GLOBAL SMOOTH RIGIDITY FOR TORAL AUTOMORPHISMS

BORIS KALININ[1], VICTORIA SADOVSKAYA[2], AND ZHENQI JENNY WANG[3]

Abstract. We study regularity of a conjugacy between a hyperbolic or partially hyperbolic toral automorphism L and a C^∞ diffeomorphism f of the torus. For a very weakly irreducible hyperbolic automorphism L we show that any C^1 conjugacy is C^∞. For a very weakly irreducible ergodic partially hyperbolic automorphism L we show that any $C^{1+\text{Hölder}}$ conjugacy is C^∞. As a corollary, we improve regularity of the conjugacy to C^∞ in prior local and global rigidity results.

1. Introduction and main results

1.1. Hyperbolic systems on tori and their topological classification.

Hyperbolic automorphisms of tori are the prime examples of hyperbolic dynamical systems. Any matrix $L \in GL(d,\mathbb{Z})$ induces an automorphism of the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$, which we denote by the same letter. An automorphism L is *hyperbolic* or *Anosov* if the matrix has no eigenvalues on the unit circle. In this case, $\mathbb{R}^d = E^s \oplus E^u$, where $E^{s/u}$ is the sum of generalized eigenspaces of L corresponding to the eigenvalues of modulus less/greater than 1. In general, a diffeomorphism f of a compact manifold $\mathcal{M}$ is *Anosov* if there exist a continuous Df-invariant splitting $T\mathcal{M} = \mathcal{E}^s \oplus \mathcal{E}^u$ and constants $K > 0$ and $\theta < 1$ such that for all $n \in \mathbb{N}$,

$$\|Df^n(v)\| \le K\theta^n\|v\| \text{ for all } v \in \mathcal{E}^s, \text{ and } \|Df^{-n}(v)\| \le K\theta^n\|v\| \text{ for all } v \in \mathcal{E}^u.$$

The sub-bundles $\mathcal{E}^s$ and $\mathcal{E}^u$ are called stable and unstable. They are tangent to the stable and unstable foliations $\mathcal{W}^s$ and $\mathcal{W}^u$.

Classical results of Franks and Manning [14, 26] establish topological classification of Anosov diffeomorphisms of $\mathbb{T}^d$. Any such diffeomorphism f is topologically conjugate to the hyperbolic automorphism L that f induces on $\mathbb{Z}^d = H_1(\mathbb{T}^d,\mathbb{Z})$. In particular, a hyperbolic automorphism L is topologically conjugate to any C^1-small perturbation. A *topological conjugacy* is a homeomorphism H of $\mathbb{T}^d$ such that

$$L \circ H = H \circ f. \tag{1.1}$$

Any two such conjugacies differ by an affine automorphism of $\mathbb{T}^d$ commuting with L [29], and in particular have the same regularity.

1.2. Regularity of conjugacy for hyperbolic automorphisms.

A conjugacy H in (1.1) is always bi-Hölder, but it is usually not even C^1, as there are various obstructions to smoothness. For example, C^1 regularity of H requires that the Lyapunov exponents of L coincide with the Lyapunov exponents of f with respect to any

Date: May 21, 2025.

Key words: Hyperbolic toral automorphism, partially hyperbolic toral automorphism, conjugacy, rigidity.

[1] Supported in part by Simons Foundation grant 855238.

[2] Supported in part by Simons Foundation grant MP-TSM-00002874.

[3] Supported in part by NSF grant DMS-1845416.

invariant measure, and in particular the uniform measure on any periodic orbit. Thus smoothness of the conjugacy is a rare phenomenon and it is often referred to as rigidity. The problem of regularity of H has been extensively studied and it can be roughly split into two questions:

(Q1) Is the conjugacy C^1 assuming vanishing of some natural obstructions?

(Q2) If the conjugacy is C^1, is it C^∞ for a C^∞ diffeomorphism f?

These questions were often considered in a local setting were f is a small perturbation of L. Answering them in the global setting would give a description of Anosov diffeomorphisms that are smoothly conjugate to algebraic models.

In dimension two, complete positive answers to both questions were obtained by de la Llave, Marco, and Moriyón [9, 4, 5].

The higher dimensional case is much more complicated and answers to both questions are negative in general, without some irreducibility assumption on L. Examples by de la Llave in [5] demonstrated that a conjugacy between a hyperbolic L and its analytic perturbation f can be C^k but is not C^{k+1}, and also that vanishing of periodic obstructions may not yield C^1 conjugacy. Automorphisms L in these examples are products, and hence reducible. We recall that L is *irreducible* if it has no nontrivial rational invariant subspace or, equivalently, if its characteristic polynomial is irreducible over $\mathbb{Q}$.

For large classes of irreducible L, positive answers to question (Q1) were obtained by Gogolev, Guysinky, Kalinin, Sadovskaya, Saghin, Yang, and DeWitt [17, 15, 18, 28, 19, 10, 11] under various assumptions on Lyapunov exponents and periodic data. These results were for the local setting, and recently they were extended to the global setting by DeWitt and Gogolev in [12]. In contrast to the two-dimensional case, they established only $C^{1+\text{Hölder}}$ regularity of H. Smoothness of H on $\mathbb{T}^2$ was obtained along one-dimensional stable and unstable foliations, which have C^∞ leaves. The higher-dimensional results for L with more than one (un)stable Lyapunov exponent showed regularity of H along intermediate invariant foliations corresponding to the Lyapunov splitting of L. However, the leaves of these foliations are typically only $C^{1+\text{Hölder}}$ and hence studying regularity along them cannot yield higher smoothness.

Nevertheless, Gogolev conjectured [15, 12] that the answer to question (Q2) is positive for an irreducible L. Until recently, the only progress in this direction for non-conformal case was the result of Gogolev for automorphisms of $\mathbb{T}^3$ with real spectrum [16]. The key step of his proof was showing that the leaves of the one-dimensional intermediate foliation are C^∞ by studying certain cohomological equation over a Diophantine translation on $\mathbb{T}^3$.

Recently in [24] we used KAM-type techniques to obtain a positive answer to question (Q2) for small perturbations of *weakly irreducible* hyperbolic L. We call a matrix $L \in GL(d, \mathbb{Z})$ weakly irreducible if all factors over $\mathbb{Q}$ of its characteristic polynomial have the same set of moduli of their roots. This assumption is strictly weaker than irreducibility and it is satisfied by some products. The KAM-type techniques required that the diffeomorphism f is close to L in C^r topology, where r depends on L and is typically quite large. Thus this result did not give a conclusive answer to question (Q2) even for the standard local setting with C^1 closeness.

In our new results, we use a completely different approach which allows us to avoid any closeness assumption, and so we obtain a general global regularity result under a very weak irreducibility assumption. For a matrix $L \in GL(d, \mathbb{Z})$ we denote the largest modulus of its eigenvalues by $\rho_{\max}$ and the smallest by $\rho_{\min}$. We say that L is *very weakly irreducible* if every factor over $\mathbb{Q}$ of the characteristic polynomial of L has a root of modulus $\rho_{\max}$

and a root of modulus $\rho_{\min}$. Clearly this condition is weaker than weak irreducibility. Essentially, for a toral automorphism L it means that any algebraic factor of L has the same largest and smallest Lyapunov exponents as L.

Now we state our main result in the hyperbolic case.

Theorem 1.1. *Let L be a very weakly irreducible hyperbolic automorphism of $\mathbb{T}^d$ and let f be a C^∞ diffeomorphism of $\mathbb{T}^d$. If f and L are conjugate by a C^1 diffeomorphism H then any conjugacy between f and L is a C^∞ diffeomorphism.*

Of course f in the theorem is Anosov, but we do not assume that it is close to L. For a C^1-small perturbation of L, certain weak differentiability of H implies that H is $C^{1+\text{Hölder}}$ [24, Theorem 1.1], and so we obtain the following corollary for the local setting.

Corollary 1.2. *Let L be a very weakly irreducible hyperbolic automorphism of $\mathbb{T}^d$ and let f be a C^∞ diffeomorphism of $\mathbb{T}^d$ which is C^1-close to L. If for some conjugacy H between f and L either H or H^{-1} is Lipschitz, or more generally is in a Sobolev space $W^{1,q}(\mathbb{T}^d)$ with $q > d$, then any conjugacy between f and L is a C^∞ diffeomorphism.*

1.3. Regularity of conjugacy for ergodic partially hyperbolic automorphisms.

A toral automorphism L is called *partially hyperbolic* if the matrix has some (but not all) eigenvalues on the unit circle. It is ergodic if and only if none of its eigenvalues is a root of unity. In contrast to the hyperbolic case, there is no topological classification of partially hyperbolic toral diffeomorphisms. The presence of neutral directions yields that even a small perturbation of L may not be topologically conjugate to it. However, if a conjugacy between L and a diffeomorphism f exists, the question of its regularity is interesting. As in the hyperbolic case, any two conjugacies differ by an affine automorphism of $\mathbb{T}^d$ commuting with L, and so have the same regularity.

Prior techniques of showing smoothness of a conjugacy along contracting and expanding foliations do not extend to the neutral foliations. Our methods, however, adapt well to the partially hyperbolic case, and we establish regularity of a conjugacy between L and f without any closeness assumption.

Theorem 1.3. *Let L be a very weakly irreducible partially hyperbolic ergodic automorphism of $\mathbb{T}^d$. If a C^∞ diffeomorphism f of $\mathbb{T}^d$ is $C^{1+\text{Hölder}}$ conjugate to L, then any conjugacy between f and L is a C^∞ diffeomorphism.*

To the best of our knowledge, this is the first bootstrap of regularity result for the partially hyperbolic case. In this theorem we assume $C^{1+\text{Hölder}}$ regularity of H, which is needed for our approach. In contrast to the hyperbolic case, we do not know whether C^1 regularity implies $C^{1+\text{Hölder}}$. This lack of Hölder estimates is one of the reasons why KAM-type iterative argument in [24] did not work in the partially hyperbolic case.

1.4. Applications: improving regularity of conjugacy in prior rigidity results.

Applying Theorems 1.1 and 1.3 we improve the regularity of conjugacy from $C^{1+\text{Hölder}}$ to C^∞ in local and global rigidity results for irreducible hyperbolic and partially hyperbolic toral automorphisms. For a large class of toral automorphisms, these corollaries describe diffeomorphisms which are C^∞ conjugate to them.

The first result is global and is called C^∞ periodic data rigidity of L. It follows from Theorem 1.1 and $C^{1+\text{Hölder}}$ periodic data rigidity, which was recently obtained in [12] as the globalization of the corresponding local result in [18]. The only prior results on C^∞

periodic data rigidity were obtained for automorphisms of $\mathbb{T}^3$ in [12, Theorem 1.1] and for some automorphisms conformal on full stable and unstable bundles [7, 8, 23].

Corollary 1.4. (of Theorem 1.1 and [12, Proposition 1.6])
Let $L : \mathbb{T}^d \to \mathbb{T}^d$ be an irreducible hyperbolic automorphism such that no three of its eigenvalues have the same modulus. Let f be a C^∞ Anosov diffeomorphism of $\mathbb{T}^d$ in the homotopy class of L such that the derivative D_pf^n is conjugate to L^n whenever $p = f^n(p)$. Then f is C^∞ conjugate to L.

We also obtain improvements in local Lyapunov spectrum rigidity results.

Corollary 1.5. (of Theorem 1.1 and [19, Theorem 1.1])
Let $L : \mathbb{T}^d \to \mathbb{T}^d$ be an irreducible hyperbolic automorphism such that no three of its eigenvalues have the same modulus and there are no pairs of eigenvalues of the form $\lambda, -\lambda$ or $i\lambda, -i\lambda$, where $\lambda \in \mathbb{R}$. Let f be a volume-preserving C^∞ diffeomorphism of $\mathbb{T}^d$ sufficiently C^1-close to L. If the Lyapunov exponents of f with respect to the volume are the same as the Lyapunov exponents of L, then f is C^∞ conjugate to L.

A similar result in a partially hyperbolic setting uses stronger assumptions on the spectrum of L and on closeness of f to L. A toral automorphism L is called *totally irreducible* if L^n is irreducible for all $n \in \mathbb{N}$. Such L is ergodic.

Corollary 1.6. (of Theorem 1.3 and [19, Theorem 1.5])
Let $L : \mathbb{T}^d \to \mathbb{T}^d$ be a totally irreducible automorphism with exactly two eigenvalues of modulus one and simple real eigenvalues away from the unit circle. Let f be a volume-preserving C^∞ diffeomorphism sufficiently C^N-close to L, where $N = 5$ if $d > 4$ and $N = 22$ if $d = 4$. If the Lyapunov exponents of f with respect to the volume are the same as the Lyapunov exponents of L then f is C^∞ conjugate to L.

