [m, m+1]\}$ with Cm^{d-1} unit cubes). This shows that (5.8) holds with $\beta = d/2$. This completes the verification of conditions of Proposition 4.1. \square

B.2 Visits to cones

Proof of Lemma 11.4 We only prove the result for the forward orbits, the proof for the backward orbits is similar.

Set $n_1 = 2$, $n_{k+1} = n_k^3$, $m_k = n_k - n_{k-1}$ and consider the sets $A_k = \{\omega : \tau_{n_k}(\omega) \in \mathcal{C}\}$.

Let $\mathcal{F}_{a,b}$ denote the σ -algebra generated by $\{\omega_j\}_{a \leq j \leq b}$. Since τ only depends on the past, A_k is measurable with respect to $\mathcal{F}_{-\infty, n_k}$.

Therefore by Lévy's extension of the Borel–Cantelli Lemma (see e.g. [107, §12.15]) it is enough to show that for almost all ω

$$\sum_k \mu(A_{k+1} | \mathcal{F}_{-\infty, n_k}) = \infty. \quad (\text{B.3})$$

Let $\hat{\mathcal{C}} = \{v \in \mathcal{C} : \text{dist}(v, \partial\mathcal{C}) \geq 1\}$, $\hat{A}_k = \left\{ \omega : \frac{\tau_{m_k}(\sigma^{n_{k-1}}\omega)}{\sqrt{m_k}} \in \hat{\mathcal{C}} \right\}$,

$A_k^* = \{\omega : \exists \hat{\omega} \in \hat{A}_k : \omega_j = \hat{\omega}_j \text{ for } j \in [n_{k-1}, n_k]\}$. Note that $A_k^* \subset A_k$ because for any $\omega \in A_k^*$ and for the corresponding $\hat{\omega}$, $\tau_{m_k}(\sigma^{n_{k-1}}\hat{\omega})$ is inside \mathcal{C} and is at least $\frac{1}{2}\sqrt{m_k}$ away from the boundary whereas

$$\begin{aligned} \tau_{n_k}(\omega) - \tau_{m_k}(\sigma^{n_{k-1}}\hat{\omega}) &= [\tau_{n_k}(\omega) - \tau_{n_k}(\hat{\omega})] + [\tau_{n_k}(\hat{\omega}) - \tau_{m_k}(\sigma^{n_{k-1}}\hat{\omega})] \\ &= O(n_{k-1}) \ll \sqrt{m_k}. \end{aligned}$$

Next

$$\mu(A_{k+1}|\mathcal{F}_{-\infty, n_k}) \geq \mu(A_{k+1}^*|\mathcal{F}_{-\infty, n_k}) \geq \frac{\mu(A_{k+1}^*|\mathcal{F}_{n_k, n_k})}{\hat{K}} \geq \frac{\mu(\hat{A}_{k+1}|\mathcal{F}_{n_k, n_k})}{\hat{K}},$$

where the second inequality is due to (9.2) (note that A_k^* is $\mathcal{F}_{n_{k-1}, n_k}$ -measurable, and hence $\mathcal{F}_{-\infty, n_k}$ -measurable and so (9.2) can be applied), and the third one holds because $A_{k+1}^* \supset \hat{A}_{k+1}$. Since μ is shift invariant

$$\mu(\hat{A}_{k+1}|\mathcal{F}_{n_k, n_k})(\omega) = \mu\left(\frac{\tau_{m_{k+1}}}{\sqrt{m_{k+1}}} \in \hat{\mathcal{C}} \mid \mathcal{F}_{0,0}\right)(\sigma^{-n_k}\omega)$$

By the mixing CLT ([44, 92]) if ω is any symbol in the alphabet of Σ_A

$$\lim_{m \rightarrow \infty} \mu\left(\frac{\tau_m(\omega)}{\sqrt{m}} \in \hat{\mathcal{C}} \mid \omega_0 = \omega\right) = \mathbb{P}(\mathcal{N} \in \hat{\mathcal{C}})$$

uniformly in ω , where \mathcal{N} is the normal random variable with zero mean and variance $D^2(\tau)$ given by (2.9). By the assumptions of Lemma 11.4 and Proposition 2.8, we see that $D^2(\tau)$ is non degenerate. Thus $\mathbb{P}(\mathcal{N} \in \hat{\mathcal{C}}) > 0$ for any cone \mathcal{C} . It follows that there exists $\varepsilon = \varepsilon(\mathcal{C})$ such that for all sufficiently large k and all ω

$$\mu(A_{k+1}|\mathcal{F}_{-\infty, n_k})(\omega) \geq \varepsilon. \quad (\text{B.4})$$

