OPTIMAL ENHANCED DISSIPATION AND MIXING FOR A TIME-PERIODIC, LIPSCHITZ VELOCITY FIELD ON $\mathbb{T}^2$

TAREK M. ELGINDI, KYLE LISS, and JONATHAN C. MATTINGLY

Abstract

We consider the advection-diffusion equation on $\mathbb{T}^2$ *with a Lipschitz and time-periodic velocity field that alternates between two piecewise linear shear flows. We prove enhanced dissipation on the timescale* $|\log \nu|$*, where* ν *is the diffusivity parameter. This is the optimal decay rate as* $\nu \to 0$ *for uniformly-in-time Lipschitz velocity fields. We also establish exponential mixing for the* $\nu = 0$ *problem.*

Contents

1. Introduction

Understanding the dynamics of a diffusive scalar advected by an incompressible flow is a fundamental problem in fluid mechanics with relevance to both engineering and the natural sciences. It is often possible in applications to treat the scalar as a *passive tracer*, that is, to assume that it has no feedback on the fluid flow. The relevant model is then the advection-diffusion equation

$$\begin{cases} \partial_t f + u \cdot \nabla f = \nu \Delta f, \\ f|_{t=0} = f_0 \end{cases} \tag{1}$$

DUKE MATHEMATICAL JOURNAL
Vol. 174, No. 7, © 2025 DOI 10.1215/00127094-2024-0057
Received 22 December 2023. Revision received 3 September 2024.
2020 *Mathematics Subject Classification*. Primary 35Q49, 37A25, 60J60, 76F25; Secondary 37A50, 37D20, 47D07.

posed on a suitable spatial domain, which we will always take to be the 2-dimensional torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. Here, $0 \leq \nu \ll 1$ is a fixed diffusivity constant, $u : [0,\infty) \times \mathbb{T}^2 \to \mathbb{R}^2$ is a given divergence-free velocity field, and $f : [0,\infty) \times \mathbb{T}^2 \to \mathbb{R}$ is the scalar unknown. In applications, f may represent, for example, temperature or the concentration of a chemical in water. The primary purpose of this paper is to study the quantitative decay rate of solutions with $\nu > 0$ for a particular velocity field u, which we also show defines an optimally mixing dynamical system on $\mathbb{T}^2$.

We first recall that since u is divergence-free, the spatial average of any solution to (1) is conserved and in fact all L^2 solutions with $\nu > 0$ converge strongly to their spatial average as $t \to \infty$. Indeed, a simple energy estimate using just the incompressibility of u gives

$$\frac{1}{2}\frac{d}{dt}\left\| f(t) - \fint f_0 \right\|_{L^2}^2 = -\nu \left\| \nabla f(t) \right\|_{L^2}^2, \tag{2}$$

where $\fint f_0 = \frac{1}{\mathrm{Vol}(\mathbb{T}^2)} \int_{\mathbb{T}^2} f_0(z)\,\mathrm{d}z$ denotes the spatial average. Hence, the Poincaré inequality implies that

$$\left\| f(t) - \fint f_0 \right\|_{L^2} \leq e^{-\nu t} \left\| f_0 - \fint f_0 \right\|_{L^2}. \tag{3}$$

When $u \equiv 0$, reducing (1) to the diffusion/heat equation, each nonzero Fourier mode of the solution decays independently at a frequency-dependent rate and the ν^{-1} decay timescale of (3) is in general sharp. An intriguing feature of (1) is that when diffusion and advection are both present, the formation of small scales induced by stirring can cause solutions to dissipate energy much faster than the heat equation as $\nu \to 0$, greatly accelerating the convergence in (3). This phenomenon is often referred to as *enhanced dissipation*. Intuitively, advection results in a conservative transfer of energy to high frequencies, which the diffusion damps much more effectively. Enhanced dissipation can be defined precisely as solutions to (1) exhibiting a decay timescale that is $o(\nu^{-1})$ as $\nu \to 0$.

Definition 1.1
A divergence-free and time-dependent velocity field $u : [0,\infty) \times \mathbb{T}^2 \to \mathbb{R}^2$ is said to be *dissipation enhancing* if there exist a constant $C \geq 1$ and a rate function $\delta(\nu) : (0,\infty) \to (0,\infty)$ with

$$\lim_{\nu \to 0} \frac{\nu}{\delta(\nu)} = 0$$

so that for all $f_0 \in L^2$ the solution f of (1) satisfies

$$\left\| f(t) - \fint f_0 \right\|_{L^2} \leq C e^{-\delta(\nu) t} \left\| f_0 - \fint f_0 \right\|_{L^2}$$

for all $t \geq 0$ and ν sufficiently small.

While enhanced dissipation is well understood in certain cases, for example, shear flows (see, e.g., [2], [10], [30], [31], [45], [46], [68]), the effects of diffusion and advection naturally do not commute, and understanding the interplay between them in full generality is extremely challenging. In fact, even though it is straightforward to prove that the optimal rate function for a uniformly-in-time Lipschitz velocity field behaves as $\delta(\nu) \approx |\log \nu|^{-1}$ as $\nu \to 0$, the only flows we are aware of that achieve this rate are those constructed in [7] from generic solutions to various stochastically forced fluid models. The setting of smooth or at least Lipschitz flows is, however, of clear physical importance. Indeed, one is naturally interested in understanding the interaction between advection and diffusion for velocity fields given by stationary or quasiperiodic solutions to the Euler or Navier–Stokes equations, which are typically regular.

Our main goal in this paper is to introduce an explicit, deterministic, time-periodic velocity field that is uniformly-in-time Lipschitz and whose dynamics we can understand deeply enough to establish *sharp* dissipation enhancement estimates. We consider in particular alternating piecewise linear shear flows. For $\alpha > 0$, let $H_\alpha : \mathbb{T}^2 \to \mathbb{R}^2$ and $V_\alpha : \mathbb{T}^2 \to \mathbb{R}^2$ (with $\mathbb{T} \cong [0,1]$) denote the shear flows

$$H_\alpha(x,y) = \begin{pmatrix} -2\alpha|y-1/2| \\ 0 \end{pmatrix} \qquad \text{and} \qquad V_\alpha(x,y) = \begin{pmatrix} 0 \\ -2\alpha|x-1/2| \end{pmatrix}.$$

We then define the Lipschitz continuous (in space), divergence-free velocity field $u_\alpha : [0,\infty) \times \mathbb{T}^2 \to \mathbb{R}^2$ by

$$u_\alpha(t,x,y) = \begin{cases} V_\alpha(x,y) & t \in [0,1/2), \\ H_\alpha(x,y) & t \in [1/2,1) \end{cases}$$

for $t \in [0,1)$ and extended for $t \geq 1$ to be time-periodic with a period of 1. An important feature of u_α is that it generates a *uniformly hyperbolic* map on $\mathbb{T}^2$ with singularities for α large. We mention that variations on the velocity field u_α have been considered in the recent works [20], [47], and [48], which will be discussed more in Remark 1.3 and Section 1.1 below.

Our main result is stated as follows.

THEOREM 1

If α is a sufficiently large even integer, then the velocity field u_α is dissipation enhancing with the optimal rate function

$$\delta(\nu) = \frac{c}{|\log \nu|},$$

where $c > 0$ is a constant depending on α but not on ν.

The fact that the decay timescale in Theorem 1 is optimal with respect to scaling in ν for Lipschitz u can be proved easily with (2) and the fact that $\|\nabla f(t)\|_{L^2}$ can grow at most exponentially fast. It should be mentioned that the situation changes significantly when the velocity field is less regular. For instance, if u is only Hölder continuous with exponent strictly less than 1, then (1) no longer necessarily conserves the L^2 norm when $\nu = 0$ and it is possible to construct examples such that solutions of the $\nu > 0$ problem decay at a ν-independent rate (see [3], [19], [25], [38], [39]). When u is not Lipschitz but instead bounded just in $W^{1,p}$ for $p < \infty$, then it is expected that the optimal rate function is ν-dependent but vanishes slower than $|\log \nu|^{-1}$ as $\nu \to 0$ (see [16]).

The estimate in (3) relies only on the dissipation produced by the Laplacian and ignores the effects of the transport term. One must leverage the mixing produced from the advection by u to obtain the optimal enhanced dissipation in Theorem 1. To this end, we study the large scale mixing properties of dynamics which hold for $\nu = 0$. Mixing refers, roughly speaking, to chaotic stretching and folding. An initially localized concentration of scalar tends to filament and spread throughout the entire domain, resulting in the conservative transfer of energy to higher Fourier modes. Precisely, a divergence-free and time-dependent velocity field $u : [0,\infty) \times \mathbb{T}^2 \to \mathbb{R}^2$ is said to be *mixing* if for every $f_0 \in L^2$ the solution to the advection equation

$$\begin{cases} \partial_t f + u \cdot \nabla f = 0, \\ f|_{t=0} = f_0 \end{cases} \tag{4}$$

converges weakly in L^2 as $t \to \infty$ to its spatial average.

As a corollary to the proof of Theorem 1 and an existing result in hyperbolic dynamics (see [35]), we are able to obtain optimal mixing estimates for the velocity field u_α. To state this result, we begin by recalling a quantitative notion of mixing rate. For a given divergence-free velocity field u, let $\phi_t : \mathbb{T}^2 \to \mathbb{T}^2$ denote the associated flow map, which solves

$$\begin{cases} \frac{d}{dt}\phi_t(z) = u(t, \phi_t(z)), \\ \phi_0(z) = z \in \mathbb{T}^2. \end{cases} \tag{5}$$

For any Lipschitz and divergence-free u, ϕ_t is well defined and a (Lebesgue) measure-preserving homeomorphism for each $t \geq 0$. The solution to the pure advection equation (4) is then given by

$$f(t) = f_0 \circ \phi_t^{-1}. \tag{6}$$

From (6), one can naturally quantify mixing in terms of correlation decay of sufficiently regular observables for the dynamical system on $\mathbb{T}^2$ generated by ϕ_t. We are interested here in *exponential mixing*, as this is easily seen to be the optimal rate for uniformly-in-time Lipschitz velocity fields.