1.5. **Proof strategy.** We reduce the conjugacy equation (1.1) to a cohomological equation over f with *hyperbolic* twist L

$$L\,h - h \circ f = R, \tag{1.2}$$

where $h = H - I$ and $R = f - L$ can be viewed as functions from $\mathbb{T}^d$ to $\mathbb{R}^d$. Projecting this equation to an *unstable* Lyapunov subspace $E^\rho \subset E^u$ for L we obtain

$$L_\rho h_\rho - h_\rho \circ f = R_\rho, \quad \text{where } L_\rho = L|_{E^\rho}. \tag{1.3}$$

The classical approach is to view it as a fixed point equation on h_ρ for an affine contraction on the space $C^0(\mathbb{T}^d, E^\rho)$. It yields existence of the unique solution given by the series

$$h_\rho = \sum_{k=0}^{\infty} L_\rho^{-k-1}(R_\rho \circ f^k), \tag{1.4}$$

and hence the continuous solution $h = \sum_\rho h_\rho$ to (1.2), which was used in the construction of the continuous conjugacy h [14].

In contrast to this, it is much harder to obtain higher regularity of solutions for cohomological equations with hyperbolic twist. We are not aware of any bootstrap of regularity results for solutions of (1.2) with a general Anosov f. In the very special case when f itself is an automorphism, using Fourier coefficients and super-exponential mixing for *linear* f, one can prove bootstrap of regularity for solutions of (1.3) to C^∞ from a sufficiently high regularity, which depends on ρ and the Lyapunov exponents of f, see [3]. Even in this case bootstrap from C^1 fails as low regularity solutions exist in general for reducible cases. We develop new analytical methods allowing us to utilize global information about f given by irreducibility of its linear model L.

Term-wise differentiation of (1.4) can be used to show that h_ρ is C^∞ along the leaves of $\mathcal{W}^s$. However, regularity of h_ρ along $\mathcal{W}^u$ cannot be analyzed in this way, as the series of derivatives diverge since Df^k grows exponentially. The regularity of a conjugacy h was studied using dynamical and geometric tools along invariant foliations of f first (rather than globally). In the case of more than one unstable Lyapunov exponent this roughly corresponds to analyzing regularity of each h_ρ along only the corresponding Lyapunov foliation $\mathcal{W}^\rho$ of f. The leaves of these foliation are only $C^{1+\text{Hölder}}$ in general, and hence this approach is unsuitable for higher regularity. In contrast, we use a novel analytical approach to estimate distributional derivatives of h_ρ along the full $\mathcal{W}^u$ and showing that weak derivatives of all orders are in $L^2(\mathbb{T}^d)$. We use some techniques developed in [13] for the context of $\mathbb{Z}^k$-actions to study regularity of conjugacy along the *neutral* rather than unstable foliation of f (and with neutral twist L_ρ). To analyze regularity of h_ρ along $\mathcal{W}^u$, instead of (1.4) we aim to use a different representation using *negative* iterates of f,

$$h_\rho = -\sum_{k=1}^{\infty} L_\rho^{k-1}(R_\rho \circ f^{-k}). \tag{1.5}$$

However, this series does not converge due to expanding twist L_ρ. The key element of our approach is to differentiate the equation once along the corresponding fast subfoliation $\mathcal{W}$ of $\mathcal{W}^u$. Such derivatives $D^1_{\mathcal{W}} h_\rho$ converge in distributional sense as the contracting derivative Df^{-k} balances the expanding L_ρ^{k-1}. Then we differentiate further and show that all distributional derivatives of $D^1_{\mathcal{W}} h_\rho$ along *all* of $\mathcal{W}^u$ are in $L^2(\mathbb{T}^d)$.

Next we need to obtain weak differentiability of h_ρ from that of $D^1_{\mathcal{W}} h_\rho$. This is the step where we use global information given by very weak irreducibility of L. We recall that the bootstrap fails in general for reducible L. In particular, it is clear that the regularity of h_ρ transverse to $\mathcal{W}$ cannot be recovered locally, even if h_ρ is constant along $\mathcal{W}$. In order to use global structure of L to study nonlinear f, we construct a *global* coordinate chart for the foliation $\mathcal{W}^u$ and the fast part $\mathcal{W}$. This allows us to use a suitable version of Diophantine property for the fast linear foliation of L, which we obtain from very weak irreducibility. Using this Diophantine property we establish a regularity result that allows us to recover sufficient information about all weak derivatives of h_ρ along $\mathcal{W}^u$ from that of $D^1_{\mathcal{W}} h_\rho$.

Our techniques are also well-adapted to treating the partially hyperbolic case, so that bootstrap from $C^{1+\text{Hölder}}$ to C^∞ requires only minor changes compared to hyperbolic one.

2. Proof of Theorem 1.1

Since L and hence f are hyperbolic, we have hyperbolic splittings for L and f

$$\mathbb{R}^d = E^s \oplus E^u \quad \text{for } L \qquad \text{and} \qquad \mathbb{R}^d = \mathcal{E}^s_x \oplus \mathcal{E}^u_x \quad \text{for } f. \tag{2.1}$$

The subspace $E^{s/u}$ is the sum of all generalized eigenspaces of L corresponding to eigenvalues of moduli less/grater than 1. The stable and unstable subbundles $\mathcal{E}^s$ and $\mathcal{E}^u$ of $T\mathbb{T}^d$ are tangent to f-invariant topological foliations $\mathcal{W}^s$ and $\mathcal{W}^u$ with *uniformly C^∞ leaves.* This means that the individual leaves are C^∞ immersed manifolds with the derivatives of any order continuous on $T\mathbb{T}^d$. For any continuous conjugacy H as in (1.1) we have

$$H(\mathcal{W}^s) = W^s \quad \text{and} \quad H(\mathcal{W}^u) = W^u, \tag{2.2}$$

where W^s and W^u are the linear foliations defined by E^s and E^u.

2.1. **The conjugacy H is $C^{1+\text{Hölder}}$.** First we show that the conjugacy H in Theorem 1.1 is in fact $C^{1+\alpha}$ for some $\alpha > 0$. Since H is C^1 we can differentiate the conjugacy equation $L \circ H = H \circ f$ and obtain

$$L \circ DH(x) = DH(fx) \circ Df(x).$$

Thus $C(x) = DH(x)$ is a continuous conjugacy between the derivative cocycles L and Df, that is,

$$L = C(fx) \circ Df(x) \circ C(x)^{-1}.$$

We consider the Lyapunov splitting $\mathbb{R}^d = \oplus_{\rho \in \Delta} E_\rho$ for L, where Δ is the set of moduli of eigenvalues of L and E_ρ is the sum of generalized eigenspaces of L corresponding to the eigenvalues of modulus ρ. Denoting $\mathcal{E}_\rho(x) = C^{-1}(x)E_\rho$ we obtain a continuous Df-invariant splitting $T\mathbb{T}^d = \oplus_{\rho \in \Delta} \mathcal{E}_\rho$ with the same expansion/contraction rates as for L. It is well known that such splitting is α-Hölder with some exponent $\alpha > 0$ which depends on Δ. Then the restriction $C_\rho(x) = C(x)|_{\mathcal{E}_\rho(x)}$ is a continuous conjugacy between α-Hölder linear cocycles $L|_{E_\rho}$ and $Df|_{\mathcal{E}_\rho(x)}$ over f. Both cocycles have one Lyapunov exponent $\log \rho$, and L is clearly fiber bunched as the nonconformality $\|L^n\| \cdot \|L^{-n}\|$ grows at most polynomially and hence is dominated by expansion/contraction along (un)stable manifolds. Hence [24, Theorem 2.1] applies and yields that the conjugacy $C_\rho(x)$ is α-Hölder for each ρ. It follows that $DH(x) = C(x)$ is α-Hölder.

2.2. **General construction and properties of H.** While we do not assume that H is close to the identity map, we may assume that H is homotopic to the identity. Indeed, the induced map $H_* \colon \mathbb{Z}^d \to \mathbb{Z}^d$ on the first homology group of $\mathbb{T}^d$ is given by a matrix $A \in GL(d, \mathbb{Z})$, which defines an automorphism of $\mathbb{T}^d$. Replacing f by $A \circ f \circ A^{-1}$ and H by $H \circ A^{-1}$, we may assume that $H_* = \text{Id}$. In particular, this yields that f is in the homotopy class of L. Since L fixes 0 we see that f fixes $H^{-1}(0)$. Conjugating f by the translation $x \mapsto x + H^{-1}(0)$ we can also assume that $f(0) = L(0) = H(0) = 0$ is a common fixed point for f and L.

From now on we will assume that H is in the homotopy class of the identity, f is in the homotopy class of L, and they satisfy $H(0) = f(0) = 0$. Under these assumptions, the conjugacy H is unique and is given by the following construction. We lift f and H to $\mathbb{R}^d$ as

$$\bar{H} = \text{Id} + \bar{h} \quad \text{and} \quad \bar{f} = L + \bar{R},$$

where $\bar{h}, \bar{R} : \mathbb{R}^d \to \mathbb{R}^d$ are $\mathbb{Z}^d$-periodic functions satisfying $\bar{R}(0) = \bar{h}(0) = 0$. The lifts satisfy the conjugacy equation

$$L \circ \bar{H} = \bar{H} \circ \bar{f} \quad \text{which yields} \quad L \circ \bar{h} - \bar{h} \circ \bar{f} = \bar{R}.$$

The latter equation projects to the following equations for functions on $\mathbb{T}^d$

$$L \circ h - h \circ f = R \quad \text{or} \quad h = L^{-1} \circ h \circ f + L^{-1}R, \tag{2.3}$$

where $h, R : \mathbb{T}^d \to \mathbb{R}^d$ satisfy $R(0) = h(0) = 0$ and are as regular as f and H respectively. Thus we have that R is C^∞ and h is $C^{1+\alpha}$.

Using the hyperbolic splitting $\mathbb{R}^d = E^s \oplus E^u$ for L we define the stable and unstable projections h^s, h^u, R_s, R_u of h and R respectively. Projecting the second equation in (2.3) to E^u we obtain

$$h^u = L_u^{-1}(h^u \circ f) + L_u^{-1} R_u, \quad \text{where } L_u = L|_{E^u}. \tag{2.4}$$

This is a fixed point equation for h^u with the affine operator

$$T_u(\psi) = L_u^{-1}(\psi \circ f) + L_u^{-1} R_u. \tag{2.5}$$

Since $\|L_u^{-1}\| < 1$, the operator T_u is a contraction on the space $C^0(\mathbb{T}^d, E^u)$, and thus h^u is its unique fixed point given by

$$h^u = \lim_{k\to\infty} T_u^k(0) = \sum_{k=0}^{\infty} L_u^{-k-1}(R_u \circ f^k). \tag{2.6}$$

Similarly, one obtains the equation for h^s and the series $h^s = -\sum_{k=1}^{\infty} L_s^{k-1}(R_s \circ f^{-k})$.

Lemma 2.1. *The unstable component h^u is uniformly C^∞ along $\mathcal{W}^s$, that is, it has derivatives of all orders along the leaves of $\mathcal{W}^s$ which vary continuously on $\mathbb{T}^d$. Similarly, h^s is uniformly C^∞ along $\mathcal{W}^u$.*

Proof. This follows from (2.2). Specifically, $H(\mathcal{W}^u) = W^u$ means that the locally defined unstable component of the conjugacy $H^u = \mathrm{Id}^u + h^u$ is locally constant along the leaves of $\mathcal{W}^s$ and hence is as regular as these leaves. Thus h^u is uniformly C^∞ along $\mathcal{W}^s$.

Alternatively, this can be established by term-wise differentiation of the series (2.6) as norms of the derivatives of f^k along $\mathcal{W}^s$ decay exponentially. This is clear for $Df^k|_{\mathcal{E}^s}$, and for higher order derivatives this is given by (5.11) in Lemma 5.2 (i). □

2.3. **Lyapunov splitting and smoothness of h^u.** We will focus on the unstable component and show that h^u is C^∞ on $\mathbb{T}^d$. For this we consider the splitting of the unstable bundles for L and f into Lyapunov subspaces. Let $1 < \rho_1 < \cdots < \rho_\ell$ be the distinct moduli of the unstable eigenvalues of L and let

$$E^u = E^{u,L} = E^1 \oplus E^2 \oplus \cdots \oplus E^\ell \tag{2.7}$$

be the corresponding splitting of E^u, where E^i is the direct sum of generalized eigenspaces corresponding to the eigenvalues with modulus ρ_i. Similarly to the above, we project the conjugacy equation to this splitting. We let $L_i = L|_{E^i}$ and denote by h_i and R_i the E^i components of h and R respectively. Then (2.3) yields

$$L_i h_i(x) - h_i(f(x)) = R_i(x), \qquad x \in \mathbb{T}^d. \tag{2.8}$$

As above, this equation has a unique solution given by the similar series

$$h_i = \sum_{k=0}^{\infty} L_i^{-k-1}(R_i \circ f^k).$$

However, this series is still not useful for studying the regularity of h_i along $\mathcal{W}^u$ as differentiating along $\mathcal{W}^u$ yields diverging series.

The main part of the proof is analyzing regularity of individual h_i along $\mathcal{W}^u$ and showing inductively that they all are C^∞ on $\mathbb{T}^d$. The inductive process is given by the next theorem, which yields that h^u is C^∞. The argument for h^s is similar and hence h and $H = \mathrm{Id} + h$ are also C^∞.

Theorem 2.2. *Suppose $i \in \{1, \dots, \ell\}$. If h_j is C^∞ for all $1 \le j < i$, then h_i is also C^∞.*

In the base case $i = 1$ the assumption becomes vacuous. We emphasize that the regularity of h_i in the theorem is the global regularity on $\mathbb{T}^d$ rather than the regularity along the corresponding foliation.