(B.3) follows competing the proof of the lemma. \square

B.3 Separation estimates for cocycles

Proof of Lemma 14.1 (m2) follows from the fact that there exists a constant C_τ such that if ω' and ω'' belong to the same cylinder of length N , then

$$|\tau_N(\omega') - \tau_N(\omega'')| \leq C_\tau.$$

To prove (m1) let

$$N_A(\omega, k) = \#\left\{(i, j) \in [0, (10k)^{100}] \times [0, (10k)^{100}], i \neq j : \frac{\|\tau_{(j-i)n_{k-1}}(\sigma^{in_{k-1}}\omega)\|}{(|j-i|n_{k-1})^{1/2}} < k^{-20}\right\}.$$

Denote $m_{ij} = |i-j|n_{k-1}$. Covering the ball with center at the origin and radius $\frac{\sqrt{m_{ij}}}{k^{20}}$ in \mathbb{R}^d by unit cubes and applying the anticoncentration inequality (B.2)

with $q = 1$ (or [35, formula (A.4)]) to each cube, we obtain that

$$\mu \left(\|\tau_{m_{ij}}(\omega)\| \leq \frac{\sqrt{m_{ij}}}{k^{20}} \right) \leq Ck^{-20d}. \quad (\text{B.5})$$

Since μ is shift invariant we conclude that

$$\mu \left(\frac{\|\tau_{m_{ij}}(\sigma^{in_{k-1}}\omega)\|}{m_{ij}^{1/2}} < \frac{1}{k^{20}} \right) \leq Ck^{-20d}.$$

Summing over i and j we obtain

$$\mu(N_A(\cdot, k)) \leq C(10k)^{200-20d}.$$

Next, by the Markov inequality,

$$\mu(\omega : N_A(\omega, k) \geq (10k)^{191}) \leq \frac{C}{k^{20d-9}}.$$

This shows that the measure of the complement of A_k is small. The estimate of measure of B_k is similar except we replace (B.5) by

$$\mu \left(\max_{n \leq m} \|\tau_n(\omega)\| \geq k^{20}\sqrt{m} \right) \leq c_1 e^{-c_2 k^{40}}. \quad (\text{B.6})$$

To prove (B.6) it is sufficient to consider the case $d = 1$ since for higher dimensions we can consider each coordinate separately. Thus it suffices to show that

$$\mu \left(\max_{n \leq m} \tau_n(\omega) \geq k^{20}\sqrt{m} \right) \leq c_1 e^{-c_2 k^{40}} \quad (\text{B.7})$$

(the bound on $\mu \left(\min_{n \leq m} \tau_n(\omega) \leq -k^{20}\sqrt{m} \right)$ is obtained by replacing τ by $-\tau$).

To prove (B.7) with $d = 1$ we use the reflection principle. Namely, [35, formula (A.3)] shows that for each L

$$\mu(|\tau_m(\omega)| \geq L\sqrt{m}) \leq \bar{c}_1 e^{-\bar{c}_2 L^2}. \quad (\text{B.8})$$

Let

$$D_m(k) = \{\omega : \exists n \leq m, \text{ and } \bar{\omega} : \bar{\omega}_j = \omega_j \text{ for } j \in 0, \dots, n-1 \text{ and } \tau_n(\omega) \geq k^{20}\sqrt{m}\}.$$

Note that $D_m(k)$ contains the LHS of (B.7) and that $D_m(k)$ is a disjoint union of the cylinders of length at most m , $D_m = \bigcup_j \mathcal{D}_j$ (to see this, take for each

ω the smallest n such that the last display holds and recall that τ only depends on the past). Next, similarly to (B.4) (since $d = 1$ the relevant cone is the cone of positive numbers) there exists $\varepsilon > 0$ such that for each cylinder \mathcal{D} of length $n = n(\mathcal{D})$ and for each $m \geq n$,

$$\mu(\tau_{m-n}(\omega) \geq 0 | \omega \in \sigma^{-n}\mathcal{D}) \geq \varepsilon.$$

Combining this with (B.8), we obtain

$$\begin{aligned} \bar{c}_1 e^{-\bar{c}_2 k^{40}/4} &\geq \mu\left(\tau_m \geq \frac{k^{20}\sqrt{m}}{2}\right) \geq \sum_j \mu\left(\omega \in \mathcal{D}_j, \tau_m \geq \frac{k^{20}\sqrt{m}}{2}\right) \\ &\geq \sum_j \mu(\mathcal{D}_j) \mu\left(\tau_m \geq \frac{k^{20}\sqrt{m}}{2} \mid \omega \in \mathcal{D}_j\right) \geq \varepsilon \sum_j \mu(\mathcal{D}_j) = \varepsilon \mu(D_m) \end{aligned}$$

proving (B.7) and completing the proof of the lemma. \square

Appendix C. The main results in general context

Here we put our results into a general context of flexibility of statistical properties in smooth dynamics.