Definition 1.2
A Lipschitz and divergence-free velocity field u is said to be an *exponential mixer* if there exist $C, c > 0$ so that

$$\left| ⨍ (f \circ \phi_t^{-1}) g - ⨍ f ⨍ g \right| \leq C \exp(-ct) \|f\|_{C^1} \|g\|_{C^1}$$

for any $f, g \in C^1(\mathbb{T}^2)$. This rate is optimal for uniformly-in-time Lipschitz u.

Remark 1.1
We mention that while there are many simple examples of exponentially mixing *maps* on $\mathbb{T}^2$ (e.g., Arnold's cat map and the baker's map), such maps cannot be realized by Lipschitz flows. This is true in the case of the baker's map because it is actually discontinuous on $\mathbb{T}^2$ (though it is known that the folded baker's map can be realized by a Hölder continuous velocity field; see [40]). The problem considered here of constructing exponentially mixing *flows* that are smooth or at least Lipschitz is fundamentally different and much more challenging than the corresponding problem for maps.

Remark 1.2
Exponential mixing in the sense of Definition 1.2 is equivalent to the existence of constants $C, c > 0$ such that for any $N \in \mathbb{N}$ and squares $R, Q \subseteq \mathbb{T}^2$ of side length 2^{-N}, there holds

$$\left| m\big(\phi_n^{-1}(R) \cap Q\big) - m(R) m(Q) \right| \leq C e^{-cn} \tag{7}$$

for all $n \geq CN$, where m denotes Lebesgue measure on $\mathbb{T}^2$. In other words, exponential mixing means that any square at a given length scale 2^{-N} that is pushed through the dynamics becomes roughly evenly distributed over $\mathbb{T}^2$ at the spatial resolution 2^{-N} after time approximately N. This is the main perspective on mixing and its relationship to dissipation enhancement that we will utilize in this paper.

We can now state our mixing result for u_α.

THEOREM 2
If α is a sufficiently large even integer, then u_α is exponentially mixing in the sense of Definition 1.2.

Remark 1.3
The restriction that α be an even integer is not important for the proof of either Theorem 1 or Theorem 2. It is merely done for convenience so that the time-1 flow map of u_α can be expressed in a simple way as a piecewise toral automorphism (see the beginning of Section 3.1). On the other hand, the assumption that α be sufficiently large is necessary, at least for Theorem 2. Indeed, it was observed in [20] that the mixing is at best algebraic in time in the cases $\alpha = 1/2, 1$ due to the presence of certain fixed structures. For instance, it was noted that when $\alpha = 1$ there exist four line segments that are each fixed by the time-2 flow map of u_α, and moreover that the gradient of the time-2 flow map locally near any one of these invariant sets is a Jordan block with eigenvalue 1. This rules out exponential mixing and implies in particular that correlations can decay with a rate no better than $\mathcal{O}(1/t)$. It was also observed numerically in [20] that when $\alpha = 1$ the mixing appears to be faster away from the invariant set, which suggests that the mixing rate in this case should be exactly $\mathcal{O}(1/t)$. The optimal $\mathcal{O}(1/t)$ mixing for $\alpha = 1$ was recently proved in [48]. The fact that the mixing rate for u_α is only algebraic when $\alpha = 1/2, 1$ suggests also that in these cases it can be dissipation enhancing with a rate function that is at best polynomial in ν, but such an implication is not strictly true in general.

Remark 1.4
For both of our results, the velocity field can be made smooth in time with the proof unchanged by replacing H_α and V_α in the definition of u_α with $\varphi(t)H_1$ and $\varphi(1/2 - t)V_1$ for any smooth function $\varphi : [0, 1/2] \to [0, \infty)$ with compact support in $(0, 1/2)$ satisfying $\int_0^{1/2} \varphi(t)\,\mathrm{d}t = \alpha$.

We emphasize here that while the properties of the $\nu = 0$ dynamics are fundamental to obtaining enhanced dissipation on the $|\log \nu|$ timescale, exponential mixing is not known to imply sharp dissipation enhancement for the $\nu > 0$ problem in general. There is a general result (see [31]) that one can take $\delta(\nu) = C|\log \nu|^{-2}$ for *any* Lipschitz continuous exponentially mixing flow. This is not quite optimal because of the 2 in the exponent of the logarithm, which seems very hard to remove in general because of difficulties with approximating the $\nu > 0$ dynamics by the $\nu = 0$ problem. Trying to prove in complete generality that exponential mixing of the $\nu = 0$ dynamics implies optimal enhanced dissipation appears to be quite challenging and perhaps unnecessary in practice. Instead, it seems more practical to show that suitable mixing properties hold for sufficiently small $\nu > 0$, and this is indeed the viewpoint taken in all of the known examples of regular velocity fields that induce sharp dissipation enhancement (see [7], [28], [66]). This is also the general philosophy taken here. In our proof of Theorem 1 we consider the $\nu > 0$ dynamics directly from the viewpoint of Markov processes and the stochastic representation of (1) and do not explicitly use exponential mixing of the $\nu = 0$ problem. We rely only on proving that a fixed fraction of mass of any initial scalar is mixed to the diffusive $\sqrt{\nu}$ scale after time $t \approx |\log \nu|$ and moreover that this mixing crucially persists in some sense for the $\nu > 0$ dynamics. This is described by Lemma 2.1, which we are able to prove in a direct and self-contained way.

1.1. Previous results

The study of mixing and dissipation enhancing flows has seen an explosion of work in the last few decades. In this section, we will outline the existing works and then discuss how the present contribution fits into the picture.

1.1.1. PDE literature

Early work on enhanced dissipation was done by Kelvin [54] and Kolmogorov [56]. An important foundational work is the paper [26] in which the authors establish, as a corollary of a more general result, necessary and sufficient conditions for *autonomous* flows to be dissipation enhancing. Extensions to the case of time-periodic flows were established in [55] and [73]. Building off of [26], the authors of [29] and [43] established quantitative enhanced dissipation rates based on quantitative mixing rates. A restricted form of enhanced dissipation has also been established in a number of specific settings like for shear flows (see, e.g., [2], [10], [24], [45], [46], [68]) in various domains and using a variety of proofs as well as circular flows (see [30], [44]), cellular flows (see [51], [52]), and general flows with autonomous Hamiltonian (see [17], [67]). Enhanced dissipation for certain shear flows has also been established in various nonlinear settings (see, e.g., [11], [12], [32], [61], [69], [70]).

A number of works on mixing in the PDE literature have also appeared recently in relation to Bressan's mixing conjecture in [14]. There, a central question is whether *low regularity* velocity fields can mix faster than exponentially. This is still open in general (in particular in the setting of Bressan's conjecture), though significant progress has been made both on the positive side (see [27], [33], [49]) and the negative side (see [1], [40], [74]). For additional works focused on optimal mixing subject to certain physical constraints on the flow see, for example, [58] and [60].

1.1.2. Chaotic dynamics literature

The invariant manifold structure and statistical properties of hyperbolic maps have been studied extensively in the dynamics literature (see, e.g., [5], [18], [21]–[23], [35], [53], [59], [71]). Exponential decay of correlations is known in various situations, including certain examples of billiards (see [5], [22], [71]) and some piecewise linear maps (see [21], [47]). Mixing has also seen new developments in the context of *random* dynamical systems. The first examples of spatially smooth exponentially mixing flows on the torus were constructed in [8] and [9] and from solutions to various stochastic fluid models with sufficiently nondegenerate noise. These flows are also known to give enhanced dissipation on the $|\log \nu|$ timescale (see [7]). Note that the velocity fields studied in [7]–[9] are smooth but are not uniformly bounded in time. Motivated by questions in passive scalar turbulence, a simple candidate for an exponentially mixing flow which is both bounded and spatially smooth was introduced and studied numerically by Pierrehumbert in the classical work [65]. Pierrehumbert's model, which consists of alternating sine shears with random phases, was proved to be an exponential mixer just recently (see [13]). The general random dynamics framework developed in [13] has since been applied to alternating sine shears with different randomization procedures (see [27]).

1.1.3. Comparison with existing works

In two spatial dimensions, there cannot exist a dissipation enhancing flow with autonomous velocity field on the torus (see [26]). This is because, under suitable conditions, divergence-free vector fields on $\mathbb{T}^2$ admit a periodic Hamiltonian, which would be preserved under the flow. Because of this, the velocity field must be sufficiently complicated for the flow to be dissipation enhancing, especially so if we want the flow to enjoy sharp decay rates. A natural step after looking for a flow with autonomous velocity field is to look for one with a time-periodic velocity field. Here, the restrictions imposed by the Hamiltonian structure are relaxed, though not entirely removed. It follows from our construction that the simple (time-periodic) alternation of two autonomous and Lipschitz continuous velocity fields can lead to optimal dissipation enhancement (and mixing). As remarked earlier, the only previous example of dissipation enhancement on the $|\log \nu|$ timescale that we know of are the flows constructed as solutions to stochastic PDEs (see [7]), which while spatially regular are neither time-periodic nor uniformly bounded. It appears that Theorem 1 provides the first example in the setting of time-periodic flows and also in the setting of uniformly Lipschitz flows. An instance of convergence on the $|\log \nu|$ timescale in a discrete-time setting has very recently appeared for one-sided Bernoulli shifts on the torus with axis-aligned cylinder sets perturbed by small noise in [50] (here, ν is the parameter that controls the strength of the noise, which plays a role analogous to the Laplacian in (1)). Such maps, however, cannot be realized as the time-1 flow map of a velocity field on $\mathbb{T}^2$. Additionally, the analysis in [50] seems not to apply to invertible dynamics. We mention that the viewpoint of analyzing enhanced dissipation in terms of the convergence rate for an associated Markov process, which we utilize in our work, is present also in [50] as well as [52].