3. Proof of Theorem 2.2

3.1. **Outline of the proof.** We recall that the conjugacy H is a $C^{1+\alpha}$ diffeomorphism. Using the Lyapunov splitting (2.7) for L we obtain the corresponding splitting for f

$$\mathcal{E}^u = \mathcal{E}^{u,f} = \mathcal{E}^1 \oplus \mathcal{E}^2 \oplus \cdots \oplus \mathcal{E}^\ell \tag{3.1}$$

into α-Hölder Df-invariant sub-bundles $\mathcal{E}^j = DH^{-1}(E^j)$ for $j = 1, \dots, \ell$.

The bundles $\mathcal{E}^i$ are tangent to foliations $\mathcal{W}^i$, which are mapped by H to the corresponding linear ones: $H(\mathcal{W}^j) = W^j$ for all $j = 1, \dots, \ell$. In particular, each $\mathcal{W}^j$ is a $C^{1+\alpha}$ foliation. However, even its individual leaves are not more regular in general. This is the main reason why the regularity of H cannot be bootstrapped by studying its restrictions to $\mathcal{W}^i$. Instead, we show global smoothness of the component h_i on $\mathbb{T}^d$.

We will work with the fast subbundle $\mathcal{E}^{i,\ell}$ and the corresponding foliations $\mathcal{W}^{i,\ell}$, where

$$\mathcal{E}^{i,\ell} = \mathcal{E}^i \oplus \cdots \oplus \mathcal{E}^\ell = T\mathcal{W}^{i,\ell}.$$

We note that each $\mathcal{W}^{i,\ell}$ has uniformly C^∞ leaves, moreover it gives a C^∞ subfoliation of the leaves of $\mathcal{W}^u$. While this holds in general, in our case it follows from Proposition 3.1.

Our approach is to represent the first derivatives of h_i along $\mathcal{W}^{i,\ell}$ as series over the *negative* iterates of f. Using exponential mixing we show that these series converge as distributions as the "expanding twist" by L_i is now balanced by the contracting derivative of f^{-1} along $\mathcal{W}^{i,\ell}$. Then we further differentiate these series in distributional sense along all directions in $\mathcal{W}^u$. We obtain estimates of such derivatives of all orders in terms of fractional Sobolev norms of test functions and then show that they are in $L^2(\mathbb{T}^d)$. This is done in Proposition 3.2.

In Proposition 3.1 we construct an appropriate global coordinate chart which sends foliations $\mathcal{W}^u$ and $\mathcal{W}^{i,\ell}$ to the corresponding linear foliations W^u and $W^{i,\ell}$ for L. This allows us to use the Diophantine property of $W^{i,\ell} = E^{i,\ell}$ to show that *all* derivatives of h_i along $\mathcal{W}^u$ are in $L^2(\mathbb{T}^d)$. This is done in Proposition 3.3. Since by Lemma 2.1 h_i is also uniformly C^∞ along $\mathcal{W}^s$, and we will conclude that h_i is C^∞ on $\mathbb{T}^d$ using [6, Theorem 3].

3.2. **Global charts.** Modifying the $C^{1+\alpha}$ conjugacy H, we now construct a global chart Γ_i for foliations $\mathcal{W}^u$ and $\mathcal{W}^{i,\ell}$ such that Γ_i is C^∞ along $\mathcal{W}^u$.

Proposition 3.1. *Suppose that $i \in \{1, \dots, \ell\}$ and that $h_j \in C^\infty(\mathbb{T}^d)$ for $1 \le j < i$. Then there exists a $C^{1+\alpha}$ diffeomorphism Γ_i of $\mathbb{T}^d$ such that*

(1) *Γ_i maps $\mathcal{W}^u$ to W^u and is uniformly C^∞ along $\mathcal{W}^u$;*
(2) *Γ_i maps $\mathcal{W}^{i,\ell}$ to $W^{i,\ell}$.*

We note that Γ_i may not map $\mathcal{W}^s$ to W^s.

If $i = 1$ in the proposition, we have $\mathcal{W}^{i,\ell} = \mathcal{W}^{1,\ell} = \mathcal{W}^u$ and the assumption that h_j is C^∞ becomes vacuous. The argument applies and gives a $C^{1+\alpha}$ diffeomorphism Γ_1 which is just a smoothing of H along the leaves $\mathcal{W}^u$.

Proof. Recall that the conjugacy $H = \mathrm{Id} + h$, where $h : \mathbb{T}^d \to \mathbb{R}^d$, is a $C^{1+\alpha}$ diffeomorphism. We write $h = (h^s, h_1, \ldots, h_\ell)$ and smooth each h_j, $j \geq i$, while keeping the other components unchanged. More precisely, we define

$$\tilde{h}_\varepsilon = \big(h^s, h_1, \ldots, h_{i-1}, s_\varepsilon(h_i), \ldots, s_\varepsilon(h_\ell)\big), \tag{3.2}$$

where $s_\varepsilon(h_j) \in C^\infty(\mathbb{T}^d)$ is obtained using a standard smoothing by convolution operator. Since $\tilde{h}_\varepsilon \to h$ in C^1 as $\varepsilon \to 0$, we see that $\mathrm{Id} + \tilde{h}_\varepsilon$ is C^1 close to the diffeomorphism H for small ε. It follows that $\mathrm{Id} + \tilde{h}_\varepsilon$ is invertible and hence is a C^1 diffeomorphism. We fix such $\epsilon > 0$ and define

$$\Gamma_i = \mathrm{Id} + \tilde{h}_\varepsilon. \tag{3.3}$$

Since $\tilde{h}_\varepsilon$ remains $C^{1+\alpha}$ on $\mathbb{T}^d$, both Γ_i and Γ_i^{-1} are also $C^{1+\alpha}$.

For $j \geq i$ components $(\tilde{h}_\varepsilon)_j = s_\varepsilon(h_j)$ are C^∞ by smoothing, and for $1 \leq j < i$ components $(\tilde{h}_\varepsilon)_j = h_j$ are C^∞ by the assumption. The stable component $\tilde{h}^s_\varepsilon = h^s$ remains unchanged and hence is uniformly C^∞ along $\mathcal{W}^u$ by Lemma 2.1. Thus the whole $\tilde{h}_\varepsilon$ is uniformly C^∞ along $\mathcal{W}^u$, and hence so is Γ_i.

Also since $\tilde{h}^s_\varepsilon = h^s$, the property $H(\mathcal{W}^u) = W^u$ ensures that $\Gamma_i(\mathcal{W}^u) = W^u$. Similarly, $(\tilde{h}_\varepsilon)_j = h_j$, for $1 \leq j < i$, and so $H(\mathcal{W}^{i,\ell}) = W^{i,\ell}$ yields $\Gamma_i(\mathcal{W}^{i,\ell}) = W^{i,\ell}$. □

3.3. **Derivatives along $\mathcal{W}^u$.** We denote the dimension of E^u by d^u and we fix an orthonormal basis $x_1, \ldots, x_{d^u}$ of E^u such that $x_1, \ldots, x_{\dim E^{i,\ell}}$ is a basis of $E^{i,\ell}$. We denote by $m = (m_1, \ldots, m_{d^u})$ a multi-index with nonnegative integer components, and the corresponding derivative of a function ω on $\mathbb{T}^d$ by

$$D^m_{E^u}(\omega) = \partial^{m_1}_{x_1} \partial^{m_2}_{x_2} \cdots \partial^{m_{d^u}}_{x_{d^u}} \omega.$$

The global foliation chart Γ_i allows us to conveniently describe the derivatives along the foliations $\mathcal{W}^u$ and $\mathcal{W}^{i,\ell}$ in a similar way:

$$D^m_{\mathcal{W}^u}\omega = D^m_{E^u}(\omega \circ \Gamma_i^{-1}) = \partial^{m_1}_{x_1} \partial^{m_2}_{x_2} \cdots \partial^{m_{d^u}}_{x_{d^u}}(\omega \circ \Gamma_i^{-1}).$$

The proof of Theorem 2.2 has two main parts. The first one is the following proposition, which shows that all derivatives of h_i along $\mathcal{W}^u$ are in L^2, as long as the first derivative is taken along $\mathcal{W}^{i,\ell}$.

Proposition 3.2. *Suppose that $i \in \{1, \ldots, \ell\}$ and that h_j are C^∞ for $1 \leq j < i$. Then for every $1 \leq k \leq \dim E^{i,\ell}$ and every multi-index $m = (m_1, \ldots, m_{d^u})$*

$$D^m_{E^u} \partial_{x_k}(h_i \circ \Gamma_i^{-1}) \in L^2(\mathbb{T}^d).$$

We will prove this proposition in Section 5. The proof is dynamical and uses derivative estimates and exponential mixing.

The second part is Proposition 3.3 where we improve the conclusion of Proposition 3.2 by showing that *all* derivatives $D^m_{E^u}(h_i \circ \Gamma_i^{-1})$ are in $L^2(\mathbb{T}^d)$.

Proposition 3.3. *Suppose that $D^m_{E^u} \partial_{x_k}(h_i \circ \Gamma_i^{-1}) \in L^2(\mathbb{T}^d)$ for every $1 \leq k \leq \dim E^{i,\ell}$ and every multi-index $m = (m_1, \ldots, m_{d^u})$. Then*

$$D^m_{E^u}(h_i \circ \Gamma_i^{-1}) \in L^2(\mathbb{T}^d) \quad \textit{for every } m = (m_1, \ldots, m_{d^u}).$$

We will prove this proposition in Section 4. The proof is analytical and relies on a Diophantine property of the linear foliation $E^{i,\ell}$, which we obtain from the very weak irreducibility of L. This is the only place where we use the irreducibility assumption.

Combining Propositions 3.2 and 3.3 we conclude that $D^m_{E^u}(h_i \circ \Gamma_i^{-1}) \in L^2(\mathbb{T}^d)$ for all m, that is, all derivatives $D^m_{\mathcal{W}^u} h_i$ of h_i along $\mathcal{W}^u$ are in $L^2(\mathbb{T}^d)$. Recall that by Lemma 2.1 the map h^u and hence each component h_i are uniformly C^∞ along $\mathcal{W}^s$, that is, $D^m_{\mathcal{W}^s} h_i$ are continuous and so are also in $L^2(\mathbb{T}^d)$. Now we apply [6, Theorem 3]. It yields that all derivatives of order m of a function ϕ are in $L^2(\mathbb{T}^d)$ if ϕ has derivatives of order m in $L^2(\mathbb{T}^d)$ along finitely many transverse foliations, each with uniformly C^∞ leaves and admitting foliation charts whose Jacobians are uniformly C^∞ along the foliation. The last assumption holds for $\mathcal{W}^s$ and $\mathcal{W}^u$ [6, Theorem 2]. Thus we conclude that all derivatives of h_i are in $L^2(\mathbb{T}^d)$, and hence $h_i \in C^\infty(\mathbb{T}^d)$ by Sobolev embedding theorems.

This completes the proof of Theorem 2.2 modulo the proofs of Propositions 3.2 and 3.3. For convenience of the exposition we will prove Proposition 3.3 first.

4. Proof of Proposition 3.3

The proof of Proposition 3.3 is structured as follows. In Section 4.1 we give corollaries of very weak irreducibility and then state and prove a certain Diophantine property for spaces $E^{i,\ell}$. In Section 4.2 we discuss fractional Sobolev spaces $\mathcal{H}^\beta$ that we will use in our regularity results. In Section 4.3 we state and prove the main technical result, Proposition 4.5, and in Section 4.4 we use it to complete the proof of Proposition 3.3.

4.1. **Very weak irreducibility and Diophantine property.** For a matrix $L \in GL(d, \mathbb{Q})$ we denote the largest modulus of its eigenvalues by $\rho_{\max}$. Let $E_{\max}$ be the direct sum of its generalized eigenspaces corresponding to the eigenvalues with modulus $\rho_{\max}$. We denote by $\hat{E}_{\max}$ the direct sum of its generalized eigenspaces corresponding to the eigenvalues with modulus different from $\rho_{\max}$. Thus we have an L-invariant splitting $\mathbb{R}^d = E_{\max} \oplus \hat{E}_{\max}$. We also denote by $(E_{\max})^\perp$ the orthogonal complement of $E_{\max}$ with respect to the standard inner product.

For be the characteristic polynomial p of L we consider its prime decomposition over $\mathbb{Q}$

$$p(t) = \prod_{k=1}^{K} (p_k(t))^{d_k}$$

and the corresponding splitting of $\mathbb{R}^d$ into L-invariant rational subspaces

$$\mathbb{R}^d = \oplus V_k\,, \quad \text{where} \quad V_k = \ker\,(p_k^{d_k}(L)).$$

Lemma 4.1. *For any matrix $L \in GL(d, \mathbb{Q})$ the following are equivalent.*

1. *Each p_k has a root of modulus $\rho_{\max}$,*
2. *$\hat{E}_{\max} \cap \mathbb{Z}^d = \{0\}$,*
3. *$(E_{\max})^\perp \cap \mathbb{Z}^d = \{0\}$.*

Since we focus on $\mathcal{W}^u$ we deal only with $\rho_{\max}$. A similar result holds for $\rho_{\min}$, the smallest modulus of eigenvalues of L, and the corresponding spaces $(E_{\min})^\perp$ and $\hat{E}_{\min}$. It would be used in the proof of smoothness of h^s.