There is a vast literature on statistical properties of dynamical systems. A survey by Sinai [100] lists the following hierarchy of chaotic properties for dynamical systems preserving a smooth measure (the properties marked with * are not on the list in [100] but we added them to obtain a more complete list ¹⁸).

(1) **(Erg)** Ergodicity; (2*) **(WM)** Weak Mixing (3) **(M)** Mixing; (4*) **(PE)** Positive entropy; (5) **(K)** K property; (6) **(B)** Bernoulli property; (7) **(CLT)** Central Limit Theorem¹⁹; (8) **(PM)** Polynomial mixing; (9) **(EM)** Exponential mixing.

Recall that a formal definition of **(CLT)**, **(PM)**, and **(EM)** were given in Sect. 1. The definitions of the other properties are standard.

Properties (1)–(6) are qualitative. They make sense for any measure preserving dynamical system. Properties (7)–(9) are quantitative. They require smooth structure but provide quantitative estimates. Currently there are many examples of systems enjoying a full array of chaotic properties which follow from either uniform hyperbolicity or non-uniform hyperbolicity, in case there

¹⁸ Other interesting statistical properties include Large Deviations, Poisson Limit Theorem, and Local Limit Theorem. We do not include them into our list since our paper does not contain new results or counter examples pertaining to these properties

¹⁹ [100] refers to classical CLT, but since the time it was written several CLTs with non classical normalization has been proven, cf. footnote 4.

is a control on the region where hyperbolicity is weak [11, 14, 26, 109]. 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.

Thus we have the following list of statistical properties of dynamical systems.

(Erg), (WM), (M), (PE), (K), (M), (CLT), (PM), (EM).

While properties on the bottom of the list are often more difficult to establish especially in the context of nonuniformly hyperbolic systems discussed in [100], property (j) of the list in general does not imply property i for $i \leq j$. Thus it is desirable to study the following *realizability problem*: given two disjoint subsets $\mathcal{A}_1, \mathcal{A}_2 \subset \{1, \dots, 9\}$, is there a smooth²⁰ map preserving a smooth probability measure that satisfies all properties in \mathcal{A}_1 and does not have any of the properties in \mathcal{A}_2 ?

The simplest version of the realizability problem is when $|\mathcal{A}_1| = |\mathcal{A}_2| = 1$, which case is presented in the following table. Here Y in cell (i, j) means that the property in row i implies the property in the column j . (k) in cell (i, j) means that a diffeo number (k) on the list below has property (i) but not property (j) .

The examples in the table below are the following (the papers cited in the list contain results needed to verify some properties in the table):

(1) irrational rotation; (2) horocycle flow ([20]); (3) Anosov diffeo \times identity; (4) maps from Theorem 1.3; (5) skew products on $\mathbb{T}^2 \times \mathbb{T}^2$ of the form $(Ax, y + \alpha\tau(x))$ where A is linear Anosov map, α is Liouvillian and τ is not a coboundary [33]; (6) Anosov diffeo \times Diophantine rotation (see [27, 71] and Theorem 3.1).

	Erg	WM/M	PE	K/B	CLT	PM	EM
Erg	♣	(1)	(1)	(1)	(1)	(1)	(1)
WM/M	Y	♣	(2)	(2)	(5)	(5)	(5)
PE	(3)	(3)	♣	(3)	(3)	(3)	(3)
K/B	Y	Y	Y	♣	(5)	(5)	(5)
CLT	Y	(6)	(4)	(6)	♣	(6)	(6)
PM	Y	Y	(2)	(2)	(2)	♣	(2)
EM	Y	Y	Y	Y	??	Y	♣

We combined **(WM)** and **(M)** (as well as **(K)** and **(B)**) together since the same counter examples work for both properties. It is well known that weak

²⁰ Realizability problem also makes sense and is interesting in other settings such as for symbolic or hamiltonian systems.

mixing does not imply mixing (see Sect. 8.3) and that K does not imply Bernoulli (see Part V).

The positive implications in the top left 4×4 corner are standard and can be found in most textbooks on ergodic theory. It is also clear that Exponential Mixing \Rightarrow Polynomial Mixing \Rightarrow Mixing and that CLT implies the weak law of large numbers which in turn entails ergodicity. The fact that the exponential mixing implies the Bernoulli property (and hence both K property and positive entropy) is more recent [37].