In regards to exponential mixing, our construction can be contrasted to the examples in [13] and [27] where the velocity has *analytic* regularity in space, though the velocity is not time-periodic and is based on *random* switching in time, while ours is deterministic and periodic. The construction can also be contrasted with the previous work [40], where less regular exponential mixers were constructed by finding a velocity field whose time-1 map is the baker's transformation. The velocity field considered in the present paper was analyzed numerically in [20] and observed to be exponentially mixing when α is not too small. Theorem 2 above establishes exponential mixing, at least if α is large enough. As mentioned earlier in Remark 1.3, the fact that exponential mixing *does not* occur for certain values of α small was first observed in [20], and the fact that when $\alpha = 1$ correlations decay like $\mathcal{O}(1/t)$ was proved in the recent paper [48]. Most similar to our setting is the recent work [47], in which the general theorems from [23] and [53] are used to obtain exponential mixing for a class of non-monotone, piecewise linear alternating shear flows without the dissipation ($\nu = 0$). The shears have fixed amplitude and are taken sufficiently asymmetric to ensure good mixing properties of the map, whereas in our example the shears are symmetric and uniform hyperbolicity is guaranteed by taking the amplitude large. We provide a short and self-contained proof of the mixing described by Lemma 2.1, which is particularly well suited for proving our principle result concerning the dissipation enhancement as $\nu \to 0$.

1.2. Discussion of the proof

The proof of the Theorem 1 is based primarily on considering the "pulsed" diffusion obtained by applying the dissipation discretely in time and a careful study of the chaotic properties of the dynamical system generated by u_α. We then deduce enhanced dissipation in continuous time by using an approximation argument that compares the solution to (1) with the pulsed diffusion.

It is important to not think of this scheme as treating the $\nu > 0$ dynamics as a perturbation of the $\nu = 0$ dynamics even though we are interested in the $\nu \to 0$ limit. The effect of the noise is critical to our results as it produces the small scale mixing. Equally important is the fact that the structures which produce the large scale mixing at $\nu = 0$ persist when the system is viewed correctly when $\nu > 0$. As we will study the pulsed diffusion using its probabilistic representation, it is best to think of the pulsed diffusion as a time splitting scheme for the stochastic differential equation (SDE) whose probability flow produces (1).

We now discuss the main steps in our approach, which hinges on a few basic ideas from probability, hyperbolic dynamics, and PDE.

1.2.1. Probabilistic setup for the pulsed diffusion

The dynamics given by (1) and (4) describe the spread of an initial particle distribution f_0 transported by the velocity u with and without diffusion. The dynamics of an individual particle starting at z without diffusion is given by $t \mapsto \phi_t(z)$, where ϕ_t denotes the flow map defined in (5). In the presence of diffusion, the dynamics of an individual particle is given by the SDE

$$\varphi_{t,W}(z) = z + \int_0^t u\big(s, \varphi_{s,W}(z)\big)\, ds + \sqrt{2\nu} W_t, \tag{8}$$

where W_t is a standard 2-dimensional Brownian motion $[0,\infty)$ with $W_0 = 0$. Then when $\nu > 0$, $f(t,z) = \mathbf{E}(f_0 \circ \varphi_{t,W}^{-1}(z))$ where $\mathbf{E}$ is the expectation over the Brownian paths.

Here and throughout the remainder of the paper, we write ϕ_t for the deterministic flow map associated with u_α, suppressing the dependence on α for simplicity. The starting point for our analysis is to consider the pulsed diffusion

$$\Phi_\nu g = e^{\nu\Delta}(g \circ T), \tag{9}$$

where $T = \phi_1^{-1}$ and $e^{\nu\Delta}$ denotes convolution with the periodized heat kernel. Then $\Phi_\nu f_0$ is an approximation for $\mathbf{E}(f_0 \circ \varphi_{t,W}^{-1}(z))$. To parallel the pathwise/particle level description given in (5) and (8) for the pulsed diffusion, we introduce the "kicked" version of T by $T_v(z) = T(z + \pi(v))$ for $v \in \mathbb{R}^2$, where $\pi : \mathbb{R}^2 \to \mathbb{T}^2$ denotes the projection $\pi(x,y) = (x \mod 1, y \mod 1)$. Observe that if ξ is a normal random variable on $\mathbb{R}^2$ with mean zero and identity covariance, then

$$\Phi_\nu g(z) = \mathbf{E}_\xi\big(g \circ T_{\sqrt{2\nu}\xi}(z)\big) := \int_{\mathbb{R}} g\big(T_{\sqrt{2\nu}v}(z)\big)\mathcal{N}(dv) \tag{10}$$

for $z \in \mathbb{T}^2$, where $\mathcal{N}$ denotes the standard unit normal distribution on $\mathbb{R}^2$.

To consider iterates of Φ_ν, we introduce the following probabilistic framework. Define the probability space $(\Omega, \mathcal{F}, \mathbf{P}) = (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mathcal{N})^{\mathbb{N}}$. Then, for $\xi = (\xi_1, \xi_2, \ldots) \in \Omega$, define the "randomly kicked" map

$$T_\xi^n = T_{\xi_n} \circ \cdots \circ T_{\xi_1}.$$

For notational convenience, we also set $T_{\nu,\xi}^n = T_{\sqrt{2\nu}\xi}^n$. Given any $z \in \mathbb{T}^2$, $\{T_{\nu,\xi}^n(z)\}_{n=0}^\infty$ defines a Markov chain on $\mathbb{T}^2$ with initial condition z. As before, the connection between $T_{\nu,\xi}^n$ and Φ_ν is given by

$$\Phi_\nu^n g(z) = \int_\Omega g\big(T_{\nu,\xi}^n(z)\big)\mathbf{P}(\mathrm{d}\xi). \tag{11}$$

1.2.2. Uniform hyperbolicity and the cone condition

A fundamental property of the $\nu = 0$ dynamics is the exponential separation of particle trajectories both forward and backward in time. Indeed, taking $t \in \mathbb{N}$ without loss of generality, it is not hard to see that since ∇V_α and ∇H_α are both piecewise constant, for every $z \in \mathbb{T}^2$ away from a measure zero set of finitely many lines we have

$$\nabla T^N(z) := \nabla \phi_N^{-1}(z) = \Pi_{i=1}^N A_{j_i}, \tag{12}$$

where $j_i \in \{1, 2, 3, 4\}$ and A_{j_i} is one of the four matrices

$$A_1 = \begin{pmatrix} 1+\alpha^2 & \alpha \\ \alpha & 1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1-\alpha^2 & \alpha \\ -\alpha & 1 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 1-\alpha^2 & -\alpha \\ \alpha & 1 \end{pmatrix}, \qquad A_4 = \begin{pmatrix} 1+\alpha^2 & -\alpha \\ -\alpha & 1 \end{pmatrix}.$$

Of course, the particular product depends on the point z. There are four possible matrices since the vertical and horizontal shears are each linear on two halves of the torus. A key property that these matrices share is that for α large they are all hyperbolic with "close" expanding and contracting directions. In particular, when $\alpha \gg 1$ the expanding direction of each is very close to $e_1 = (1, 0)$, while the contracting direction is very close to $e_2 = (0, 1)$. This can be made precise by establishing that the cones

$$C_u := \left\{(x, y) \in \mathbb{R}^2 : |y| \le \frac{|x|}{\alpha\delta_1}\right\},$$

$$C_s := \left\{(x, y) \in \mathbb{R}^2 : |x| \le \frac{|y|}{\alpha\delta_1}\right\}$$

are, respectively, forward and backward invariant under the action of any of the A_i. Here, $\delta_1 < 1$ is fixed and α is sufficiently large. In particular, it can be shown using basic linear algebra that for any $v_u \in C_u$ and $v_s \in C_s$ we have

$$\left|(\Pi_{i=1}^N A_{j_i}) v_u\right| \ge \exp(cN), \qquad \left|(\Pi_{i=1}^N A_{j_i})^{-1} v_s\right| \ge \exp(cN),$$

for *any* choice of the sequence $\{j_i\}_{i=1}^N$, with $c > 0$ a constant growing with α. Intuitively, the fact that each A_i is hyperbolic guarantees a minimal amount of expansion in the unstable direction under each iteration, while the cone condition ensures that the expanding and contracting directions do not mix to cancel out the long-time effects. The observations above imply that the discrete-time dynamical system on $\mathbb{T}^2$ defined by $T = \phi_1^{-1}$ is uniformly hyperbolic for α large. A critical fact in all that follows is that the representation given in (12) also holds for $\nabla T_{\nu,\xi}^N(z)$, as the additive shifts only rigidly translate the solutions and do not change the gradients.

1.2.3. *Complexity of intersections of approximate stable and unstable foliations*

The cone conditions and uniform hyperbolicity established above allow us to infer exponentially fast gradient growth as $t \to \infty$ for generic Lipschitz solutions to (4). This itself does not imply mixing or dissipation enhancement (at the sharp rate for the latter). For that, we need to take more advantage of the above structure to get geometric information on the evolution of areas. In view of (7), we would like to show that squares of side length 2^{-N} are stretched horizontally across the whole torus in time on the order of N under the *forward* iterates of T and similarly stretched vertically under the *backward* iterates. If this were the case, we would be able to argue that

$$m\big(T^{2n}(R) \cap Q\big) = m\big(T^{n}(R) \cap T^{-n}(Q)\big) \gtrsim m(R)m(Q),$$

for $n \approx N$ and any squares R and Q of side length 2^{-N}. Given the forward invariance of C_u and the fact that T maps any line segment to a union of line segments, it is natural to first study the expansion of short horizontal lines. The challenge here is that while a line segment W tangent to the unstable cone is expanded forwards in time, it is also repeatedly cut across the singularity curves of ∇T (see Figure 1 for a representative singularity set). Thus, there is a competition between expansion and cutting, making it not immediately clear the extent to which "long lines dominate" in $T^n(W)$. Nevertheless, using the cone conditions and the relatively simple structure of the singularity set of ∇T we are able to prove that horizontal lines are stretched across the whole torus at the correct rate. Growth bounds of this general type on unstable curves are well known to be crucial in the chaotic dynamics literature and are often referred to as "complexity estimates." We follow closely here the strategy of Baladi and Demers (see [4], [34]), with some modifications and simplifications that are available in our specific setting.