Proof. (2) $\Rightarrow$ (1) If some p_k has no roots of modulus $\rho_{\max}$, then $V_k \subset \hat{E}_{\max}$. Since V_k is a rational subspace, it contains nonzero points of $\mathbb{Z}^d$ and hence so does $\hat{E}_{\max}$.

(1) $\Rightarrow$ (2) Suppose there is $0 \neq n \in (\mathbb{Z}^d \cap \hat{E}_{\max})$. Then for some k the component n_k of n in V_k is nonzero and rational. We note that $n_k \in \hat{E}_{\max}$ as $\hat{E}_{\max} = \oplus_k(\hat{E}_{\max} \cap V_k)$. Then

$$W = span\{L^m n_k : m \in \mathbb{Z}\}$$

is a rational L-invariant subspace contained in $\hat{E}_{\max} \cap V_k$. Then the characteristic polynomial of $L|_W$ is a power of p_k and hence has a root of modulus $\rho_{\max}$ by (1). Thus $W \cap E_{\max} \neq 0$, contradicting $W \subset \hat{E}_{\max}$.

(1) $\Leftrightarrow$ (3) Since the transpose L^τ has the same characteristic polynomial p, (1) is the same as the corresponding property (1^τ) for L^τ, and hence it is equivalent to (2^τ) with the corresponding subspace for L^τ: $\hat{E}^\tau_{\max} \cap \mathbb{Z}^d = \{0\}$. It remans to note that $(E_{\max})^\perp = \hat{E}^\tau_{\max}$. Indeed, the polynomial

$$q(x) = \prod_{|\lambda|=\rho_{\max}} (x-\lambda)^d,$$

where the product is over all eigenvalues of L of modulus $\rho_{\max}$, is real and we obtain

$$(E_{\max})^\perp = (\ker q(L))^\perp = \text{range}\,(q(L))^\tau = \text{range}\,(q(L^\tau)) = \hat{E}^\tau_{\max} \tag{4.1}$$

since $q(L^\tau)$ is invertible on $\hat{E}^\tau_{\max}$ and zero on $E^\tau_{\max}$. $\square$

Definition 4.2. *We say that a subspace V of $\mathbb{R}^d$ has* Diophantine property *if there exists $K > 0$ such that for any $n \in \mathbb{Z}^d$ and any orthonormal basis $\{v_1, \dots, v_{\dim V}\}$ of V we have*

$$\sum_{i=1}^{\dim V} |n \cdot v_i| \geq K\|n\|^{-d}, \tag{4.2}$$

where $\|n\|$ is the standard norm of n and $n \cdot v_i$ is the standard inner product in $\mathbb{R}^d$.

Lemma 4.3. *The space $E^{i,\ell}$ has Diophantine property for each $i \in \{1, \dots, \ell\}$,*

Proof. It suffices to prove the lemma for $i = \ell$ as it gives the smallest subspace E^ℓ. For any orthonormal basis $\{v_1, \dots, v_{\dim E^\ell}\}$ of E^ℓ we have

$$\sum_{j=1}^{\dim E^\ell} |n \cdot v_i| \geq \|\pi_{E^\ell}(n)\| = \text{dist}(n, (E^\ell)^\perp),$$

where π_{E^ℓ} denotes the orthogonal projection to E^ℓ. To complete the proof we need to show that $\text{dist}(n, (E^\ell)^\perp) \geq K\|n\|^{-d}$. Here we use Katznelson's Lemma, see e.g. [3, Lemma 4.1] for a proof.

Lemma 4.4 (Katznelson's Lemma)**.** *Let A be a $d \times d$ integer matrix. Assume that $\mathbb{R}^d$ splits as $\mathbb{R}^d = V_1 \oplus V_2$, where V_1 and V_2 are invariant under A, and $A|_{V_1}$ and $A|_{V_2}$ have no common eigenvalues. If $V_1 \cap \mathbb{Z}^d = \{0\}$, then there exists a constant K such that*

$$dist(n, V_1) \geq K\|n\|^{-d} \quad \textit{for all } 0 \neq n \in \mathbb{Z}^d,$$

where $\|n\|$ denotes Euclidean norm and dist is Euclidean distance.

We apply this lemma with the matrix $A = L^\tau$ and its invariant splitting $\mathbb{R}^d = \hat{E}^\tau_{\max} \oplus E^\tau_{\max}$. We use the notations of Lemma 4.1, so that $E^\ell = E_{\max}$ and by (4.1) we have $(E_{\max})^\perp = \hat{E}^\tau_{\max}$. By very weak irreducibility of L and Lemma 4.1(3) we have $(E_{\max})^\perp \cap \mathbb{Z}^d = \{0\}$. Thus the assumptions of Katznelson's Lemma are satisfied and it yields $\text{dist}(n, (E_\ell)^\perp) \geq K\|n\|^{-d}$ as desired. $\square$

4.2. **Fractional Sobolev spaces.** We will use fractional Sobolev spaces $\mathcal{H}^\beta$ on $\mathbb{T}^d$ which can be defined in terms of Fourier coefficients as follows. For any function $\omega \in L^2(\mathbb{T}^d, \mathbb{C})$ we denote its Fourier coefficients by $\hat{\omega}_n$, $n \in \mathbb{Z}^d$, and write its Fourier series

$$\omega(x) = \sum_{n \in \mathbb{Z}^d} \widehat{\omega}_n e^{2\pi \mathbf{i}\, n \cdot x}.$$

For any $\beta > 0$ we define the norm

$$\|\omega\|_{\mathcal{H}^\beta} = \Big(\sum_{n \in \mathbb{Z}^d} (1 + \|n\|^2)^\beta \cdot |\hat{\omega}_n|^2 \Big)^{1/2} \geq \|\omega\|_{L^2} \tag{4.3}$$

and the fractional Sobolev space

$$\mathcal{H}^\beta = \{\omega \in L^2(\mathbb{T}^d) : \ \|\omega\|_{\mathcal{H}^\beta} < \infty\} \quad \text{and} \quad \mathcal{H}^\beta_0 = \{\omega \in \mathcal{H}^\beta : \ \hat{\omega}_0 = 0\}. \tag{4.4}$$

For $k \in \mathbb{N}$, the space $\mathcal{H}^k$ coincides with the usual Sobolev space $W^{k,2}$ of L^2 functions whose weak derivatives of order up to k are in L^2. By the Sobolev embedding theorem, for any $k, r \in \mathbb{N}$ such that $k > r + d/2$ we have

$$\mathcal{H}^k \subset C^r \quad \text{and} \quad \|\omega\|_{C^r} \leq M \|\omega\|_{\mathcal{H}^k}. \tag{4.5}$$

We will work with $\mathcal{H}^\beta$ for $0 < \beta < 1$ and use the inclusion of the space of α-Hölder functions

$$C^\alpha \subset \mathcal{H}^\beta \quad \text{for any } 0 < \beta < \alpha\ . \tag{4.6}$$

This can be easily seen by using the Hölder estimate in the numerator of the norm

$$(\|\omega\|'_{\mathcal{H}^\beta})^2 = \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \frac{|\omega(x) - \omega(y)|^2}{|x - y|^{d+2\beta}} \, dx\, dy.$$

which is equivalent to the norm (4.3), see e.g. [2].

4.3. **Diophantine regularity result.** Now we state the main analytical result which uses the Diophantine property. It relates differentiability of a function to that of its first derivatives along a foliation V with the Diophantine property. Here we consider distributional derivatives. For a function $\omega \in L^2(\mathbb{T}^d)$ and a C^∞ test function $\psi : \mathbb{T}^d \to \mathbb{C}$ we denote their pairing as

$$\langle \omega, \psi \rangle = \int_{T^d} \omega(x)\, \bar{\psi}(x)\, dx, \tag{4.7}$$

where the integral is with respect to the Lebesgue measure.

For a multi-index $m = (m_1, \dots, m_{\dim E})$ the distributional derivative $D^m_E \omega$ of ω is defined as the functional on the space of C^∞ test functions by

$$\langle D^m_E \omega, \psi \rangle = (-1)^{|m|} \langle \omega, D^m_E \psi \rangle, \quad \text{where } |m| = \sum_{k=1}^{\dim E} m_k. \tag{4.8}$$

We write that $D^m_E \omega \in L^2(\mathbb{T}^d)$ if this distribution is given by an L^2 function.

Proposition 4.5. *Let V be a subspace in $\mathbb{R}^d$ with Diophantine property and let E be any subspace of $\mathbb{R}^d$. Let $\{v_1, \dots, v_{\dim V}\}$ and $\{e_1, \dots, e_{\dim E}\}$ be their orthonormal bases. Suppose that $\omega \in L^2(\mathbb{T}^d)$ satisfies $D^m_E \partial_{v_j} \omega \in L^2(\mathbb{T}^d)$ for every $1 \leq j \leq \dim V$ and every multi-index $m = (m_1, \dots, m_{\dim E})$. Then for any $0 < \beta < 1$ and every multi-index $m = (m_1, \dots, m_{\dim E})$ there exists a constant $K = K(d, m, \beta, V, \omega)$ such that*

$$|\langle D^m_E \omega, \psi \rangle| = |\langle \omega, D^m_E \psi \rangle| \leq K \|\psi\|_{\mathcal{H}^\beta} \quad \textit{for every } C^\infty \textit{ test function } \psi. \tag{4.9}$$

Proof. For any C^∞ function ω, multi-index $m = (m_1, \dots, m_{\dim E})$, and $n \in \mathbb{Z}^d$ we have

$$(\widehat{D_E^m \omega})_n = \langle D_E^m \omega, e^{2\pi \mathbf{i}\, n\cdot x} \rangle = (-1)^{|m|} \langle \omega, D_E^m e^{2\pi \mathbf{i}\, n\cdot x} \rangle = (-1)^{|m|}\, \overline{(2\pi \mathbf{i})^{|m|} n^m}\, \widehat{\omega}_n \tag{4.10}$$

and hence

$$|(\widehat{D_E^m \omega})_n| = (2\pi)^{|m|} |n^m|\, |\widehat{\omega}_n| \quad \text{ where } \quad n^m = \prod_{k=1}^{\dim E} (n \cdot e_k)^{m_k}. \tag{4.11}$$

These equalities hold for any L^2 function ω by the definition of distributional derivative.

First we express the assumption $D_E^m \partial_{v_j} \omega \in L^2(\mathbb{T}^d)$ in terms of Fourier coefficients using $|\widehat{\partial_v \omega_n}| = 2\pi |n \cdot v|\, |\hat{\omega}_n|$ and (4.11)

$$\sum_{n \in \mathbb{Z}^d} |\hat{\omega}_n|^2\, |n^m|^2\, |n \cdot v_j|^2 = (2\pi)^{-2(|m|+1)} \| D_E^m\, \partial_{v_j} \omega \|_{L^2}^2 < \infty.$$

Using Cauchy-Schwarz inequality we obtain

$$\begin{aligned} &\sum_{n \in \mathbb{Z}^d} |\hat{\omega}_n|^2\, |n^m|\, |n \cdot v_j| = \sum_{n \in \mathbb{Z}^d} (|\hat{\omega}_n|\, |n^m|\, |n \cdot v_j|)\, |\hat{\omega}_n| \\ &\le \Big(\sum_{n \in \mathbb{Z}^d} |\hat{\omega}_n|^2\, |n^m|^2\, |n \cdot v_j|^2 \Big)^{1/2} \cdot \Big(\sum_{n \in \mathbb{Z}^d} |\hat{\omega}_n|^2 \Big)^{1/2} \le \; \| D_E^m\, \partial_{v_j} \omega \|_{L^2} \cdot \|\omega\|_{L^2} < \infty. \end{aligned}$$

Using this inductively to divide exponent of $|n^m|\, |n \cdot v_j|$ by 2 we obtain that for any m and any $k \in \mathbb{N}$

$$\begin{aligned} &\sum_{n \in \mathbb{Z}^d} |\hat{\omega}_n|^2\, |n^m|^{2/2^k}\, |n \cdot v_j|^{2/2^k} = \sum_{n \in \mathbb{Z}^d} (|\hat{\omega}_n|\, |n^m|^{2/2^k} |n \cdot v_j|^{2/2^k})\, |\hat{\omega}_n| \le \\ &\le \Big(\sum_{n \in \mathbb{Z}^d} |\hat{\omega}_n|^2\, |n^m|^{2/2^{k-1}} |n \cdot v_j|^{2/2^{k-1}} \Big)^{1/2} \cdot \|\omega\|_{L^2} \; \le \; K_1(k, m, \omega) < \infty. \end{aligned}$$

Since m is arbitrary, taking $m = 2^k m'$ we can rewrite it as

$$\sum_{n \in \mathbb{Z}^d} |\hat{\omega}_n|^2\, |n^{m'}|^2\, |n \cdot v_j|^{2/2^k} \le \; K_1(k, 2^k m', \omega) \quad \text{ for all } k \in \mathbb{N} \text{ and all } m'. \tag{4.12}$$

Informally, the last inequality means that $D_E^{m'} \partial_{v_j}^{1/2^k} \omega \in L^2(\mathbb{T}^d)$. More precisely, for any $\theta \in \mathcal{H}^\beta(\mathbb{T}^d)$ we define the following "fractional derivative"

$$|\partial_{v_j}|^\beta\, \theta \; = \sum_{n \in \mathbb{Z}^d} \widehat{\theta}_n\, |n \cdot v_j|^\beta\, e^{2\pi \mathbf{i}\, n \cdot x} \; \in L^2(\mathbb{T}^d). \tag{4.13}$$

To prove (4.9) we will use (4.12) along with a representation of the test function ψ as a sum of suitable fractional derivatives. The latter is given by the next lemma, which relies on the Diophantine property of V.