The only open problem in the above table, namely the existence of a system satisfying **(EM)** but not **(CLT)** seems hard. Recall from Sect. 4 that the classical CLT follows if the system enjoys exponential mixing of all orders. Therefore the problem whether **(EM)** implies **(CLT)** is related to the question whether exponential mixing implies multiple exponential mixing which can be thought of as a quantitative version of the famous open problem of Rokhlin. Except for this specific question, the realizability problem is well understood in case $|\mathcal{A}_1| = |\mathcal{A}_2| = 1$.

Next, we study the realizability problem with $|\mathcal{A}_1| = 2$, $|\mathcal{A}_2| = 1$ and $\text{CLT} \in \mathcal{A}_1$. The table below lists in cell (i, j) a map which has both property (i) and satisfies CLT but does not have property j . Clearly the question makes sense only if we have an example of a system which has property (i) but not property (j).

WM	M	PE	K	B	PM
WM	♣	(8)	(9)	(9)	(10)
M	♣	♣	(9)	(9)	(10)
PE	(6)	(6)	♣	(6)	(6)
K	♣	♣	♣	(7)	??
B	♣	♣	♣	♣	??
PM	♣	♣	(9)	(9)	♣

Here, (6) refers to the diffeomorphisms from the previous table, while (7), (8), (9), and (10) and refer to the maps from Theorems 1.5, 1.4(a), (b) and 1.3(a). To see that the example of Theorem 1.3(a) is not polynomially mixing we note that for polynomially mixing systems the growth of ergodic integrals can not be regularly varying with index one. Namely (see e.g. [35, §8.1]), for polynomially mixing systems there exists $\delta > 0$ such that the ergodic averages of smooth functions H satisfy $\lim_{T \rightarrow \infty} \frac{H_T}{T^{1-\delta}} = 0$ almost surely, and hence, in law.

Appendix D. Open problems

Here we list some open problems related to our results that we believe should be studied in the future.

In the examples in Theorem 1.3(b), $\dim(M_r)$ grows with r which leads to the following natural problem:

Problem D.1 (a) Construct a C^∞ diffeomorphism with zero entropy satisfying the classical CLT.

(b) Construct a C^∞ flow with zero entropy satisfying the classical CLT.

The next problem is also motivated by Theorem 1.3:

Problem D.2 For which α does there exist a smooth system satisfying the CLT with normalization which is regularly varying of index α ?

We mention that several authors [8, 18, 29, 43] obtained the Central Limit Theorem for circle rotations where normalization is a slowly varying function. However, firstly, the functions considered in those papers are only piecewise smooth and, secondly, they require an additional randomness or remove zero density subset of times. Similar results in the context of substitutions are obtained in [15, 91].

In the examples in Theorem 1.4(b) the rate of polynomial mixing is rather slow (slower than linear). This motivates the following problem:

Problem D.3 Given $m \in \mathbb{N}$ construct a diffeomorphism which is mixing at rate n^{-m} and satisfies at least one of the following: (a) is not K ; (b) has zero entropy; (c) does not satisfy the CLT.

Theorem 1.5 motivates the following problems:

Problem D.4 Construct an example of K (or even Bernoulli) diffeomorphism which satisfies the CLT but is not polynomially mixing.

Problem D.5 Let M a compact manifold of dimension at least two. Does there exists a C^∞ diffeomorphism of M preserving a smooth measure satisfying a Central Limit Theorem?

Currently it is known that any compact manifold of dimension at least two admits an ergodic diffeomorphism of zero entropy [3], a Bernoulli diffeomorphism [17], and, moreover, a nonuniformly hyperbolic diffeomorphism [41]. We note that a recent preprint [97] constructs area preserving diffeomorphisms on any surface of class $C^{1+\beta}$ (with β small) which satisfy (CLT). It seems likely that similar constructions could be made in higher dimensions. However, the method of [97] requires low regularity to have degenerate saddles where a typical orbit does not spent too much time, and so those methods do not work

in higher smoothness such as C^2 . We also note that [21] shows that for any aperiodic dynamical system there exists some measurable observable satisfying the CLT²¹ (see [76, 77, 79, 105] for related results). In contrast Problem D.5 asks to construct a system where the CLT holds for most smooth functions.

Problem D.6 Let M be a compact manifold of dimension at least three. Does there exist a diffeomorphism of M preserving a smooth measure which is K but not Bernoulli?

We note that in case of dimension two, the answer is negative due to Pesin theory [7]. At present there are no example of K but not Bernoulli maps in dimension three. We refer the reader to [62] for more discussion on this problem.