In order to make a somewhat precise statement, let us define $\mathcal{W}_u$ to be the collection of line segments on $\mathbb{T}^2$ that are tangent to some $v \in C_u$.

LEMMA 1.1 (Informal statement)
Let $W \in \mathcal{W}_u$. Then, if $n \gtrsim |\log|W||$, we have that $T^n(W)$ can be written as a union of line segments in $\mathcal{W}_u$ most of which have length greater than $5/4$. In particular, if α is sufficiently large, they stretch across $\mathbb{T}^2$.

A similar statement holds backwards in time for line segments tangent to the stable cone. The complexity lemma implies that horizontal segments forward in time *must* intersect vertical segments backward in time. From this, we can show that squares that are mapped forward in time must intersect squares that are mapped backwards in time with a significant area of intersection. To make this rigorous, we rely on an argument using the area formula from geometric measure theory.

As in the preceding section, all of the arguments made here largely carry over without modification to the randomly kicked version $T^n_{\nu,\xi}$, as rigid translations of the solutions do not change the estimates made.

1.2.4. Area formula argument

The area formula allows us to reduce the estimate of $m(T^n(R) \cap T^{-n}(Q))$ to tracking the intersection points between horizontal lines mapped through T^n and vertical lines mapped through T^{-n}. Define the Lipschitz continuous function $F_n : \mathbb{T}^2 \to \mathbb{T}^2$ by

$$F_n(z) = \big(\Pi_x \circ T^n(z), \Pi_y \circ T^{-n}(z)\big),$$

where Π_x and Π_y denote the usual projections $\Pi_x(x, y) = x$ and $\Pi_y(x, y) = y$. Since the unstable direction is approximately e_1 and the stable direction is approximately e_2, the first and second coordinates of $F_n(z)$ roughly give the forward and backward dynamics, respectively, in the unstable direction. More precisely, $F_n^{-1}(x, y)$ consists exactly of the intersection points between $T^n(\ell_y)$ and $T^{-n}(\ell_x)$, where ℓ_y denotes the horizontal line with second coordinate y that spans across the entire torus and ℓ_x similarly denotes the vertical line spanning the torus with first coordinate x. Using this fact, one can show by the area formula, recalled in Section 4 for the convenience of the reader, that

$$m\big(T^n(R) \cap T^{-n}(Q)\big) = \int_{\Pi_x(Q)} \int_{\Pi_y(R)} \Big(\sum_{z \in T^n(W_{y'}) \cap T^{-n}(V_{x'})} \frac{1}{J_{F_n}(z)} \Big) \,\mathrm{d}y' \,\mathrm{d}x',$$

where $W_{y'} = \Pi_x(R) \times \{y'\}$, $V_{x'} = \{x'\} \times \Pi_y(Q)$, and J_{F_n} denotes the Jacobian of F_n. Note that $W_{y'}$ and $V_{x'}$ are horizontal and vertical lines, respectively, of length 2^{-N}. To estimate the sum for a fixed (x', y'), we take $n \gtrsim N$ and decompose $T^n(W_{y'})$ and $T^{-n}(V_{x'})$ each into the union of "mostly long lines" as guaranteed by the complexity lemma. The precise formulation of the statement that most elements of $T^n(W_{y'})$ and $T^{-n}(V_{x'})$ span the entire torus ensures that $\#(T^n(W_{y'}) \cap T^{-n}(V_{x'}))$ balances the Jacobian factor, yielding the bound

$$\sum_{z \in T^n(W_{y'}) \cap T^{-n}(V_{x'})} \frac{1}{J_F(z)} \gtrsim |W_{y'}||V_{x'}| = 2^{-2N},$$

from which it follows immediately that

$$m\big(T^{2n}(R) \cap Q\big) = m\big(T^n(R) \cap T^{-n}(Q)\big) \gtrsim m(R)m(Q) \tag{13}$$

for $n \gtrsim N$.

Since R and Q are both squares of side length 2^{-N}, the bound (13) holding for $n \gtrsim N$ describes a kind of exponential mixing estimate. It extends to the randomly kicked maps in the form

$$m\big(T_\xi^{2n}(R) \cap Q\big) \gtrsim m(R)m(Q)$$

for any $\xi \in \Omega$ and is the content of Lemma 2.1, which we refer to as the *geometric mixing lemma*. The motivation for this name is that it follows as a corollary that the *geometric mixing scale* (see Definition 2.1) of any solution to (4) with mean-zero $f_0 \in C^1$ decays exponentially fast. The geometric mixing scale has been considered in various works (see [1], [40], [74]), largely in connection with Bressan's conjecture. Exponential decay of the geometric mixing scale does not imply exponential mixing in the sense of Definition 1.2 (see, e.g., discussions in [72]), but we observe that Lemma 2.1 can similarly be used to show that T^k is ergodic with respect to Lebesgue measure for every $k \in \mathbb{N}$, from which exponential decay of correlations follows by [35, Theorem 2.8]. We do not require this fact for the proof of Theorem 1, which relies only on the geometric mixing lemma for the kicked maps.

1.2.5. Enhanced dissipation

Using (11) and standard ideas from the ergodic theory of Markov processes, we show that the convergence $\Phi_\nu^n g \to \fint g$ on the $|\log \nu|$ timescale (in L^∞ for bounded g) follows provided that for every $\xi \in \Omega$ and rectangles R, Q with side length $\sqrt{\nu}$ there holds

$$m\big(T_\xi^{n}(R) \cap Q\big) \gtrsim m(R)m(Q)$$

for $n \approx |\log \nu|$. This is exactly the geometric mixing lemma applied with squares at the diffusive $\sqrt{\nu}$ length scale.

To extend the results for the pulsed diffusion to continuous time, we use an approximation scheme based on a contradiction argument. In particular, we show that *if the solution to (1)* decays slower than the expected rate, then the pulsed diffusion well-approximates the solution to (1). The dissipation for solutions to the discrete problem are then transferred to solutions of (1), which gives a contradiction. Compared with previous similar arguments in [29] and [43], here we are comparing two *dissipative* problems rather than a dissipative problem with a conservative one. This gives us more control and in particular allows us to remove the exponent 2 from the $|\log \nu|^2$ timescale often obtained in such arguments.

Remark 1.5

Tracking the complexity and geometry of the intersection of invariant manifolds is a central element of modern smooth ergodic theory and Pesin theory in particular (see [5], [6]). Here we use some of the elements of the theory replacing other elements with ideas and estimates from nondegenerate Markov process theory. In particular, we use the geometry of the intersections of stable and unstable manifolds to prove a kind of scale dependent, topological irreducibility and control the stretching and contraction of the transition density. In Theorem 1, we do not make use of any absolute continuity or smoothness of measures restricted to these manifolds in proving our ergodic theorems as one often does in smooth ergodic theory. Rather we leverage the smoothing properties of the noise to deduce the measure-theoretic irreducibility needed to prove our ergodic and mixing results at $\nu > 0$.

1.3. Outline of the article

In Section 2 we prove Theorems 1 and 2 assuming the geometric exponential mixing estimate described above. This includes the necessary convergence results for the pulsed diffusion and the approximation argument that uses them to conclude optimal dissipation enhancement for u_α. In Section 3 we prove the complexity lemma, and in Section 4 we conclude the proof of the geometric mixing lemma using the area formula argument.

2. Proof of Theorems 1 and 2

In this section, we first state precisely the geometric mixing lemma. Then, assuming it, we complete the proof of Theorems 1 and 2.

2.1. Geometric mixing lemma

For $N \in \mathbb{Z}_{\geq 0}$, we define a tiling of $\mathbb{T}^2$ with squares at scale 2^{-N} by

$$\mathcal{R}_N = \big\{\big[j2^{-N}, (j+1)2^{-N}\big] \times \big[k2^{-N}, (k+1)2^{-N}\big] : k, j \in \mathbb{Z}, 0 \leq k, j \leq 2^N - 1\big\}.$$

The geometric exponential mixing lemma says that

$$m\big(T_\xi^n(R) \cap Q\big) \gtrsim m(R)m(Q)$$

for any $R, Q \in \mathcal{R}_N$, $n \gtrsim N$, and $\xi \in (\mathbb{R}^2)^{\mathbb{N}} = \Omega$. In the statement below and all those that follow, we assume that α is an even integer without explicit mention of it.

LEMMA 2.1

There exist $C_0, \alpha_0 \geq 1$ and $\lambda \in (0,1)$ such that for $\alpha \geq \alpha_0$, $n \geq C_0 N$, $\xi \in \Omega$, and any $R, Q \in \mathcal{R}_N$ there holds

$$m\big(T_\xi^n(R) \cap Q\big) \geq \lambda m(R)m(Q).$$

The proof of Lemma 2.1 is the main technical step in the paper and the content of Sections 3 and 4.

Remark 2.1

It is also possible to use the ideas of Sections 3 and 4 to prove that the analogous upper bound $m(T_\xi^n(R) \cap Q) \leq Cm(R)m(Q)$ holds for a suitable constant C. The lower bound however is the fundamental estimate which describes mixing, and for the proof of Theorems 1 and 2 it is the only direction that we require.

Remark 2.2

It is possible to establish Lemma 2.1 (and often using a much simpler proof) with T_ξ^n replaced by randomly kicked versions of other maps, such as Arnold's cat map and the baker's map. For the cat map, the proof becomes a simple exercise. In the case of the baker's map, some work is required but the proof is still significantly simpler than in the case considered here. However, as discussed in Remark 1.1, these maps are not realized by Lipschitz flows. It is also worth noting that the argument in Section 2.4 that extends enhanced dissipation for the pulsed diffusion to continuous time relies on Lipschitz regularity, and hence is no longer valid in such settings.