Lemma 4.6. *Suppose that $0 < \beta < 1$ and $k \in \mathbb{N}$ satisfy $d/2^k < \beta/2$. Then there exists $K_3 = K_3(V, k)$ such that for any $\psi \in \mathcal{H}^\beta$ there exist*

$$\theta_j \in \mathcal{H}_0^{\beta/2} \;\; \textit{with} \;\; \|\theta_j\|_{\beta/2} \le K_2 \|\psi\|_{\mathcal{H}^\beta}, \quad 1 \le j \le \dim V, \qquad \textit{such that} \tag{4.14}$$

$$\psi = \hat{\psi}_0 + \sum_{j=1}^{\dim V} |\partial_{v_j}|^{1/2^k} \theta_j. \tag{4.15}$$

Proof. For each $0 \neq n \in \mathbb{Z}^d$ we define $\iota(n) \in \{1, \dots, \dim V\}$ to be the smallest index for which $|n \cdot v_{\iota(n)}|$ is maximal, and hence

$$|n \cdot v_{\iota(n)}| \;\geq\; \frac{1}{\dim V} \sum_{i=1}^{\dim V} |n \cdot v_i|.$$

Then we can write $\psi = \hat{\psi}_0 + \sum_{j=1}^{\dim V} \psi_j,$ where the function $\psi_j = \sum_{0 \neq n \in \mathbb{Z}^d} \widehat{\psi}_n \, e^{2\pi \mathbf{i}\, n \cdot x}$ is defined by

$$(\widehat{\psi_j})_n = \widehat{\psi}_n \;\text{ if }\; \iota(n) = j \;\text{ and otherwise }\; (\widehat{\psi_j})_n = 0.$$

Now for each $1 \leq j \leq \dim V$ we construct $\theta_j = \sum_{0 \neq n \in \mathbb{Z}^d} \widehat{\theta}_n \, e^{2\pi \mathbf{i}\, n \cdot x}$ satisfying

$$|\partial_{v_j}|^{1/2^k} \theta_j = \psi_j \quad \text{by taking} \quad (\widehat{\theta_j})_n = |n \cdot v_j|^{-1/2^k} (\widehat{\psi_j})_n \quad \forall\, 0 \neq n \in \mathbb{Z}^d.$$

We note that $(\widehat{\theta_j})_n = 0$ unless $\iota(n) = j$, in which case we have

$$|n \cdot v_j| \;\geq\; \frac{1}{\dim V} \sum_{i=1}^{\dim V} |n \cdot v_i| \;\geq\; K_2(V) \|n\|^{-d}$$

by the Diophantine property (4.2) of V. Thus every nonzero Fourier coefficient of θ_j can be estimated as

$$|(\widehat{\theta_j})_n| = \frac{|(\widehat{\psi_j})_n|}{|n \cdot v_j|^{1/2^k}} \;\leq\; |(\widehat{\psi_j})_n| \, K_3(V,k) \, \|n\|^{d/2^k} \;=\; K_3(V,k) \, |\widehat{\psi}_n| \, \|n\|^{d/2^k}.$$

Since $\psi \in \mathcal{H}^\beta$, see (4.3) and (4.4), it follows that $\theta_j \in \mathcal{H}^{\beta - (d/2^k)}$ and $\|\theta_j\|_{\beta-(d/2^k)} \leq C\|\psi\|_{\mathcal{H}^\beta}$. Since $\frac{d}{2^k} < \frac{\beta}{2}$ we have $\beta - \frac{d}{2^k} > \frac{\beta}{2}$, and hence θ_j satisfies (4.14). $\square$

Now we finish the proof of Proposition 4.5 using Lemma 4.6. We consider $\psi \in \mathcal{H}^\beta$ and choose $k = k(d, \beta)$ sufficiently large so that the lemma applies. Then we can estimate

$$\begin{aligned}
|\langle D_E^m \omega, \psi \rangle| \;&\overset{(1)}{\leq}\; \sum_{j=1}^{\dim V} \left| \left\langle D_E^m \omega,\, |\partial_{v_j}|^{1/2^k} \theta_j \right\rangle \right| \;\overset{(2)}{\leq}\; \sum_{j=1}^{\dim V} \sum_{n \in \mathbb{Z}^d} (2\pi)^{|m|} |\hat{\omega}_n| \, |n^m| \, |n \cdot v_j|^{1/2^k} |(\hat{\theta}_j)_n| \\
&\overset{(3)}{\leq}\; (2\pi)^{|m|} \sum_{j=1}^{\dim V} \Big(\sum_{n \in \mathbb{Z}^d} |\hat{\omega}_n|^2 \, |n^m|^2 \, |n \cdot v_j|^{2/2^k} \Big)^{1/2} \cdot \Big(\sum_{n \in \mathbb{Z}^d} |(\hat{\theta}_j)_n|^2 \Big)^{1/2} \\
&\overset{(4)}{\leq}\; (2\pi)^{|m|} \big(K_1(k, 2^k m, \omega) \big)^{1/2} \sum_{j=1}^{\dim V} \|\theta_j\|_{L^2} \\
&\overset{(5)}{\leq}\; (2\pi)^{|m|} \big(K_1(k, 2^k m, \omega) \big)^{1/2} (K_3 \dim V) \, \|\psi\|_{\mathcal{H}^\beta} = K(\beta, m, d, V, \omega) \, \|\psi\|_{\mathcal{H}^\beta}.
\end{aligned}$$

Here in (1) we use (4.15) and $\langle D_E^m \omega, \hat{\psi}_0 \rangle = 0$, in (2) we use (4.11) and (4.13), in (3) we use Cauchy-Schwarz inequality, in (4) we use (4.12), and in (5) we use (4.14).

This completes the proof of Propositions 4.5. $\square$

4.4. **Completing the proof of Proposition 3.3.** We fix coordinates in E^i and we write the vector-valued function $\omega = h_i \circ \Gamma_i^{-1} : \mathbb{T}^d \to E^i$ as $(\omega_1, \dots, \omega_{\dim E^i})$. We apply Proposition 4.5 with

$$V = E^{i,\ell}, \quad E = E^u, \;\text{ and }\; \omega_j, \;\; j = 1, \dots, \omega_{\dim E^i}.$$

Since $h_i \circ \Gamma_i^{-1}$ is α-Hölder continuous, all ω_j are in $\mathcal{H}^\beta$ for any $\beta < \alpha$ by (4.6).

Now Proposition 3.3 follows from the next lemma, which upgrades the derivatives to L^2. While the statement is natural, we do not know if it holds with $\mathcal{H}^\beta$ replaced by C^β.

Lemma 4.7. *Suppose that $\omega \in \mathcal{H}^\beta$ and that for some multi-index m and $K > 0$ we have*

$$|\langle D_E^{2m}\omega, \psi\rangle| \le K\|\psi\|_{\mathcal{H}^\beta} \tag{4.16}$$

for every C^∞ test function ψ. Then $D_E^m\omega \in L^2(\mathbb{T}^d)$.

Proof. We consider the smoothing of ω by truncation of its Fourier series,

$$\omega_N = \sum_{n\in\mathbb{Z}^d,\, \|n\|\le N} \widehat{\omega}_n e^{2\pi \mathbf{i}\, n\cdot x}. \tag{4.17}$$

Each ω_N is $C^\infty$and satisfies $\|\omega_N\|_{\mathcal{H}^\beta} \le \|\omega\|_{\mathcal{H}^\beta}$. Then (4.16) and (4.11) yield

$$\begin{aligned} K\|\omega\|_{\mathcal{H}^\beta} \;&\ge\; K\|\omega_N\|_{\mathcal{H}^\beta} \;\ge\; |\langle D_E^{2m}\omega,\, \omega_N\rangle| \;=\; |\langle \omega,\, D_E^{2m}\omega_N\rangle| \;=\; \Big|\int_{\mathbb{T}^d} \omega(x)\, \overline{D_E^{2m}\omega_N(x)}\, dx\,\Big| \\ &= \Big|\sum_{\|n\|\le N} \widehat{\omega}_n\, \overline{\widehat{(D_E^{2m}\omega)}_n}\,\Big| \;=\; \Big|\sum_{\|n\|\le N} \widehat{\omega}_n\, (2\pi\mathbf{i})^{2|m|}\, n^{2m}\, \overline{\widehat{\omega}_n}\,\Big| \\ &= \sum_{\|n\|\le N} (2\pi)^{2|m|}\, n^{2m}\, |\widehat{\omega}_n|^2 \;=\; \sum_{\|n\|\le N} \big((2\pi)^{|m|}|n^m|\, |\widehat{\omega}_n|\big)^2 \qquad \text{for any } N\in\mathbb{N}. \end{aligned}$$

We conclude using (4.11) that

$$\sum_{n\in\mathbb{Z}^d} |\widehat{(D_E^m\omega)}_n|^2 = \sum_{n\in\mathbb{Z}^d} \big((2\pi)^{|m|}|n^m|\, |\widehat{\omega}_n|\big)^2 \le K\|\omega\|_{\mathcal{H}^\beta}$$

and hence the distribution $D_E^m\omega$ is given by the L^2 function $\sum_{n\in\mathbb{Z}^d} \widehat{(D_E^m\omega)}_n\, e^{2\pi\mathbf{i}\, n\cdot x}$. $\square$

5. Proof of Proposition 3.2

We prove the proposition in Section 5.1. In Section 5.2 we state and prove derivative estimates that are used in this proof as well as in Section 6.

5.1. **Main part of the proof.** For a vector-valued function $f = (f_1,\dots,f_N) : \mathbb{T}^d \to \mathbb{R}^N$ we define its norm as the maximum $\|f\| = \max_i \|f_i\|$ of the corresponding norms of the components. We also adopt vector-valued notations for the inner product in $L^2(\mathbb{T}^d)$ with respect to the Lebesgue measure and for pairing of a distribution with a test function $\psi : \mathbb{T}^d \to \mathbb{R}$

$$\langle f, \psi\rangle := \big(\langle f_1,\psi\rangle, ,\cdots, \langle f_N,\psi\rangle\big) \in \mathbb{R}^N. \tag{5.1}$$

First, we obtain a suitable series representation for the distribution $D_{E^u}^m \partial_{x_j}(h_i\circ\Gamma_i^{-1})$, where $1 \le j \le \dim E^{i,\ell}$. Specifically, we will show that

$$\big\langle h_i\circ\Gamma_i^{-1},\, D_{E^u}^m\partial_{x_j}\psi\big\rangle \;=\; \sum_{k=1}^{\infty} L_i^{k-1}\big\langle R_i\circ f^{-k}\circ\Gamma_i^{-1},\, D_{E^u}^m\partial_{x_j}\psi\big\rangle. \tag{5.2}$$

As we discussed in the introduction, this is different from differentiating series (1.4) for h_i and instead corresponds to differentiating series (1.5). However, (1.5) does not converge in C^0 and even its distributional convergence is a priory not clear. In contrast, the series (5.2) converges distributionally. To show this we work directly with a finite iteration of the conjugacy equation (2.8). Rewriting $L_ih_i(x) - h_i(f(x)) = R_i(x)$ we obtain

$$h_i(x) = L_ih_i(f^{-1}(x)) - R_i(f^{-1}(x)) \quad \text{for all } \; x\in\mathbb{T}^d.$$

Iterating this, we see that for any $n \in \mathbb{N}$,

$$h_i = L_i^n h_i \circ f^{-n} - \sum_{k=1}^{n} L_i^{k-1} R_i \circ f^{-k},$$

and hence

$$h_i \circ \Gamma_i^{-1} \;=\; L_i^n h_i \circ f^{-n} \circ \Gamma_i^{-1} - \sum_{k=1}^{n} L_i^{k-1} R_i \circ f^{-k} \circ \Gamma_i^{-1}.$$

Now we show that the error term $L_i^n h_i \circ f^{-n} \circ \Gamma_i^{-1}$ paired with $D^m_{E^u} \partial_{x_j} \psi$ tends to zero as $n \to \infty$. Recalling that Γ_i, Γ_i^{-1}, H, H^{-1}, and h_i are $C^{1+\alpha}$ we obtain

$$\begin{aligned}
(5.3) \qquad & L_i^n \left\langle h_i \circ f^{-n} \circ \Gamma_i^{-1},\, D^m_{E^u} \partial_{x_j} \psi \right\rangle \\
&\overset{(1)}{=} -L_i^n \left\langle \partial_{x_j} (h_i \circ f^{-n} \circ \Gamma_i^{-1}),\, D^m_{E^u} \psi \right\rangle \\
&\overset{(2)}{=} -L_i^n \left\langle \partial_{x_j} (h_i \circ H^{-1} \circ L^{-n} \circ H \circ \Gamma_i^{-1}),\, D^m_{E^u} \psi \right\rangle \\
&= -L_i^n \left\langle D_{E^{i,\ell}} (h_i \circ H^{-1})_{L^{-n} \circ H \circ \Gamma^{-1} x} \cdot L^{-n}|_{E^{i,\ell}} \cdot \partial_{x_j} (H \circ \Gamma_i^{-1})_x,\, D^m_{E^u} \psi \right\rangle \\
&\overset{(3)}{=} -\rho_i^{-n} L_i^n \left\langle D_{E^{i,\ell}} (h_i \circ H^{-1})_{L^{-n} y} \cdot B_n \cdot \partial_{x_j} (H \circ \Gamma^{-1})_{\Gamma_i \circ H^{-1} y},\, D^m_{E^u} (\psi)_{\Gamma_i \circ H^{-1} y} \cdot J(y) \right\rangle
\end{aligned}$$

Here in (1) we integrate by parts, noting that the left side is $C^{1+\alpha}$ and so the pairing is given by integration; in (2) we use that H conjugates L^{-n} and f^{-n}; in (3) we denote $B_n = \rho_i^n L^{-n}|_{E^{i,\ell}}$, change the variable to $y = H \circ \Gamma_i^{-1}(x)$, and denote by $J(y)$ the Jacobian of the coordinate change.