References

1. Aaronson, J., Denker, M.: Local limit theorems for partial sums of stationary sequences generated by Gibbs–Markov maps. *Stoch. Dyn.* **1**, 193–237 (2001)
2. Abramov, L.M., Rokhlin, V.A.: Entropy of a skew product of mappings with invariant measure. *Vestnik Leningrad. Univ.* **17**, 5–13 (1962)
3. Anosov, D.V., Katok, A.B.: New examples in smooth ergodic theory: ergodic diffeomorphisms. *Trans. Moscow Math. Soc.* **23**, 1–35 (1970)
4. Arnold, V.I.: Topological and ergodic properties of closed 1-forms with incommensurable periods. *Funktionalnyi Analiz i Ego Prilozheniya* **25**, 1–12 (1991)
5. Bálint, P., Chernov, N., Dolgopyat, D.: Limit theorems for dispersing billiards with cusps. *Commun. Math. Phys.* **308**, 479–510 (2011)
6. Bálint, P., Gouëzel, S.: Limit theorems in the stadium billiard. *Commun. Math. Phys.* **263**, 461–512 (2006)
7. Barreira, L., Pesin, Y.: Nonuniform Hyperbolicity: Dynamics of Systems with Nonzero Lyapunov Exponents. Cambridge University Press, Cambridge (2007)
8. Beck, J.: Randomness of the square root of 2 and the giant leap. *Period. Math. Hungar.* Part 1: **60**, 137–242 (2010); Part 2: **62**, 127–246 (2011)
9. Björklund, M., Einsiedler, M., Gorodnik, A.: Quantitative multiple mixing. *JEMS* **22**, 1475–1529 (2020)
10. Björklund, M., Gorodnik, A.: Central limit theorems for group actions which are exponentially mixing of all orders. *J. d’Analyse Mathematiques* **141**, 457–482 (2020)
11. Bonatti, C., Diaz, L.J., Viana, M.: Dynamics Beyond Uniform Hyperbolicity: A Global Geometric and Probabilistic Perspective. Springer, Berlin (2005)
12. Bolthausen, E.: A central limit theorem for two-dimensional random walks in random sceneries. *Ann. Probab.* **17**, 108–115 (1989)
13. Bowen, R.: Symbolic dynamics for hyperbolic flows. *Am. J. Math.* **95**, 429–460 (1973)
14. Bowen, R.: Equilibrium States and Ergodic Theory of Anosov Diffeomorphisms. Springer, New York (1975)

²¹ One can also ask which limit distributions can appear in the limit theorems in the context of measurable dynamics and which normalizations are possible. These issues are discussed in [55, 102].