2.2. *Probabilistic preliminaries*

We now introduce some standard probabilistic notation that will be useful in the study of the pulsed diffusion and also recall a basic convergence theorem from the ergodic theory of Markov processes.

Recall that for $\nu > 0$ we write $T_{\nu,\xi} = T_{\sqrt{2\nu}\xi}$ and that $T_{\nu,\xi}(z)$ defines a Markov chain for any initial condition $z \in \mathbb{T}^2$. The associated Markov transition kernel is defined by

$$\Phi_\nu(z, A) = \mathbf{P}\big(T_{\nu,\xi}(z) \in A\big),$$

where $A \in \mathscr{B}(\mathbb{T}^2)$, the Borel σ-algebra on $\mathbb{T}^2$. Recall that a Markov transition kernel on a measurable space (X, Σ) is a mapping $Q : X \times \Sigma \to [0, 1]$ such that $Q(x, \cdot)$ is a probability measure for every $x \in X$ and $x \mapsto Q(x, A)$ is measurable for every fixed $A \in \Sigma$. We similarly define $\Phi_\nu^n(z, A) = \mathbf{P}(T_{\nu,\xi}^n(z) \in A)$ and note that by the Markov property, for every $n \geq 2$ we have

$$\Phi_\nu^n(z, A) = \int_{\mathbb{T}^2} \Phi_\nu^{n-1}(z', A)\Phi_\nu(z, \mathrm{d}z'). \tag{14}$$

This can be seen as simply the semigroup property $\Phi_\nu^{n+m} = \Phi_\nu^n \Phi_\nu^m$ for the Markov semigroup Φ_ν defined in the introduction in (9) and (10) since $\Phi_\nu(z, A) = \Phi_\nu \mathbf{1}_A(z)$, where the indicator function $\mathbf{1}_A(z)$ is 1 if $z \in A$ and zero otherwise. It will also be convenient for us later on to define the shift operator $\theta : \Omega \to \Omega$, which acts on $\xi = (\xi_1, \xi_2, \dots) \in \Omega$ by

$$\theta\xi = (\xi_2, \xi_3, \dots).$$

To study the convergence of $\Phi_\nu^n g$ to $⨍ g$, we will make use of the following basic result from the ergodic theory of Markov processes (see [36], [37], [57], [62]).

Recall that if Q is a Markov transition kernel on a measurable space (X, Σ), then a probability measure μ on (X, Σ) is said to be *invariant* for Q if

$$\mu(A) = \int_X Q(x, A)\mu(\mathrm{d}x)$$

for every $A \in \Sigma$.

LEMMA 2.2 (Doeblin)
Let X be a Polish space, and let Σ be the Borel σ-algebra on X. Let Q be a Markov transition kernel on (X, Σ), and define Q^n as in (14). Suppose that there exists $\lambda \in (0, 1)$ and a Borel probability measure μ on (X, Σ) such that

$$Q(x, \cdot) \geq \lambda\mu(\cdot) \tag{15}$$

for every $x \in X$. Then there exists a unique invariant measure μ_* for Q such that for every bounded, Borel measurable function $f : X \to \mathbb{R}$ there holds

$$\left\| Q^n f - \int_X f(x)\mu_*(\mathrm{d}x) \right\|_\infty \leq 2(1 - \lambda)^n \|f\|_\infty,$$

where $\|f\|_\infty = \sup_{x \in X} |f(x)|$.

Since Lebesgue measure is invariant for Φ_ν, it follows from Lemma 2.2 that we can obtain an optimal convergence result for the pulsed diffusion by showing that $Q = \Phi_\nu^{n_0}$ satisfies (15) for some $n_0 \approx |\log(\nu)|$ and $\lambda \in (0, 1)$ that does not depend on ν.

2.3. *Convergence results for the pulsed diffusion*

In this section, we use Lemma 2.2 as described above to prove an optimal convergence estimate for the pulsed diffusion. The main result here is stated as follows.

THEOREM 3
For all α sufficiently large and $p \in (1, \infty]$, there exist constants $C, c > 0$ such that for any $\nu \in (0, 1/2]$, and mean-zero $f \in L^p(\mathbb{T}^2)$ there holds

$$\|\Phi_\nu^n f\|_{L^p} \leq Ce^{-cn|\log \nu|^{-1}} \|f\|_{L^p}. \tag{16}$$

The key step in the proof of Theorem 3 is to use Lemma 2.1 to obtain a uniform-in-ν lower bound on the density of $T^n_{\nu,\xi}(z_0)$ with respect to Lebesgue measure. For $n \geq 2$ we denote this density by $p_n(z_0, \cdot)$, and for $n = 1$ we write $p(z_0, \cdot)$.

LEMMA 2.3
Let α_0, C_0, and λ be as in the statement of Lemma 2.1. There exist constants $\delta_1, C_1 > 0$ that do not depend on ν so that for all $\alpha \geq \alpha_0$, $z, z_0 \in \mathbb{T}^2$, and $n \geq C_1|\log\nu|$ we have

$$p_n(z_0,z) \geq \delta_1. \tag{17}$$

Proof
It suffices to consider the case where $\nu \ll 1$. Let N be the first natural number such that $2^{-N} \leq \sqrt{\nu}$, and let $M \geq N$ be the first natural number such that $2^{-M} \leq \nu^2$. Suppose that $n \geq C_0 M + 2$, and note that this is satisfied provided that $n \geq C|\log\nu|$ for some constant C that does not depend on ν. We first claim that for ν sufficiently small we have

$$\inf_{R\in\mathcal{R}_N} \inf_{z_0\in\mathbb{T}^2} \frac{\Phi_\nu^{n-1}(z_0,R)}{m(R)} \geq \frac{\lambda}{2}, \tag{18}$$

where λ is as in Lemma 2.1. By the semigroup property followed by Fubini's theorem, for any $Q \in \mathcal{R}_M$ we have

$$\begin{aligned}\Phi_\nu^{n-1}(z_0,Q) &= \int_{\mathbb{T}^2} \Phi_\nu^{n-2}(z,Q)\Phi_\nu(z_0,\mathrm{d}z)\\ &= \int_\Omega \Big(\int_{\mathbb{T}^2} \chi_Q\big(T_{\nu,\xi}^{n-2}(z)\big)p(z_0,z)\,\mathrm{d}z\Big)\mathbf{P}(\mathrm{d}\xi),\end{aligned} \tag{19}$$

where χ_E denotes the characteristic function of a measurable set $E \subseteq \mathbb{T}^2$. Let $\bar{p}(z_0,\cdot)$ be the function that is constant on each $Q_j \in \mathcal{R}_M$ and such that

$$⨍_{Q_j} \bar{p}(z_0,z)\,\mathrm{d}z = ⨍_{Q_j} p(z_0,z)\,\mathrm{d}z := c_j.$$

Since $p(z_0,z) = G(z_0 - T^{-1}(z))$, where G is the fundamental solution to the heat equation on $\mathbb{T}^2$, we have $\|\nabla p(z_0,\cdot)\|_{L^\infty} \leq C_\alpha \nu^{-3/2}$ for some constant C_α depending on α. It follows then that $\|\bar{p}(z_0,\cdot) - p(z_0,\cdot)\|_{L^\infty} \leq C_\alpha\sqrt{2\nu}$. Now, by Lemma 2.1, for every $\xi \in \Omega$ we have

$$\begin{aligned}\int_{\mathbb{T}^2} \chi_Q\big(T_{\nu,\xi}^{n-2}(z)\big)\bar{p}(z_0,z)\,\mathrm{d}z &= \sum_{Q_j\in\mathcal{R}_M} c_j \int_{\mathbb{T}^2} \chi_Q\big(T_{\nu,\xi}^{n-2}(z)\big)\chi_{Q_j}(z)\,\mathrm{d}z\\ &\geq \lambda \sum_{Q_j\in\mathcal{R}_M} c_j m(Q_j)m(Q)\\ &= \lambda m(Q).\end{aligned}$$

Putting this estimate into (19) and writing $p(z_0, z) = p(z_0, z) - \bar{p}(z_0, z) + \bar{p}(z_0, z)$, we obtain

$$\begin{aligned}\Phi_\nu^{n-1}(z_0, Q) &\geq \lambda m(Q) - C_\alpha \sqrt{2\nu} \int_\Omega \Big(\int_{\mathbb{T}^2} \chi_Q\big(T_{\nu,\xi}^{n-2}(z)\big)\, \mathrm{d}z \Big) \mathbf{P}(\mathrm{d}\xi) \\ &= (\lambda - C_\alpha \sqrt{2\nu}) m(Q).\end{aligned}$$

Thus, for ν sufficiently small we have

$$\inf_{Q \in \mathcal{R}_M} \inf_{z_0 \in \mathbb{T}^2} \frac{\Phi_\nu^{n-1}(z_0, Q)}{m(Q)} \geq \frac{\lambda}{2}. \tag{20}$$

The bound (18) then follows by writing

$$\Phi_\nu^{n-1}(z_0, R) = \sum_{Q \in \mathcal{R}_M, Q \subseteq R} \Phi_\nu^{n-1}(z_0, Q)$$

and using (20).