To estimate (5.3) we will use exponential mixing of Hölder functions with respect to the volume for ergodic automorphisms of tori. It was obtained by Lind [25, Theorem 6] and by Gorodnik and Spatizer [21, Theorem 1.1] for nilmanifolds, which use for the estimate.

Theorem 5.1. [25, 21] *Let L be an ergodic automorphism of $\mathbb{T}^d$. Then for any $0 < \alpha < 1$ and any $g_1,\, g_2 \in C^\alpha(\mathbb{T}^d)$ there exists $\gamma = \gamma(\alpha) > 0$ and $K = K(\alpha)$ such that*

$$\left| \langle g_1 \circ L^n, g_2 \rangle - \Big(\int_{\mathbb{T}^d} g_1(x) dx \Big) \Big(\int_{\mathbb{T}^d} g_2(x) dx \Big) \right| \;\le\; K \|g_1\|_{C^\alpha} \|g_2\|_{C^\alpha} \, e^{-\gamma |n|} \ \text{ for all } n \in \mathbb{Z}.$$

Now we estimate the last line of (5.3). The norms $\|\rho_i^{-n} L_i^n\|$ and $\|B_n\|$ grow at most as $K_1 n^d$ since the moduli of all their eigenvalues are at most one. Also, the derivative $D_{E^{i,\ell}}(h_i \circ H^{-1})$ is α-Hölder with zero average, and $\partial_{x_j}(H \circ \Gamma^{-1})_{\Gamma_i \circ H^{-1} y}$ and $J(y)$ are α-Hölder functions as well. Hence we can use Theorem 5.1 to estimate

$$\begin{aligned}
& \|\rho_i^{-n} L_i^n \big\langle D_{E^{i,\ell}} (h_i \circ H^{-1})_{L^{-n} y} \cdot B_n \cdot \partial_{x_j} (H \circ \Gamma_i^{-1})_{\Gamma_i \circ H^{-1} y},\, D^m_{E^u} (\psi)_{\Gamma_i \circ H^{-1} y} \cdot J(y) \big\rangle \| \\
(5.4) \qquad & \le K_2 \, n^{2d} \, \| D_{E^{i,\ell}} (h_i \circ H^{-1}) \|_{C^\alpha} \, \| \partial_{x_j} (H \circ \Gamma_i^{-1})_{\Gamma_i \circ H^{-1} y} \cdot D^m_{E^u}(\psi) \cdot J(y) \|_{C^\alpha} \cdot e^{-\gamma n}.
\end{aligned}$$

Thus the pairing (5.3) decays exponentially, specifically, there is $C = C(K_2, h_i, H, \Gamma_i, J)$ such that

$$(5.5) \qquad \| L_i^n \left\langle h_i \circ f^{-n} \circ \Gamma_i^{-1},\, D^m_{E^u} \partial_{x_j} \psi \right\rangle \| \le C n^{2d} e^{-\gamma n} \, \| D^m_{E^u} (\psi) \|_{C^\alpha} \ \text{ for all } n \in \mathbb{N}.$$

Now we estimate the terms in representation (5.2), which are similar to the error term (5.3) with h_i replaced by R_i. However, we want to estimate by Hölder norm of ψ in place of

$\|D^m_{E^u}(\psi)\|$ and so we want to move the derivative $D^m_{E^u}$ to the left and use differentiability of $R = f - L$:

$$\big\langle R_i \circ f^{-k} \circ \Gamma_i^{-1},\, D^m_{E^u}\partial_{x_j}\psi\big\rangle = (-1)^{|m|+1}\big\langle D^m_{E^u}\partial_{x_j}(R_i \circ f^{-k} \circ \Gamma_i^{-1}),\, \psi\big\rangle. \tag{5.6}$$

However, having higher order derivatives on the left does not allow to use exponential mixing directly.

Instead, we split ψ into its truncation smoothing ψ_N as in (4.17) and the error $\psi - \psi_N$. Then for any $\beta > 0$ and $N \geq 1$ the following estimates hold

$$\|\psi - \psi_N\|_{L^2} \leq N^{-\beta}\|\psi\|_{\mathcal{H}^\beta} \quad \text{and} \tag{5.7}$$

$$\|\psi_N\|_{\mathcal{H}^\beta} \leq 2^{\beta/2}\, N^\beta \|\psi\|_{L^2}. \tag{5.8}$$

Indeed,

$$\|\psi - \psi_N\|^2_{L^2} = \sum_{\|n\|>N} |\widehat{\psi}_n|^2 \leq N^{-2\beta} \sum_{\|n\|>N} \|n\|^{2\beta}|\widehat{\psi}_n|^2 \leq N^{-2\beta}\|\psi\|^2_{\mathcal{H}^\beta} \quad \text{and}$$

$$\|\psi_N\|^2_{\mathcal{H}^\beta} = \sum_{\|n\|\leq N} (1+\|n\|^2)^\beta |\widehat{\psi}_n|^2 \leq (1+N^2)^\beta \sum_{\|n\|>N} |\widehat{\psi}_n|^2 \leq 2^\beta N^{2\beta}\|\psi\|^2_{L^2}.$$

To estimate the pairing with ψ_N we use bounds (5.8) on the derivative $D^m_{E^u}\psi_N$ in terms of ψ. This allows us to repeat the estimates (5.3) and (5.5) above replacing n by k, h_i by R_i, and ψ by ψ_N. Thus we obtain

$$\begin{aligned}
&\left\| L_i^{k-1}\big\langle R_i \circ f^{-k} \circ \Gamma_i^{-1},\, D^m_{E^u}\partial_{x_j}(\psi_N)\big\rangle \right\| \leq Ck^{2d} \cdot \|D^m_{E^u}(\psi_N)\|_{C^\alpha} \cdot e^{-\gamma k} \\
&\leq C_1 k^{2d} \cdot \|\psi_N\|_{C^{|m|+1}} \cdot e^{-\gamma k} \leq C_1 k^{2d} \cdot \|\psi_N\|_{\mathcal{H}^{|m|+1+d}} \cdot e^{-\gamma k} \quad \text{using (4.5)} \\
&\leq C_2 k^{2d} \cdot N^{|m|+d+1} \cdot \|\psi\|_{L^2} \cdot e^{-\gamma k} \quad \text{using (5.8).}
\end{aligned}$$

To estimate the pairing with $\psi_N - \psi$ we use the bounds on norms higher order derivatives of f, which we obtain in Lemma 5.3 in Section 5.2 below. Specifically, using the second part of (5.15) with $g = f^{-1}$, $\phi = R_i$, $\mathcal{W} = \mathcal{W}^u$, $\mathcal{W}' = \mathcal{W}^{i,\ell}$, and $\lambda = \rho_i^{-1}$ we obtain that for each m and $\delta > 0$ there exists $K = K(\delta, m, \|R_i|_{\mathcal{W}^u}\|_{C^{m+1}})$ such that for all $k \in \mathbb{N}$,

$$\|D^m_{\mathcal{W}^u}(D_{\mathcal{W}^{i,\ell}}(R_i \circ f^{-k}))\|_{C^0} \leq K(\rho_i^{-1} + \delta)^k\, \|R_i|_{\mathcal{W}^u}\|_{C^{m+1}}.$$

Using this and (5.6) we estimate

$$\begin{aligned}
&\left\| L_i^{k-1}\big\langle R_i \circ f^{-k} \circ \Gamma_i^{-1},\, D^m_{E^u}\partial_{x_j}(\psi - \psi_N)\big\rangle \right\| \\
&= \left\| L_i^{k-1}\big\langle D^m_{E^u}\partial_{x_j}(R_i \circ f^{-k} \circ \Gamma_i^{-1}),\, (\psi - \psi_N)\big\rangle \right\| \\
&\leq \|L_i^{k-1}\| \cdot \|D^m_{\mathcal{W}^u}(D_{\mathcal{W}^{i,\ell}}(R_i \circ f^{-k}))\|_{C^0} \cdot \|\psi - \psi_N\|_{L^2} \\
&\leq Ck^d \rho_i^k \cdot \|R_i|_{\mathcal{W}^u}\|_{C^{m+1}} \cdot K(\rho_i^{-1} + \delta)^k \cdot \|\psi - \psi_N\|_{L^2} \\
&\leq C_3 k^d \cdot (1 + \rho_i\delta)^k \cdot N^{-\beta}\|\psi\|_{\mathcal{H}^\beta} \quad \text{using (5.7).}
\end{aligned}$$

Finally we combine the two estimates above to get an estimate for terms in (5.2)

$$\|L_i^{k-1}\big\langle R_i \circ f^{-k} \circ \Gamma_i^{-1},\, D^m_{E^u}\partial_{x_j}\psi\big\rangle\| \;\leq\; C_4 k^{2d}\, \|\psi\|_{\mathcal{H}^\beta}\, (N^{|m|+d+1}e^{-\gamma k} + (1+\rho_i\delta)^k N^{-\beta}).$$

For each k we choose $N = N(k) = e^{\gamma k/(2(|m|+d+1))}$ so that we can write the last term as

$$N^{|m|+d+1}e^{-\gamma k} + (1+\rho_i\delta)^k N^{-\beta} = e^{-\gamma k/2} + (1+\rho_i\delta)^k e^{-\beta\gamma k/(2(|m|+d+1))}$$

and get an exponentially converging series for small enough δ.

We conclude that there is $0<\xi<1$ and a constant C such that for all $k\in\mathbb{N}$ we have

$$\|L_i^{k-1}\langle R_i\circ f^{-k}\circ\Gamma_i^{-1},\, D^m_{E^u}\partial_{x_j}\psi\rangle\|\le C\xi^k\|\psi\|_{\mathcal{H}^\beta} \tag{5.9}$$

and hence

$$\|\langle D^m_{E^u}\partial_{x_j}(h_i\circ\Gamma_i^{-1}),\,\psi\rangle\|=\|\sum_{k=1}^{\infty}L_i^{k-1}\langle R_i\circ f^{-k}\circ\Gamma_i^{-1},\, D^m_{E^u}\partial_{x_j}\psi\rangle\|\le C_m\|\psi\|_{\mathcal{H}^\beta}. \tag{5.10}$$

Now we complete the proof of Proposition 3.2 by showing that $D^m_{E^u}\partial_{x_j}(h_i\circ\Gamma_i^{-1})\in L^2(\mathbb{T}^d)$ for any m. We denote $\omega=\partial_{x_j}(h_i\circ\Gamma_i^{-1})$ and recall that it is α-Hölder since the conjugacy H and the global chart Γ are $C^{1+\alpha}$ diffeomorphisms. It follows that ω is in $\mathcal{H}^\beta$ for any $\beta<\alpha$ by (4.6). Then (5.10) shows that for any m and any C^∞ function ψ we have

$$|\langle D^{2m}_{E^u}\omega,\,\psi\rangle|=|\langle\omega,\,D^{2m}_{E^u}\psi\rangle|\le C_{2m}\|\psi\|_{\mathcal{H}^\beta}.$$

Hence we can apply Lemma 4.7 to conclude that $D^m_{E^u}\omega$ is in $L^2(\mathbb{T}^d)$.

5.2. **Derivative estimates.** In this section we prove the derivative estimates used above. While estimates of derivatives of compositions along the stable manifold are not new, a precise reference is hard to give as we need a specific result involving faster part. We will also need similar estimates along the center foliation in the next section. For a function ϕ, a foliation $\mathcal{W}$, and $m\in\mathbb{N}$ we will denote by $D^m_{\mathcal{W}}\phi$ the derivative of order m of ϕ restricted to the leaves of $\mathcal{W}$. We view it here as an m-linear form on $T\mathcal{W}$ and we denote its norm by $\|D^m_{\mathcal{W}}\phi\|$.