15. Bressaud, X., Bufetov, A.I., Hubert, P.: Deviation of ergodic averages for substitution dynamical systems with eigenvalues of modulus 1. *Proc. Lond. Math. Soc.* **109**, 483–522 (2014)
16. Brin, M.I.: The topology of group extensions of C systems. *Mat. Zametki* **18**, 453–465 (1975)
17. Brin, M.I., Feldman, J., Katok, A.: Bernoulli diffeomorphisms and group extensions of dynamical systems with nonzero characteristic exponents. *Ann. Math.* **113**, 159–179 (1981)
18. Bromberg, M., Ulcigrai, C.: A temporal central limit theorem for real-valued cocycles over rotations. *Ann. Inst. Henri Poincaré Probab. Stat.* **54**, 2304–2334 (2018)
19. Bruin, H.: On volume preserving almost Anosov flows. [arXiv:1908.05675](https://arxiv.org/abs/1908.05675)
20. Bufetov, A., Forni, G.: Limit theorems for horocycle flows. *Ann. Sci. Ec. Norm. Supér.* **47**, 851–903 (2014)
21. Burton, R., Denker, M.: On the central limit theorem for dynamical systems. *Trans. AMS* **302**, 715–726 (1987)
22. Burton, R., Shields, P.: A skew-product which is Bernoulli. *Monatsh. Math.* **86**, 155–165 (1978/79)
23. Butzter, P.L., Westphal, W.: The mean ergodic theorem and saturation. *Indiana Univ. Math. J.* **20**, 1163–1174 (1970/71)
24. Chazottes, J.-R., Gouëzel, S.: Optimal concentration inequalities for dynamical systems. *Commun. Math. Phys.* **316**, 843–889 (2012)
25. Chernov, N.I.: Limit theorems and Markov approximations for chaotic dynamical systems. *Probab. Theory Rel. Fields* **101**, 321–362 (1995)
26. Chernov, N., Markarian, R.: Chaotic billiards. In: *AMS Mathematical Surveys and Monographs*, vol. 127 (2006)
27. Cohen, G., Conze, J.-P.: The CLT for rotated ergodic sums and related processes. *Discrete Contin. Dyn. Syst.* **33**, 3981–4002 (2013)
28. Cohen, G., Conze, J.-P.: CLT for random walks of commuting endomorphisms on compact abelian groups. *J. Theor. Probab.* **30**, 143–195 (2017)
29. Conze, J.-P., Isola, S., Le Borgne, S.: Diffusive behavior of ergodic sums over rotations. *Stoch. Dyn.* **19**, 1950016 (2019)
30. Darling, D.A., Kac, M.: On occupation times for Markoff processes. *Trans. AMS* **84**, 444–458 (1957)
31. den Hollander, F., Keane, M.S., Serafin, J., Steif, J.E.: Weak Bernoullicity of random walk in random scenery. *Jpn. J. Math.* **29**, 389–406 (2003)
32. den Hollander, F., Steif, J.E.: Mixing properties of the generalized T, T^{-1} -process. *J. Anal. Math.* **72**, 165–202 (1997)
33. Dolgopyat, D.: On mixing properties of compact group extensions of hyperbolic systems. *Isr. Math. J.* **130**, 157–205 (2002)
34. Dolgopyat, D.: Limit theorems for partially hyperbolic systems. *Trans. AMS* **356**, 1637–1689 (2004)
35. Dolgopyat, D., Dong, C., Kanigowski, A., Nándori, P.: Mixing properties of generalized T, T^{-1} transformations. *Isr. J. Math.* **2022**, 1–53 (2022)
36. Dolgopyat, D., Fayad, B., Liu, S.: Multiple Borel Cantelli lemma in dynamics and multilog law for recurrence. *J. Mod. Dyn.* [arXiv:2103.08382](https://arxiv.org/abs/2103.08382)
37. Dolgopyat, D., Kanigowski, A., Rodriguez, H.F.: Exponential mixing implies Bernoulli. [arXiv:2106.03147](https://arxiv.org/abs/2106.03147)
38. Dolgopyat, D., Nándori, P.: Non equilibrium density profiles in Lorentz tubes with thermostated boundaries. *Commun. Pure Appl. Math.* **69**, 649–692 (2016)
39. Dolgopyat, D., Nándori, P.: Infinite measure renewal theorem and related results. *Bull. LMS* **51**, 145–167 (2019)