We now use (18) to conclude the lower bound claimed in the lemma for $p_n(z_0, z)$. For ease of notation we write $h(z) = p_{n-1}(z_0, z)$. Note that $p_n(z_0, z) = (e^{\nu\Delta} h)(T^{-1}(z))$. Since composition with T^{-1} preserves global lower bounds, it suffices to show that there exists $\delta_1 > 0$ that does not depend on ν or z_0 such that

$$(e^{\nu\Delta} h)(z) \geq \delta_1$$

for all $z \in \mathbb{T}^2$. Let $\bar{h}$ denote the periodic extension of h to $\mathbb{R}^2$, and define

$$P = [0, 2^{-N+1}] \times [0, 2^{-N+1}] \subseteq \mathbb{R}^2.$$

Observe that (18) implies

$$\int_P \bar{h}(z - z')\, \mathrm{d}z' \geq \lambda 2^{-2N-1} \geq \frac{\lambda\nu}{8}$$

for any $z \in \mathbb{T}^2$. Since $|z'|^2 \leq 8\nu$ for $z' \in P$, we thus have

$$\begin{aligned}(e^{\nu\Delta} h)(z) &= \frac{1}{4\pi\nu} \int_{\mathbb{R}^2} \exp\Big(\frac{-|z'|^2}{4\nu}\Big) \bar{h}(z - z')\, \mathrm{d}z' \\ &\geq \frac{1}{4\pi\nu} \int_P \exp\Big(\frac{-|z'|^2}{4\nu}\Big) \bar{h}(z - z')\, \mathrm{d}z' \geq \frac{\lambda}{32\pi e^2}.\end{aligned}$$ □

With Lemma 2.3 at hand we are now ready to prove Theorem 3. We will use Lemma 2.2 and the lower bound in (17) to deduce convergence in L^∞, and then employ the Riesz–Thorin interpolation theorem to extend to the range $p \in (1, \infty)$.

Proof of Theorem 3

Let $\delta_1, C_1 > 0$ be as in the statement of Lemma 2.3, and let n_0 be the first natural number greater than or equal to $C_1|\log\nu|$. The lower bound in (17) implies that

$$\Phi_\nu^{n_0}(z,\cdot) \geq \delta_1 m(\cdot)$$

for every $z \in \mathbb{T}^2$. Therefore, by Lemma 2.2 there exists a constant $\delta_2 > 0$ depending only on δ_1 such that for every $k \in \mathbb{N}$ and bounded, Borel measurable function g with mean zero we have

$$\|\Phi_\nu^{kn_0} g\|_\infty \leq 2e^{-\delta_2 k}\|g\|_\infty.$$

In other words, for every $n \in n_0\mathbb{N}$ we have

$$\|\Phi_\nu^n g\|_{L^\infty} \leq 2\exp\Big(-\frac{\delta_2 n}{n_0}\Big)\|g\|_{L^\infty} \leq 2\exp\Big(-\frac{\delta_2 n}{(C_1+2)|\log\nu|}\Big)\|g\|_{L^\infty}. \tag{21}$$

This gives the desired estimate for $n \in n_0\mathbb{N}$. The estimate at intermediate times follows in a standard way from (21) and the fact that Φ_ν^m is bounded on L^∞ uniformly in $m \in \mathbb{N}$. This completes the proof of (16) in the case $p = \infty$.

We now prove (16) for general $p \in (1,\infty)$ using the $p = \infty$ case established above. Observe that for $f \in L^1$ we have

$$\begin{aligned}\|\Phi_\nu^n f\|_{L^1} &= \int_{\mathbb{T}^2}\Big|\int_{\mathbb{T}^2} p_n(z,z')f(z')\,\mathrm{d}z'\Big|\,\mathrm{d}z \leq \int_{\mathbb{T}^2}|f(z')|\int_{\mathbb{T}^2} p_n(z,z')\,\mathrm{d}z\,\mathrm{d}z' \\ &= \|f\|_{L^1},\end{aligned}$$

where in the last equality $\int_{\mathbb{T}^2} p_n(z,z')\,\mathrm{d}z = 1$ for all z' because Lebesgue measure is invariant for Φ_ν^n. Thus,

$$\|\Phi_\nu^n\|_{L^1\to L^1} \leq 1, \tag{22}$$

and (16) for $p \in (1,\infty)$ follows by interpolating (22) with the $p = \infty$ bound and applying the Riesz–Thorin theorem. □

2.4. *Continuous-time enhanced dissipation*

In this section, we prove that u_α is dissipation enhancing with the optimal rate function $\delta(\nu) = C|\log\nu|^{-1}$ by approximating the continuous-time diffusion with the pulsed diffusion and applying Theorem 3. The approximation relies only on the fact that the velocity field is Lipschitz and is thus quite general. We will therefore formulate it as an abstract result, as it may be of independent interest.

We begin by describing the general setting that we consider. Let $v : [0,1]\times\mathbb{T}^2 \to \mathbb{R}^2$ be a Lipschitz continuous (in space), divergence-free velocity field satisfying

$$\sup_{t\in[0,1]} \left\| v(t) \right\|_{\mathrm{Lip}} < \infty.$$

We assume moreover that there is a partition $0 = t_0 < t_1 < \cdots < t_m = 1$ of $[0,1]$ so that $v \in C([t_{i-1}, t_i) \times \mathbb{T}^2)$ for each $i = 1, \ldots, m$. Let ψ_t denote the flow map associated with v, and for $\nu > 0$ let Ψ_ν be the operator given by

$$\Psi_\nu f = e^{\nu\Delta}(f \circ \psi_1^{-1}).$$

The iterates of Ψ_ν define the pulsed diffusion generated by v.

It is clear that $v = u_\alpha$ satisfies the assumptions above. Our general statement in the present setting is that the L^2 decay timescale of the advection-diffusion equation defined by the periodic-in-time extension of v is at least as fast as the L^2 decay timescale of the pulsed diffusion Ψ_ν.

LEMMA 2.4

Let v and Ψ_ν be as above, and for $f_0 \in L^2$ let f denote the solution of (1) with u the periodic-in-time extension of $v : [0,1) \times \mathbb{T}^2 \to \mathbb{R}^2$. Suppose that there exists $n_0 \in \mathbb{N}$ such that for every mean-zero $g \in L^2$ there holds

$$\|\Psi_\nu^{n_0} g\|_{L^2} \le \frac{1}{2}\|g\|_{L^2}. \tag{23}$$

Then there exists $\delta \in (0,1)$ depending only on $\sup_{0\le t\le 1} \|v(t)\|_{\mathrm{Lip}}$ such that for every mean-zero $f_0 \in L^2$ we have

$$\left\| f(n_0+1) \right\|_{L^2} \le (1-\delta)\|f_0\|_{L^2}. \tag{24}$$

By Theorem 3 applied with $p = 2$, Φ_ν satisfies (23) for some $n_0 \approx |\log(\nu)|$. The desired enhanced dissipation estimate for u_α at integer multiples of the time $t_0 = n_0$ then follows by iterating Lemma 2.4. The bound at the intermediate times is obtained using the monotonicity of $t \mapsto \|f(t)\|_{L^2}$ for L^2 solutions of (1). To complete the proof of Theorem 1, it only remains to prove Lemma 2.4.

Throughout the remainder of this section, for a given mean-zero $f_0 \in L^2$ we write $f(t)$ for the solution of (1) with u the periodic-in-time extension of $v : [0,1) \times \mathbb{T}^2 \to \mathbb{R}^2$. In order to approximate the (continuous-time) f by the (discrete-time) pulsed diffusion, it will be convenient to introduce some auxiliary continuous-time functions. First, let $F : [0,\infty) \times \mathbb{T}^2 \to \mathbb{R}$ be defined on the time interval $[n, n+1)$ for each integer $n \ge 0$ by

$$F(t) = \begin{cases} f(n/2 + (t-n)) & n \text{ even}, \\ f((n+1)/2) & n \text{ odd}. \end{cases}$$

That is, F evolves in the same way as f but is chosen to be constant on every other unit time interval. Next, we define the natural continuous-time operator $S_\nu^t : L^2 \to L^2$ satisfying $S_\nu^{2n} = \Psi_\nu^n$ for every $n \in \mathbb{N}$. In particular, for $g \in L^2$ let

$$S_\nu^t g = \begin{cases} (\Psi_\nu^{\lfloor t/2 \rfloor} g) \circ \psi_{t-2\lfloor t/2 \rfloor}^{-1} & 0 \le t - 2\lfloor t/2 \rfloor \le 1, \\ e^{\nu(t-2\lfloor t/2 \rfloor - 1)\Delta}((\Psi_\nu^{\lfloor t/2 \rfloor} g) \circ \psi_1^{-1}) & 1 < t - 2\lfloor t/2 \rfloor < 2, \end{cases}$$

where $\lfloor t/2 \rfloor$ denotes the largest integer less than or equal to $t/2$. Then, let $F_d(t) = S_\nu^t f_0$. This definition is simply such that F_d solves the transport equation with velocity v for $t \in (0, 1)$, the heat equation for $t \in (1, 2)$, and so on.

A key step in the proof of Lemma 2.4 is an estimate on the error between F_d and the continuous-time diffusion f.

LEMMA 2.5
There exists a constant $C_1 > 0$ depending on $\sup_{0 \le t \le 1} \|v(t)\|_{\mathrm{Lip}}$ *but not ν so that for every $n \in \mathbb{N}$, $\epsilon > 0$, and mean-zero $f_0 \in L^2$ we have*

$$\begin{aligned} \|F_d(2n) - f(n)\|_{L^2}^2 &= \|F_d(2n) - F(2n)\|_{L^2}^2 \\ &\le C_1 \epsilon \nu \int_0^{2n} \|\nabla F_d(s)\|_{L^2}^2 \,\mathrm{d}s + C_1 \epsilon^{-1} \nu \int_0^n \|\nabla f(s)\|_{L^2}^2 \,\mathrm{d}s. \end{aligned}$$

Proof
Here and throughout the remainder of this subsection, C denotes a generic constant which may depend on $\sup_{0 \le t \le 1} \|v(t)\|_{\mathrm{Lip}}$ but not on ν or the initial data f_0. We will prove that for every $n \in \mathbb{N}$ there holds

$$\begin{aligned} \|F_d(2n) - F(2n)\|_{L^2}^2 \le C \epsilon \nu \int_0^{2n} \|\nabla F_d(s)\|_{L^2}^2 \,\mathrm{d}s \\ + C \epsilon^{-1} \nu \int_0^{2n} \|\nabla F(s)\|_{L^2}^2 \,\mathrm{d}s. \end{aligned} \tag{25}$$