Lemma 5.2. *Let $\mathcal{W}$ be a foliation of $\mathbb{T}^d$ with uniformly C^∞ leaves invariant under a C^∞ diffeomorphism g such that $\|Dg|_{\mathcal{W}}\|\le\sigma$.*

(i) *If $\sigma<1$ then for each m and $\delta>0$ there exists $C=C(\delta,m,\|g|_{\mathcal{W}}\|_{C^m})$ such that*

$$\|D^m_{\mathcal{W}}g^n\|\le C(\sigma+\delta)^n\quad\text{and}\quad\|D^m_{\mathcal{W}}(\phi\circ g^n)\|\le C(\sigma+\delta)^n\|\phi|_{\mathcal{W}}\|_{C^m} \tag{5.11}$$

for any $\phi\in C^\infty(\mathbb{T}^d)$, where $D^m_{\mathcal{W}}$ is the derivative of order m along $\mathcal{W}$.

(ii) *If $\sigma>1$ then for each m there exists $C=C(m,\|g|_{\mathcal{W}}\|_{C^m})$ such that*

$$\|D^m_{\mathcal{W}}g^n\|\le C\sigma^{mn}\quad\text{and}\quad\|D^m_{\mathcal{W}}(\phi\circ g^n)\|\le C\sigma^{m^2n}\|\phi|_{\mathcal{W}}\|_{C^m}\quad\text{for any }\phi\in C^\infty(\mathbb{T}^d). \tag{5.12}$$

Proof. (i) We abbreviate $D_{\mathcal{W}}$ to D in this proof. We will show inductively that for some c and all $m\le m_0$ we have $\|D^mg^n\|\le c^{m-1}(\sigma+\delta)^n$ for all $n\in\mathbb{N}$. In the base case $m=1$ by the assumption we have $\|Dg^n\|\le\sigma^n$ for all $n\in\mathbb{N}$. Suppose the estimate holds for derivatives of orders up to $m-1$.

Now we show that $\|D^mg^n\|\le c^{m-1}(\sigma+\delta)^n$ for all $n\in\mathbb{N}$ by induction on n. The base case $n=0$ is trivial. For the inductive step we apply Faà di Bruno's formula to $D^mg^{n+1}=D^m(g\circ g^n)$. We slightly abuse notations by suppressing the base points, as they are not important in the estimate.

$$D^m(g\circ g^n)=\sum_{k_1,\dots,k_m}C_{k_1,\dots,k_m}D^kg\,[(D^1g^n)^{\otimes k_1}\otimes\cdots\otimes(D^mg^n)^{\otimes k_m}], \tag{5.13}$$

where $k=k_1+\cdots+k_m$ and the sum is taken over all $k_1,\dots,k_m$ such that $k_1+2k_2+\cdots+mk_m=m$. We note that $k_m=0$ unless $k_m=1=k$ and we can separate the corresponding term as

$$D^m(g\circ g^n)=Dg\,[D^mg^n]+\sum_{k_1,\dots,k_{m-1}}C_{k_1,\dots,k_{m-1}}D^kg\,[(D^1g^n)^{\otimes k_1}\otimes\cdots\otimes(D^{m-1}g^n)^{\otimes k_{m-1}}].$$

We need to show $\|D^m(g\circ g^n)\| \le c^{m-1}(\sigma+\delta)^{n+1}$ provided that $\|D^m g^n\| \le c^{m-1}(\sigma+\delta)^n$, which yields

$$\|Dg\,[D^m g^n]\| \le \|Dg\|\cdot\|D^m g^n\| \le \sigma\, c^{m-1}(\sigma+\delta)^n.$$

Hence it suffices to estimate the norm of the sum from above by the difference

$$c^{m-1}(\sigma+\delta)^{n+1} - \sigma\, c^{m-1}(\sigma+\delta)^n = \delta\, c^{m-1}(\sigma+\delta)^n. \tag{5.14}$$

We estimate the norms of the terms using inductive assumptions as

$$\|D^k g\| \prod_{j=1}^{m-1} \|D^j g^n\|^{k_j} \;\le\; \|g|_{\mathcal W}\|_{C^k} \prod_{j=1}^{m-1} [c^{j-1}(\sigma+\delta)^n]^{k_j} \;=\; \|g|_{\mathcal W}\|_{C^k}\, c^{m-k}(\sigma+\delta)^{nk}.$$

All terms in the sum have $k>1$ and hence each can be estimated as

$$\|D^k g\,[(D^1 g^n)^{\otimes k_1}\otimes\cdots\otimes(D^{m-1}g^n)^{\otimes k_{m-1}}]\| \;\le\; \|g|_{\mathcal W}\|_{C^k}\, c^{m-2}(\sigma+\delta)^n.$$

for $\sigma+\delta<1$. The ratio of this to (5.14) is $\|g|_{\mathcal W}\|_{C^k}\,(c\delta)^{-1}$. Hence the inductive step will hold if we choose $c > N(m)\,\|g|_{\mathcal W}\|_{C^m}\,\delta^{-1} > 1$, where $N(m)$ is the sum of the coefficients $C_{k_1,\dots,k_{m-1}}$. We can choose the same c for all $m\le m_0$. Thus the first estimate in the lemma holds with $C=c^{m-1}$ for any $m\le m_0$.

To prove the second estimate in (5.11) we apply Faà di Bruno's formula to $D^m(\phi\circ g^n)$. Each term can be estimates as

$$\|D^k_{\mathcal W} g\,[(D^1_{\mathcal W} g^n)^{\otimes k_1}\otimes\cdots\otimes(D^{m-1}_{\mathcal W}g^n)^{\otimes k_m}]\| \;\le\; \|\phi|_{\mathcal W}\|_{C^k} \prod_{j=1}^{m} \|D^j_{\mathcal W} g^n\|^{k_j} \le$$

$$\|\phi|_{\mathcal W}\|_{C^k} \prod_{j=1}^{m} [c^{j-1}(\sigma+\delta)^n]^{k_j} \le\; \|\phi|_{\mathcal W}\|_{C^m}\, c^{m-k}(\sigma+\delta)^{nk} \le\; \|\phi|_{\mathcal W}\|_{C^m}\, C(\sigma+\delta)^n$$

since $\sigma+\delta<1$ and $c^{m-k}\le c^{m-1}=C$. The estimate for the sum follows by adjusting C.

(ii) The proof of the second part is similar so we just indicate the changes. We look for the inductive estimate of the form $\|D^m g^n\|\le c^{m-1}\sigma^{mn}$, with the base $m=1$ given by the assumption. Writing $D^m(g\circ g^n) = Dg\,[D^m g^n] + \sum\dots$ we need to estimate $\|\sum\dots\|$ from above by the gap similar to (5.14):

$$c^{m-1}\sigma^{m(n+1)} - \sigma c^{m-1}\sigma^{mn} \;=\; c^{m-1}\sigma^{mn}(\sigma^m-\sigma) \;\ge\; c^{m-1}\sigma^{mn}(\sigma^2-\sigma) \quad \text{for } m\ge 2.$$

The terms in $\|\sum\dots\|$ have $k>1$ and can be estimated as before

$$\|D^k g\,[(D^1 g^n)^{\otimes k_1}\otimes\cdots\otimes(D^{m-1}g^n)^{\otimes k_{m-1}}]\| \le \|g|_{\mathcal W}\|_{C^k}\, c^{m-k}\sigma^{nk} \le \|g|_{\mathcal W}\|_{C^k}\, c^{m-2}\sigma^{mn}$$

as $2\le k\le m$ and $\sigma>1$. Hence we can again choose c large enough to obtain the estimate $\|D^m_{\mathcal W} g^n\|\le C\sigma^{mn}$ with $C=c^{m-1}$. To prove the second inequality in (5.12) we estimate each term $D^m(\phi\circ g^n)$ similarly to the above with $\sigma>1$

$$\|D^k_{\mathcal W} g\,[(D^1_{\mathcal W} g^n)^{\otimes k_1}\otimes\cdots\otimes(D^{m-1}_{\mathcal W}g^n)^{\otimes k_m}]\| \;\le\; \|\phi|_{\mathcal W}\|_{C^k} \prod_{j=1}^{m} \|D^j_{\mathcal W} g^n\|^{k_j} \le$$

$$\|\phi|_{\mathcal W}\|_{C^k} \prod_{j=1}^{m} [c^{j-1}\sigma^{mn}]^{k_j} \le\; \|\phi|_{\mathcal W}\|_{C^m}\, c^{m-k}\sigma^{mnk} \;\le\; \|\phi|_{\mathcal W}\|_{C^m}\, C\sigma^{m^2 n}.$$

$\square$

Lemma 5.3. *Let $\mathcal{W}$ and $\mathcal{W}'$ be foliations of $\mathbb{T}^d$ invariant under a C^∞ diffeomorphism g with uniformly C^∞ leaves such that $\mathcal{W}'$ is a C^∞ foliation of each leaf of $\mathcal{W}$. Suppose that $\|Dg|_{\mathcal{W}}\| \le \sigma < 1$ and $\|Dg|_{\mathcal{W}'}\| \le \lambda$. Then for each m and $\delta > 0$ there exists $K = K(\delta, m, \|g|_{\mathcal{W}}\|_{C^{m+1}})$ such that for all $n \in \mathbb{N}$,*

$$\begin{aligned} &\|D^m_{\mathcal{W}}(D_{\mathcal{W}'}g^n)\| \le K(\lambda+\delta)^n \quad \text{and} \\ &\|D^m_{\mathcal{W}} D_{\mathcal{W}'}(\phi \circ g^n)\| \le K(\lambda+\delta)^n \, \|\phi|_{\mathcal{W}}\|_{C^{m+1}} \quad \text{for any } \phi \in C^\infty(\mathbb{T}^d). \end{aligned} \tag{5.15}$$

Proof. We again abbreviate $D_{\mathcal{W}}$ to D in this proof. We also denote $G_x = Dg|_{\mathcal{W}'(x)}$ and $G_x^n = Dg^n|_{\mathcal{W}'(x)}$. We will show inductively that for some c and all $m \le m_0$ we have

$$\|D^m G_x^n\| \le c^m(\lambda+\delta)^n \quad \text{for all } n \in \mathbb{N}.$$

In the base case $m = 0$ we have $\|G_x^n\| \le \lambda^n$ for all $n \in \mathbb{N}$ by the assumption. Suppose the estimate holds for the derivatives of orders up to $m-1$.

Now we show that $\|D^m G_x^n\| \le c^m(\lambda+\delta)^n$ for all $n \in \mathbb{N}$ by induction on n. The base case $n = 1$ holds if c is large. For the inductive step we differentiate $G_x^{n+1} = G_{g^n x} \circ G_x^n$ as the product of two linear maps which depend on x:

$$D^m(G_{g^n x} \cdot G_x^n) = \sum_{k=0}^{m} \binom{m}{k} D^k(G_{g^n x}) \cdot D^{m-k}(G_x^n) = \tag{5.16}$$

$$= G_{g^n x} \cdot D^m(G_x^n) + \sum_{k=1}^{m} \binom{m}{k} D^k(G_{g^n x}) \cdot D^{m-k}(G_x^n). \tag{5.17}$$

We want to show that $\|D^m(G_{g^n x} \cdot G_x^n)\| \le c^m(\lambda+\delta)^{n+1}$ provided that $\|D^m(G_x^n)\| \le c^m(\lambda+\delta)^n$, which yields

$$\|G_{g^n x} \cdot D^m(G_x^n)\| \le \|G_{g^n x}\| \|D^m(G_x^n)\| \le \lambda\, c^m(\lambda+\delta)^n.$$

Hence it suffices to estimate the norm of the sum on the right from above by the difference

$$c^m(\lambda+\delta)^{n+1} - \lambda\, c^m(\lambda+\delta)^n = \delta\, c^m(\lambda+\delta)^n \tag{5.18}$$

Now we estimate the norms of the terms in the sum. By the inductive assumption we get

$$\|D^{m-k}(G_x^n)\| \le c^{m-k}(\lambda+\delta)^n.$$

To estimate $D^k(G_{g^n x}) = D^k[(G \circ g^n)(x)]$ we use Lemma 5.2(i) with $\phi = G$:

$$\|D^k(G_{g^n x})\| \le C(\sigma+\delta)^n \|G|_{\mathcal{W}}\|_{C^m} \le C(\sigma+\delta)^n \|g|_{\mathcal{W}}\|_{C^{m+1}} \le C\|g|_{\mathcal{W}}\|_{C^{m+1}}.$$

for $\sigma + \delta < 1$. Hence terms in the sum with $k \ge 1$ can be estimated as

$$\|D^k(G_{g^n x}) \cdot D^{m-k}(G_x^n)\| \le Cg|_{\mathcal{W}}\|_{C^{m+1}} \cdot c^{m-k}(\lambda+\delta)^n \le Cc^{m-1}(\lambda+\delta)^n \|g|_{\mathcal{W}}\|_{C^{m+1}}.$$

The ratio of this to (5.18) is $C\|g|_{\mathcal{W}}\|_{C^{m+1}}(c\delta)^{-1}$. Hence the inductive step will hold if we choose $c > 2^m C\|g|_{\mathcal{W}}\|_{C^{m+1}}\delta^{-1}$, where 2^m is the sum of the binomial coefficients. Thus we can choose the same c for all $m \le m_0$ and the lemma holds with $K = c^m$ for any $m \le m_0$.