40. Dolgopyat, D., Nándori, P.: On mixing and the local central limit theorem for hyperbolic flows. *Ergod. Theory Dyn. Syst.* **40**, 142–174 (2020)
41. Dolgopyat, D., Pesin, Y.: Every compact manifold carries a completely hyperbolic diffeomorphism. *Ergod. Theory Dyn. Syst.* **22**, 409–435 (2002)
42. Dolgopyat, D., Sarig, O.: Temporal distributional limit theorems for dynamical systems. *J. Stat. Phys.* **166**, 680–713 (2017)
43. Dolgopyat, D., Sarig, O.: Asymptotic windings of horocycles. *Isr. J. Math.* **228**, 119–176 (2018)
44. Eagleson, G.K.: Some simple conditions for limit theorems to be mixing. *Teor. Verojatnost. i Primenen.* **21**, 653–660 (1976)
45. Einsiedler, M., Lind, D.: Algebraic \mathbb{Z}^d -actions of entropy rank one. *Trans. AMS* **356**, 1799–1831 (2004)
46. Erdős, P., Kac, M.: The Gaussian law of errors in the theory of additive number theoretic functions. *Am. J. Math.* **62**, 738–742 (1940)
47. Fayad, B.: Polynomial decay of correlations for a class of smooth flows on the two torus. *Bull. SMF* **129**, 487–503 (2001)
48. Flaminio, L., Forni, G.: Invariant distributions and time averages for horocycle flows. *Duke Math. J.* **119**, 465–526 (2003)
49. Fraczek, K., Lemańczyk, M.: On symmetric logarithm and some old examples in smooth ergodic theory. *Fund. Math.* **180**, 241–255 (2003)
50. Galatolo, S.: Dimension and hitting time in rapidly mixing systems. *Math. Res. Lett.* **14**, 797–805 (2007)
51. Gordin, M.I.: The central limit theorem for stationary processes. *Dokl. Akad. Nauk SSSR* **188**, 739–741 (1969)
52. Gorodnik, A., Spatzier, R.: Mixing properties of commuting nilmanifold automorphisms. *Acta Math.* **215**, 127–159 (2015)
53. Gouëzel, S.: Central limit theorem and stable laws for intermittent maps. *Probab. Theory Rel. Fields* **128**, 82–122 (2004)
54. Gouëzel, S.: Limit theorems in dynamical systems using the spectral method. *Proc. Symp. Pure Math.* **89**, 161–193 (2015)
55. Gouëzel, S.: Growth of normalizing sequences in limit theorems for conservative maps. *Electron. Commun. Prob.* **23**, 99 (2018)
56. Gouëzel, S., Melbourne, I.: Moment bounds and concentration inequalities for slowly mixing dynamical systems. *Electron. J. Probab.* **19**, 30 (2014)
57. Guivarc'h, Y.: Propriétés ergodiques, en mesure infinie, de certains systèmes dynamiques fibrés. *Ergod. Theory Dyn. Syst.* **9**, 433–453 (1989)
58. Hall, P., Heyde, C.: *Martingale Limit Theory and Its Application*. Academic Press, New York (1980)
59. Halmos, P.R.: *Lectures on Ergodic Theory*, Chelsea Publishing, New York, vii+101 pp (1960)
60. Kalikow, S.A.: T, T^{-1} transformation is not loosely Bernoulli. *Ann. Math.* **115**, 393–409 (1982)
61. Kanigowski, A.: Bernoulli property for homogeneous systems. [arXiv:1812.03209](https://arxiv.org/abs/1812.03209)
62. Kanigowski, A., Rodriguez, H.F., Vinhage, K.: On the non-equivalence of the Bernoulli and K properties in dimension four. *J. Mod. Dyn.* **13**, 221–250 (2018)
63. Katok, A.: Smooth non-Bernoulli K-automorphisms. *Invent. Math.* **61**, 291–299 (1980)
64. Katok, A., Kononenko, A.: Cocycles' stability for partially hyperbolic systems. *Math. Res. Lett.* **3**, 191–210 (1996)
65. Katok, A., Katok, S., Schmidt, K.: Rigidity of measurable structure for \mathbb{Z}^d actions by automorphisms of a torus. *Comment. Math. Helv.* **77**, 718–745 (2002)
66. Katok, A., Spatzier, R.: First cohomology of Anosov actions of higher rank abelian groups and applications to rigidity. *Publ. IHES* **79**, 131–156 (1994)

67. Katsuda, A., Sunada, T.: Homology and closed geodesics in a compact Riemann surface. *Am. J. Math.* **110**, 145–155 (1988)
68. Kesten, H., Spitzer, F.: A limit theorem related to a new class of self-similar processes. *Z. Wahrsch. Verw. Gebiete* **50**, 5–25 (1979)
69. Khanin, K.M., Sinai, Y.G.: Mixing for some classes of special flows over rotations of the circle. *Funct. Anal. Appl.* **26**, 155–169 (1992)
70. Khintchine, A.: Zur metrischen Theorie der diophantischen Approximationen. *Math. Zeitschr.* **24**, 706–714 (1926)
71. Kifer, Y.: Limit theorems for random transformations and processes in random environments. *Trans. AMS* **350**, 1481–1518 (1998)
72. Kleinbock, D.Y., Margulis, G.A.: Bounded orbits of nonquasiunipotent flows on homogeneous spaces. *AMS Transl.* **171**, 141–172 (1996)
73. Kochergin, A.V.: Nondegenerate saddles and the absence of mixing. *Math. Notes* **19**, 277–286 (1976)
74. Kochergin, A.V.: Mixing in special flows over a shifting of segments and in smooth flows on surfaces. *Mat. Sb.* **96**, 471–502 (1975)
75. Korepanov, A., Kosloff, Z., Melbourne, I.: Martingale-coboundary decomposition for families of dynamical systems. *Ann. Inst. H. Poincare* **35**, 859–885 (2018)
76. Kosloff, Z., Volný, D.: Local limit theorem in deterministic systems. *Annales de l’Institut Henri Poincaré Prob. Stat.* **58**, 548–566 (2022)
77. Lacey, M.: On central limit theorems, modulus of continuity and Diophantine type for irrational rotations. *J. Anal. Math.* **61**, 47–59 (1993)
78. Ledrappier, F., Sarig, O.: Unique ergodicity for non-uniquely ergodic horocycle flows. *Discrete Contin. Dyn. Syst.* **16**, 411–433 (2006)
79. Lesigne, E.: Almost sure central limit theorem for strictly stationary processes. *Proc. AMS* **128**, 1751–1759 (2000)
80. Le Borgne, S.: Exemples de systemes dynamiques quasi-hyperboliques a decorrelations lentes (preprint); research announcement. *CRAS* **343**, 125–128 (2006)
81. Le Jan, Y.: The central limit theorem for the geodesic flow on noncompact manifolds of constant negative curvature. *Duke Math. J.* **74**, 159–175 (1994)
82. Livsic, A.N.: Cohomology of dynamical systems. *Math. USSR Izvestija* **6**, 1278–1301 (1972)
83. Marcus, B., Newhouse, S.: Measures of maximal entropy for a class of skew products. *Lect. Notes Math.* **729**, 105–125 (1979)
84. Margulis, G.A.: On Some Aspects of the Theory of Anosov Systems. Springer, Berlin (2004)
85. Melijison, I.: Mixing properties of a class of skew-products. *Isr. J. Math.* **19**, 266–270 (1974)
86. Melbourne, I.: Large and moderate deviations for slowly mixing dynamical systems. *Proc. AMS* **137**, 1735–1741 (2009)
87. Moore, C.C.: Ergodicity of flows on homogeneous spaces. *Am. J. Math.* **88**, 154–178 (1966)
88. Ornstein, D.: Bernoulli shifts with the same entropy are isomorphic. *Adv. Math.* **4**, 337–352 (1970)
89. Ornstein, D., Weiss, B.: Entropy and isomorphism theorems for actions of amenable groups. *J. d’Anal. Math.* **8**, 1–141 (1987)
90. Ornstein, D., Weiss, B.: The Shannon–McMillan–Breiman theorem for a class of amenable groups. *Isr. J. Math.* **44**, 53–60 (1983)
91. Paquette, E., Son, Y.: Birkhoff sum fluctuations in substitution dynamical systems. *Ergod. Theory Dyn. Syst.* **39**, 1971–2005 (2019)
92. Parry, W., Pollicott, M.: Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Asterisque* **187–188**, 268 (1990)