Given (25), the lemma follows because

$$\int_0^{2n} \|\nabla F(s)\|_{L^2}^2 \,\mathrm{d}s = \int_0^n \|\nabla f(s)\|_{L^2}^2 \,\mathrm{d}s + \sum_{k=1}^n \|\nabla f(k)\|_{L^2}^2 \le C \int_0^n \|\nabla f(s)\|_{L^2}^2 \,\mathrm{d}s,$$

where in the inequality we used that

$$\|\nabla f(k)\|_{L^2}^2 \le e^{Ct} \|\nabla f(k-t)\|_{L^2}^2 \implies \|\nabla f(k)\|_{L^2}^2 \le C \int_{k-1}^k \|\nabla f(s)\|_{L^2}^2 \,\mathrm{d}s,$$

which holds by standard energy estimates for (1) with a uniformly-in-time Lipschitz velocity field. It remains then just to prove (25). We will assume $n = 1$, as the case

$n > 1$ follows by iterating the same argument. For $t \in (0,1)$, we have

$$\partial_t F_d + v \cdot \nabla F_d = 0$$

and

$$\partial_t F + v \cdot \nabla F = \nu \Delta F.$$

Thus, an energy estimate gives

$$\frac{d}{dt}\| F_d(t) - F(t) \|_{L^2}^2 + 2\int_{\mathbb{T}^2} (F_d - F) v \cdot \nabla (F_d - F) = 2\nu \int_{\mathbb{T}^2} (F - F_d)\Delta F.$$

Integrating by parts, we obtain

$$\frac{d}{dt}\| F_d(t) - F(t) \|_{L^2}^2 \leq 2\nu \| \nabla F_d(t) \|_{L^2} \| \nabla F(t) \|_{L^2},$$

so

$$\| F_d(1) - F(1) \|_{L^2}^2 \leq 2\nu \int_0^1 \| \nabla F_d(t) \|_{L^2} \| \nabla F(t) \|_{L^2} \,\mathrm{d}t. \tag{26}$$

A similar computation using that for $t \in (1,2)$ we have $\partial_t F_d = \nu \Delta F_d$ and $\partial_t F = 0$ shows that

$$\| F_d(2) - F(2) \|_{L^2}^2 \leq \| F_d(1) - F(1) \|_{L^2}^2 + 2\nu \int_1^2 \| \nabla F_d(t) \|_{L^2} \| \nabla F(t) \|_{L^2} \,\mathrm{d}t. \tag{27}$$

Combining (26) and (27) completes the proof of (25) with $n = 1$. ☐

We now give a uniform-in-ν bound on $\nu \int_0^\infty \| \nabla F_d(s) \|_{L^2}^2 \,\mathrm{d}s$ which we will use in the approximation estimate of Lemma 2.5.

LEMMA 2.6
Suppose that $f_0 \in L^2$ *is such that*

$$\nu \| \nabla f_0 \|_{L^2}^2 \leq C_0 \| f_0 \|_{L^2}^2$$

for some constant $C_0 > 0$. *Then there exists a constant* C_2 *depending on* C_0 *and* $\sup_{0 \leq t \leq 1} \| v(t) \|_{\mathrm{Lip}}$, *but not on* ν *or* f_0, *such that*

$$\nu \int_0^\infty \| \nabla F_d(t) \|_{L^2}^2 \,\mathrm{d}t \leq C_2 \| f_0 \|_{L^2}^2.$$

Proof
First observe that by the definition of F_d, the conservation of the L^2 norm for solutions of (4), and the energy equality for the heat equation, we have

$$\nu \sum_{n=1}^{\infty} \int_{2n-1}^{2n} \|\nabla F_d(t)\|_{L^2}^2 \,\mathrm{d}t \le \frac{1}{2}\|f_0\|_{L^2}^2. \tag{28}$$

Moreover, since F_d solves the transport equation for $t \in (0,1)$ and v is uniformly Lipschitz, we have

$$\nu \int_0^1 \|\nabla F_d(t)\|_{L^2}^2 \,\mathrm{d}t \le C\nu \|\nabla F_d(0)\|_{L^2}^2 = C\nu \|\nabla f_0\|_{L^2}^2 \le C C_0 \|f_0\|_{L^2}^2. \tag{29}$$

Combining (28) and (29), it suffices to show that

$$\nu \int_{2n-1}^{2n+1} \|\nabla F_d(t)\|_{L^2}^2 \,\mathrm{d}t \le C\nu \int_{2n-1}^{2n} \|\nabla F_d(t)\|_{L^2}^2 \,\mathrm{d}t \tag{30}$$

for every $n \in \mathbb{N}$. We will just prove the estimate for $n = 1$, as the case where $n > 1$ is no different. For $t \in (1,2)$, the pulsed diffusion F_d solves the heat equation, and hence $t \mapsto \|\nabla F_d(t)\|_{L^2}$ is monotone decreasing. Thus,

$$\nu \int_1^2 \|\nabla F_d(t)\|_{L^2}^2 \,\mathrm{d}t \ge \nu \|\nabla F_d(2)\|_{L^2}^2. \tag{31}$$

Now, since F_d solves the advection again for $t \in (2,3)$, we have

$$\nu \int_2^3 \|\nabla F_d(t)\|_{L^2}^2 \,\mathrm{d}t \le \nu \sup_{2 \le t \le 3} \|\nabla F_d(t)\|_{L^2}^2 \le C\nu \|\nabla F_d(2)\|_{L^2}^2. \tag{32}$$

Combining (31) and (32) gives (30) with $n = 1$, which completes the proof. ☐

With Lemmas 2.5 and 2.6 at hand, we are ready to complete the proof of Lemma 2.4.

Proof of Lemma 2.4
Fix mean-zero $f_0 \in L^2$, and assume for now that

$$\nu \|\nabla f_0\|_{L^2}^2 \le C_0 \|f_0\|_{L^2}^2 \tag{33}$$

for some constant $C_0 > 0$. Let n_0 be as in the statement of Lemma 2.4. The basic energy estimate for f on the time interval $[0, n_0]$ reads

$$\|f(n_0)\|_{L^2}^2 = \|f_0\|_{L^2}^2 - 2\nu \int_0^{n_0} \|\nabla f(t)\|_{L^2}^2 \,\mathrm{d}t. \tag{34}$$

We claim that there exists a constant $c_0 > 0$ that does not depend on ν such that

$$2\nu \int_0^{n_0} \|\nabla f(t)\|_{L^2}^2 \,\mathrm{d}t \ge c_0 \|f_0\|_{L^2}^2. \tag{35}$$

To prove this, we suppose that instead

$$2\nu \int_0^{n_0} \left\|\nabla f(t)\right\|_{L^2}^2 \,\mathrm{d}t \le \delta^2 \|f_0\|_{L^2}^2 \tag{36}$$

holds for some $\delta \in (0,1)$ will show that δ must be bounded below independently of ν. Let F_d be as defined earlier, and let C_1 and C_2 be as in the statements of Lemmas 2.5 and 2.6, respectively. By (36), Lemma 2.5 applied with $\epsilon = \delta$, and Lemma 2.6, we have

$$\left\|F_d(2n_0) - f(n_0)\right\|_{L^2}^2 \le C_1(C_2+1)\delta\|f_0\|_{L^2}^2 := C_3\delta\|f_0\|_{L^2}^2. \tag{37}$$

Using the assumption (23), we have

$$\left\|F_d(2n_0)\right\|_{L^2} = \|\Psi_\nu^{n_0} f_0\|_{L^2} \le \frac{1}{2}\|f_0\|_{L^2}.$$

Thus, from the triangle inequality and (37), we get

$$\left\|f(n_0)\right\|_{L^2}^2 \le 2\left\|f(n_0) - F_d(2n_0)\right\|_{L^2}^2 + \frac{1}{2}\|f_0\|_{L^2}^2 \le \Big(2\delta C_3 + \frac{1}{2}\Big)\|f_0\|_{L^2}^2,$$

which together with (34) implies

$$2\nu \int_0^{n_0} \left\|\nabla f(t)\right\|_{L^2}^2 \,\mathrm{d}t = \|f_0\|_{L^2}^2 - \left\|f(n_0)\right\|_{L^2}^2 \ge \Big(\frac{1}{2} - 2\delta C_3\Big)\|f_0\|_{L^2}^2.$$

The previous estimate and (36) yield

$$\frac{1}{2} - 2\delta C_3 \le \delta^2 \implies \delta \ge \frac{1}{2(1+2C_3)},$$

which proves (35). Combining (35) with (34), we have shown that if (33) is satisfied, then there exists a constant $c > 0$ depending only on C_0 and $\sup_{0\le t\le 1}\|v(t)\|_{\mathrm{Lip}}$ such that

$$\left\|f(n_0)\right\|_{L^2} \le (1-c)\|f_0\|_{L^2}. \tag{38}$$

We now turn to the general case where (33) does not necessarily hold. If

$$\nu \int_0^1 \left\|\nabla f(t)\right\|_{L^2}^2 \,\mathrm{d}t \ge (1/4)\|f_0\|_{L^2}^2, \tag{39}$$

then since

$$\left\|f(1)\right\|_{L^2}^2 = \|f_0\|_{L^2}^2 - 2\nu \int_0^1 \left\|\nabla f(t)\right\|_{L^2}^2 \,\mathrm{d}t,$$

we see that

$$\|f(t)\|_{L^2}^2 \le \|f(1)\|_{L^2}^2 \le (1/2)\|f_0\|_{L^2}^2$$

for every $t \ge 1$, and there is nothing to prove. We may thus assume that (39) is not satisfied, in which case Chebyshev's inequality and $\|\nabla f(t)\|_{L^2} \le e^{C(t-s)}\|\nabla f(s)\|_{L^2}$ imply

$$\nu\|\nabla f(1)\|_{L^2} \le C\|f_0\|_{L^2}^2 \le C\|f(1)\|_{L^2}^2.$$

In the last step above we used that

$$\|f_0\|_{L^2}^2 \le 2\|f(1)\|_{L^2}^2$$

whenever (39) does not hold. Thus, we may assume that $f(1)$ satisfies (33). Applying (38) with f_0 replaced by $f(1)$, we conclude that

$$\|f(n_0+1)\|_{L^2} \le (1-c)\|f(1)\|_{L^2} \le (1-c)\|f_0\|_{L^2},$$

which completes the proof. ☐

Remark 2.3
As discussed in the introduction, similar arguments have appeared in previous works (see [29], [43]) and have been used to prove that a Lipschitz, exponentially mixing flow is dissipation enhancing on the timescale $|\log \nu|^2$. It is worth noting precisely what allows us to remove the 2 from the exponent of the logarithm. The standard approach is to approximate the advection-diffusion equation by the solution $f_0 \circ \psi_t^{-1}$ to the associated $\nu = 0$ problem and use a version of Lemma 2.5 with the error bound replaced by

$$\|f_0 \circ \psi_t^{-1} - f(t)\|_{L^2}^2 \le 2\nu \int_0^t \|\Delta f(s)\|_{L^2}\|f_0 \circ \psi_s^{-1}\|_{L^2}\,\mathrm{d}s.$$

Within this scheme, two derivatives are put on f because there are no a priori bounds available on the H^1 norm of the approximating $\nu = 0$ solution. Estimating $\|\Delta f\|_{L^2}$, however, leads to important losses. We are able to put a derivative on our approximating solution F_d because it experiences diffusion and hence satisfies good global H^1 bounds (Lemma 2.6 above). This is what underlies our ability to get the optimal enhanced dissipation timescale with our approximation method.