The second estimate in the lemma follows from the first one by applying Faà di Bruno's formula to $\phi \circ g^n$ and adjusting K in the same way as in Lemma 5.2. $\square$

6. Proof of Theorem 1.3

In this section we prove Theorem 1.3 by describing the adjustments we need to make in the proof of Theorem 1.1.

Since L is partially hyperbolic we have the L-invariant partially hyperbolic splitting

$$\mathbb{R}^d = E^s \oplus E^c \oplus E^u,$$

where E^c is the sum of all generalized eigenspaces of L corresponding to eigenvalues of modulus 1. Since f is $C^{1+\alpha}$ conjugate to L, it is also partially hyperbolic and it preserves the corresponding splitting

$$T\mathbb{T}^d = \mathcal{E}^s \oplus \mathcal{E}^c \oplus \mathcal{E}^u$$

into α-Hölder sub-bundles $\mathcal{E}^* = DH^{-1}(E^*)$ for $* = s, c, u$, called stable, center, and unstable respectively. Denoting the corresponding foliations for L and f by W^s, W^c, W^u and $\mathcal{W}^s$, $\mathcal{W}^c$, and $\mathcal{W}^u$ respectively, we have $H(\mathcal{W}^*) = W^*$ for $* = s, c, u$.

In general, the center foliation of a partially hyperbolic system may fail to be absolutely continuous, and the regularity of individual leaves may be lower than that of f, depending on the rate of expansion/contraction in $\mathcal{E}^c$. However, in our case $C^{1+\alpha}$ regularity of H means that foliations $\mathcal{W}^s$, $\mathcal{W}^c$, and $\mathcal{W}^u$ are $C^{1+\alpha}$, and hence all three are absolutely continuous with the conditional measures on the leaves given by the restriction of the volume form to their tangent spaces. In addition, the growth of $Df^n|_{\mathcal{E}^c}$ as $n \to \pm\infty$ is the same as for $L^n|_{E^c}$, that is, at most polynomial. This implies that f is so called *strongly r-bunched* for any r and hence $\mathcal{W}^c$ has uniformly C^∞ leaves [27]. Existence of H also implies that $\mathcal{E}^c \oplus \mathcal{E}^s$, $\mathcal{E}^c \oplus \mathcal{E}^u$, and $\mathcal{E}^s \oplus \mathcal{E}^u$ are tangent to $C^{1+\alpha}$ foliations $\mathcal{W}^{cs}$, $\mathcal{W}^{cu}$, and $\mathcal{W}^{su}$ respectively. It follows that these foliations also have uniformly C^∞ leaves, see for example [22, Lemma 4.1].

As in the proof of Theorem 1.1, we can assume that H is in the homotopy class of the identity satisfying $H(0) = 0$ and f is in the homotopy class of L satisfying $f(0) = 0$.

Using the splitting $\mathbb{R}^d = E^s \oplus E^c \oplus E^u$ for L we define the projections h^* and R_* for $* = s, u, c$, of $h = H - \mathrm{Id}$ and $R = f - L$ respectively. Similarly to the hyperbolic case, projecting the second equation in (2.3) to E^* we obtain for $* = s, u, c$

$$h^* = L_*^{-1}(h_* \circ f) + L_*^{-1} R_*, \quad \text{where } L_* = L|_{E^*}. \tag{6.1}$$

In particular, h^u is given by (2.6), and similarly $h^s = -\sum_{k=1}^{\infty} L_s^{k-1}(R_s \circ f^{-k})$. Moreover, the following analog of Lemma 2.1 holds.

Lemma 6.1. *The unstable component h^u is uniformly C^∞ along $\mathcal{W}^{cs}$. The stable component h^s is uniformly C^∞ along $\mathcal{W}^{cu}$. The center component h^c is uniformly C^∞ along $\mathcal{W}^{su}$.*

Proof. The proof is the same since $H^u = \mathrm{Id}_u + h^u$ is locally constant along the leaves of the foliation $\mathcal{W}^{sc}$, which are uniformly C^∞. Similarly, $H^c = \mathrm{Id}_c + h^c$ is locally constant along the leaves of $\mathcal{W}^{su}$, which are uniformly C^∞. □

Now we explain why h^u and h^s are C^∞ on $\mathbb{T}^d$. For h^u we use Lemma 6.1 in place of Lemma 2.1 and modify the charts Γ_i by including the center component. Then we can apply the proof of Theorem 2.2 without change, as the arguments work within the unstable foliation. Thus we obtain that h^u, and similarly h^s, are C^∞ on $\mathbb{T}^d$, and hence so are H^u and H^s.

It remains to show that h^c is C^∞ on $\mathbb{T}^d$. We give a proof by modifying our arguments for h^u. By Lemma 2.1 we already have that h^c is uniformly C^∞ along $\mathcal{W}^{su}$. Using the next proposition, we obtain global smoothness of h^c on $\mathbb{T}^d$ from [6, Theorem 3]. The required properties of $\mathcal{W}^c$ were givend above.

Proposition 6.2. *$D^m_{\mathcal{W}^c} h^c \in L^2(\mathbb{T}^d)$ for every every multi-index m.*

Proof. The proof is a significantly simplified version of the proof of Proposition 3.2. Since we study the derivatives of h^c along $\mathcal{W}^c$, so we do not need any further splitting of E^c and we do not need to separate derivative ∂_{x_k} as in that proposition. In particular we do not need Proposition 3.3 to remove ∂_{x_k}.

We will use the foliation chart Γ_c obtained as in Proposition 3.1 by smoothing h^c component of H similarly to (3.2) and (3.3)

$$\Gamma_c = \mathrm{Id} + \tilde{h}_\varepsilon, \quad \text{ where } \tilde{h}_\varepsilon = \big(h^s, s_\varepsilon(h^c), h^u\big).$$

The diffeomorphism Γ_c is $C^{1+\alpha}$ on $\mathbb{T}^d$, uniformly C^∞ along $\mathcal{W}^c$, and satisfies $\Gamma_c(\mathcal{W}^c) = W^c$.

The reason why the neutral case is simpler is that the equation (6.1) for h^c is already "neutral" since $L^n|_{E^c}$ has at most polynomial growth. Since this growth is dominated by the exponential mixing one can easily see that h^c itself, unlike h^u, can be written as a series $h^c = -\sum_{k=1}^\infty L_c^{k-1}(R_c \circ f^{-k})$ in distributional sense. More formally, we obtain the following series representation for $D^m_{\mathcal{W}^c}$ in the same way as (5.2) replacing ρ_i with $\rho_c = 1$

$$\big\langle h^c \circ \Gamma_c^{-1},\, D^m_{E^c}\psi \big\rangle = \sum_{k=1}^{\infty} L_c^{k-1} \big\langle R_c \circ f^{-k} \circ \Gamma_c^{-1},\, D^m_{E^c}\psi \big\rangle. \tag{6.2}$$

Then we use the same estimates as in the proof of (5.9) to obtain the analog of (5.10)

$$\| \big\langle D^m_{E^c}(h^c \circ \Gamma_c^{-1}),\, \psi \big\rangle \| = \| \sum_{k=1}^{\infty} L_c^{k-1} \big\langle R_c \circ f^{-k} \circ \Gamma_c^{-1},\, D^m_{E^c}\psi \big\rangle \| \le C_m \|\psi\|_{\mathcal{H}^\beta}. \tag{6.3}$$

The only differences are the absence of ∂_{x_k} term, replacing ρ_i with $\rho_c = 1$, and estimating norms of higher order derivatives of f using Lemma 5.2(ii). Specifically, we use the second part of (5.12) with $g = f^{-1}$, $\phi = R_c$, $\mathcal{W} = \mathcal{W}^c$, and $\sigma = 1 + \delta$ with δ small enough.

Finally we apply Lemma 4.7 in the same way as in the proof of Proposition 6.2 to conclude that $D^m_{E^c}(h^c \circ \Gamma_c^{-1})$ is in $L^2(\mathbb{T}^d)$ for all m.

This completes the proof of Theorem 1.3. □

References

[1] D. Anosov. *Geodesic Flows on Closed Riemannian Manifolds with Negative Curvature.* Proceedings of Steklov Institute of Mathematics, 1967, 90.

[2] A.Benyi and T. Oh. *The Sobolev inequality on the torus revisited.* Publicationes Mathematicae Debrecen, vol. 83, no. 3, pp. 359-374

[3] D. Damjanovic and A. Katok. *Local rigidity of partially hyperbolic actions. I. KAM method and $\mathbb{Z}^k$ actions on the torus.* Annals of Mathematics, 172 (2010), 1805–1858.

[4] R. de la Llave. *Invariants for smooth conjugacy of hyperbolic dynamical systems II.* Comm. Math. Phys., 109 (1987), 368-378.

[5] R. de la Llave. *Smooth conjugacy and SRB measures for uniformly and non-uniformly hyperbolic systems.* Comm. Math. Phys., 150 (1992), 289-320.

[6] R. de la Llave. *Remarks on Sobolev regularity in Anosov systems.* Ergodic Theory Dynam. Systems, 21 (2001), 1138–1180.

[7] R. de la Llave. *Rigidity of higher-dimensional conformal Anosov systems.* Ergodic Theory Dynam. Systems, 22 (2002), no. 6, 1845–1870.

[8] R. de la Llave. *Further rigidity properties of conformal Anosov systems.* Ergodic Theory Dynam. Systems, 24 (2004), no. 5, 1425–1441.
[9] R. de la Llave, J. M. Marco, R. Moriyón. *Canonical perturbation theory of Anosov systems and regularity results for the Livsic cohomology equation.* Ann. of Math. (2) 123 (1986), no. 3, 537-611.
[10] J. DeWitt. *Local Lyapunov spectrum rigidity of nilmanifold automorphisms.* J. Modern Dynamics 17 (2021): 65-109.
[11] J. DeWitt. *Periodic data rigidity of Anosov automorphisms with Jordan blocks.* Preprint.
[12] J. DeWitt and A. Gogolev. *Dominated splitting from constant periodic data and global rigidity of Anosov automorphisms.* Geometric and Functional Analysis (2024).
[13] D. Fisher, B. Kalinin, R. Spatzier. *Global rigidity of higher rank Anosov actions on tori and nilmanifolds.* J. Amer. Math. Soc., 26 (2013), no. 1, 167-198.
[14] Franks, J. *Anosov diffeomorphisms on tori.* Transactions of the AMS, vol. 145 (1969), 117-124.
[15] A. Gogolev. *Smooth conjugacy of Anosov diffeomorphisms on higher dimensional tori.* J. Modern Dynamics, 2, no. 4, 645-700 (2008).
[16] A. Gogolev. *Bootstrap for local rigidity of Anosov automorphisms of the 3-torus.* Comm. Math. Phys., 352, no. 2, 439-455 (2017).
[17] A. Gogolev and M. Guysinsky. *C^1-differentiable conjugacy of Anosov diffeomorphisms on three dimensional torus.* Discrete Contin. Dyn. Syst. 22 (2008), no. 1-2, 183-200.
[18] A. Gogolev, B. Kalinin, V. Sadovskaya. *Local rigidity for Anosov automorphisms.* Math. Research Letters, 18 (2011), no. 05, 843-858.
[19] A. Gogolev, B. Kalinin, V. Sadovskaya. *Local rigidity of Lyapunov spectrum for toral automorphisms.* Israel J. Math., 238 (2020), 389-403.
[20] A. Gogolev, B. Kalinin, V. Sadovskaya. *Center foliation rigidity for partially hyperbolic toral diffeomorphisms.* Mathematische Annalen, 387 (2023), 1579-1602.
[21] A. Gorodnik and R. Spatzier. *Mixing properties of commuting nilmanifold automorphisms.* Acta Math., 215(1):127-159, 2015.
[22] B. Kalinin, V. Sadovskaya. *Global rigidity for totally nonsymplectic Anosov $\mathbb{Z}^k$ actions.* Geometry and Topology, vol. 10 (2006), 929-954.
[23] B. Kalinin, V. Sadovskaya. *On Anosov diffeomorphisms with asymptotically conformal periodic data.* Ergodic Theory Dynam. Systems, 29 (2009), 117-136.
[24] B. Kalinin, V. Sadovskaya, Z. Wang. *Smooth local rigidity for hyperbolic toral automorphisms.* Comm. AMS, Vol. 3 (2023), 290-328.
[25] D. Lind. *Dynamical properties of quasihyperbolic toral automorphisms.* Ergodic Theory Dynamical Systems 2 (1982), no. 1, 49-68.
[26] A. Manning. *Anosov diffeomorphisms on nilmanifolds.* Proc. Amer. Math. Soc. 38 (1973), 423–426.
[27] C. Pugh, M. Shub, A. Wilkinson. *Hölder foliations.* Duke Math. J. Volume 86, Number 3 (1997), 517-546.
[28] R. Saghin and J. Yang. *Lyapunov exponents and rigidity of Anosov automorphisms and skew products.* Advances in Math., Vol. 355 (2019).
[29] P. Walters. *Conjugacy properties of affine transformations of nilmanifolds.* Math. Systems Theory 4 (1970), 327-333.

Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA.
E-mail address: kalinin@psu.edu, sadovskaya@psu.edu

Department of Mathematics, Michigan State University, East Lansing, MI 48824,USA
E-mail address: wangzq@math.msu.edu