93. Peligrad, M., Wu, W.B.: Central limit theorem for Fourier transforms of stationary processes. *Ann. Probab.* **38**, 2009–2022 (2010)
94. Pene, F.: Planar Lorentz process in random scenery. *Ann. Inst. H. Poincare* **45**, 818–839 (2009)
95. Pene, F.: Random Walks in random sceneries and related models. *ESAIM Proc. Surv.* **68**, 35–51 (2020)
96. Pene, F., Thomine, D.: Central limit theorems for the \mathbb{Z}^2 -periodic Lorentz gas. *Isr. J. Math.* **241**, 539–582 (2021)
97. Pesin, Y.B., Senti, S., Shahidi, F.: Area preserving surface diffeomorphisms with polynomial decay of correlations are ubiquitous. [arXiv: 2003.08503](https://arxiv.org/abs/2003.08503)
98. Rokhlin, V.A., Sinai, Y.G.: The structure and properties of invariant measurable partitions (in Russian). *Dokl. Akad. Nauk SSSR* **141**, 1038–1041 (1961)
99. Rudolph, D.: Asymptotically Brownian skew products give non-loosely Bernoulli K-automorphisms. *Invent. Math.* **91**, 105–128 (1988)
100. Sinai, Y.G.: The hierarchy of stochastic properties of deterministic systems. *Encycl. Math. Sci.* **100**, 106–108 (2000)
101. Szász, D., Varjú, T.: Limit laws and recurrence for the planar Lorentz process with infinite horizon. *J. Stat. Phys.* **129**, 59–80 (2007)
102. Thouvenot, J.-P., Weiss, B.: Limit laws for ergodic processes. *Stoch. Dyn.* **12**, 9 (2012)
103. Ulcigrai, C.: Weak mixing for logarithmic flows over interval exchange transformations. *J. Mod. Dyn.* **3**, 35–49 (2009)
104. Ulcigrai, C.: Absence of mixing in area-preserving flows on surfaces. *Ann. Math.* **173**, 1743–1778 (2011)
105. Volný, D.: Invariance principles and Gaussian approximation for strictly stationary processes. *Trans. AMS* **351**, 3351–3371 (1999)
106. Weiss, B.: The isomorphism problem in ergodic theory. *Bull. AMS* **78**, 668–684 (1972)
107. Williams, D.: Probability with Martingales, Cambridge Mathematical Textbooks (1991)
108. Wilkinson, A.: The cohomological equation for partially hyperbolic diffeomorphisms. *Asterisque* **358**, 75–165 (2013)
109. Young, L.-S.: Statistical properties of dynamical systems with some hyperbolicity. *Ann. Math.* **147**, 585–650 (1998)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.