2.5. *Exponential mixing for u_α*

In this section, we obtain as a corollary of Lemma 2.1 that the geometric mixing scale of solutions to (4) decays exponentially fast. We also show how exponential mixing for u_α in the sense of Definition 1.2 follows by applying in addition the main result of [35].

2.5.1. *Decay of the geometric mixing scale*

We begin by recalling the definition of the geometric mixing scale (see [1], [40], [74]).

Definition 2.1

Given fixed $\kappa \in (0,1)$, the *geometric mixing scale* of a mean-zero function $f \in L^\infty$ is defined as

$$\operatorname{mix}_\kappa(f) = \inf\Big\{2^{-N} : \Big|⨍_R f\Big| \le \kappa \|f\|_{L^\infty}\ \forall R \in \mathcal{R}_N\Big\}.$$

In the simple case where $f = \chi_A - \chi_{A^c}$, $\operatorname{mix}_\kappa(f) = 2^{-N}$ for N the largest nonnegative integer such that every square $R \in \mathcal{R}_N$ consists of at least a fraction $(1-\kappa)/2$ of points where f is both $+1$ and -1.

Remark 2.4

For $\kappa \in (0,1)$, the geometric mixing scale is typically defined as the infimum over all $\epsilon > 0$ such that

$$\Big|⨍_{B_\epsilon(z)} f\Big| \le \kappa \|f\|_{L^\infty} \tag{40}$$

for every $z \in \mathbb{T}^2$, where $B_\epsilon(z)$ denotes the open ball of radius ϵ centered at z. Decay of $\operatorname{mix}_{\kappa'}(f)$ for some $\kappa' \in (0,1)$ in the sense of Definition 2.1 implies decay of the geometric mixing scale defined using (40). Indeed, it is easy to check that if $\operatorname{mix}_{\kappa'}(f) = 2^{-N}$, then there exists $\kappa \in (0,1)$ depending only on κ' such that (40) holds for every $z \in \mathbb{T}^2$ and $\epsilon \le C2^{-N}$, where C is a constant that does not depend on κ' or f.

THEOREM 4

For all α sufficiently large, there exist constants $\kappa \in (0,1)$ and $C, c > 0$ such that for every mean-zero $f \in C^1$ we have

$$\operatorname{mix}_\kappa(f \circ \phi_t^{-1}) \le C\|f\|_{C^1}\|f\|_{L^\infty}^{-1} e^{-ct}.$$

Proof

Let $\mathcal{H}_N$ denote the set of simple functions which are mean zero and constant on each element of $\mathcal{R}_N$. We will first consider functions that are in some $\mathcal{H}_N$ and then extend to the general case by approximation. Fix $M \in \mathbb{N}$ and $f \in \mathcal{H}_M$. Let C_0 and λ be as in Lemma 2.1. Let $N \ge M$ and $n \ge C_0 N$. By Lemma 2.1, for every $R_i, R_j \in \mathcal{R}_N$ we have

$$m\big(R_i \cap T^n(R_j)\big) = m\big(T^{-n}(R_i) \cap R_j\big) \ge \lambda m(R_i)m(R_j).$$

It follows that there exists a measurable set $Q_{ij} \subseteq R_i \cap T^n(R_j)$ such that

$$m(Q_{ij}) = m\big(T^{-n}(Q_{ij}) \cap R_j\big) = \lambda m(R_i)m(R_j).$$

Note that the sets Q_{ij} are disjoint because T^n is a bijection. Define

$$P_i = R_i \setminus \bigcup_j Q_{ij}.$$

Since $f \in \mathcal{H}_M$ and $N \geq M$, we can write $f = \sum_{R_i \in \mathcal{R}_N} c_i \chi_{R_i}$ and decompose f as

$$f = \sum_{R_i \in \mathcal{R}_N} c_i \chi_{P_i} + \sum_{R_i, R_j \in \mathcal{R}_N} c_i \chi_{Q_{ij}} := f_P + f_Q.$$

Then, for any $R_k \in \mathcal{R}_N$ we have

$$\begin{aligned} \fint_{R_k} f_Q \circ T^n &= \frac{1}{m(R_k)} \sum_{R_i \in \mathcal{R}_N} c_i m\big(T^{-n}(Q_{ik}) \cap R_k\big) \\ &= \lambda \sum_{R_i \in \mathcal{R}_N} c_i m(R_i) = \lambda \int_{\mathbb{T}^2} f = 0 \end{aligned} \tag{41}$$

and

$$\begin{aligned} \Big|\fint_{R_k} f_P \circ T^n\Big| &= \frac{1}{m(R_k)} \Big| \sum_{R_i \in \mathcal{R}_N} c_i m\big(T^{-n}(P_i) \cap R_k\big)\Big| \\ &\leq \|f\|_{L^\infty} \frac{1}{m(R_k)} \sum_{R_i \in \mathcal{R}_N} m\big(T^{-n}(P_i) \cap R_k\big). \end{aligned} \tag{42}$$

Observe now that

$$\begin{aligned} \sum_i m\big(T^{-n}(P_i) \cap R_k\big) &= \sum_{R_i \in \mathcal{R}_N} m\big(T^{-n}(R_i) \cap R_k\big) \\ &\quad - \sum_{R_i, R_j \in \mathcal{R}_N} m\big(T^{-n}(Q_{ij}) \cap R_k\big) \\ &= m(R_k) - \sum_{R_i, R_j \in \mathcal{R}_N} m\big(T^{-n}(Q_{ij}) \cap R_k\big) \\ &= m(R_k) - \sum_{R_i \in \mathcal{R}_N} m\big(T^{-n}(Q_{ik}) \cap R_k\big) \\ &= m(R_k) - \sum_{R_i \in \mathcal{R}_N} \lambda m(R_i)m(R_k) = (1-\lambda)m(R_k). \end{aligned}$$

Putting this equality into (42) and applying also (41) proves that

$$\sup_{R\in\mathcal{R}_N}\left|⨍_R f\circ T^n\right|\le(1-\lambda)\|f\|_{L^\infty}\tag{43}$$

for every $N\ge M$ and $n\ge C_0N$.

Now fix $f\in C^1$, and for M to be chosen, let $\tilde f\in\mathcal{H}_M$ be such that

$$\tilde f(z)=⨍_R f$$

for every $R\in\mathcal{R}_M$ and $z\in R$. Then, by (43), for every $N\ge M$, $n\ge C_0N$, and $R\in\mathcal{R}_N$, we have

$$\begin{aligned}\left|⨍_R f\circ T^n\right|&\le\left|⨍_R \tilde f\circ T^n\right|+\left|⨍_R (f-\tilde f)\circ T^n\right|\\&\le(1-\lambda)\|\tilde f\|_{L^\infty}+\|f-\tilde f\|_{L^\infty}\\&\le(1-\lambda)\|f\|_{L^\infty}+\big(\sup_{Q\in\mathcal{R}_M}\operatorname{diam}(Q)\big)\|f\|_{\mathrm{Lip}}\\&\le(1-\lambda)\|f\|_{L^\infty}+2^{-M+1}\|f\|_{\mathrm{Lip}}.\end{aligned}$$

Taking M to be the smallest natural number such that

$$2^{-M+1}\le(\lambda/2)\|f\|_{L^\infty}\|f\|_{\mathrm{Lip}}^{-1},\tag{44}$$

it follows that for $\kappa=1-\lambda/2$ there holds

$$\operatorname{mix}_\kappa(f\circ T^n)\le 2^{-N}\tag{45}$$

for every $N\ge M$ and $n\ge C_0N$. The choice of M as the minimal natural number such that (44) holds guarantees that

$$2^{-M+2}>\frac{\lambda}{2}\|f\|_{L^\infty}\|f\|_{\mathrm{Lip}}^{-1},$$

so (45) implies that for every $n\in\mathbb{N}$ we have

$$\operatorname{mix}_\kappa(f\circ T^n)\le 2^M2^{-n/C_0}\le(8/\lambda)\|f\|_{\mathrm{Lip}}\|f\|_{L^\infty}^{-1}2^{-n/C_0},\tag{46}$$

which is the desired estimate at integer times. The estimate at general times follows by applying (46) with f replaced by $f\circ\phi_s^{-1}$ for suitable $s\in(0,1)$ and noting that

$$\|f\circ\phi_s^{-1}\|_{\mathrm{Lip}}\|f\circ\phi_s^{-1}\|_{L^\infty}^{-1}\le C\|f\|_{C^1}\|f\|_{L^\infty}^{-1}$$

for a constant C depending only on α. ☐

2.5.2. *Decay of correlations*

We now establish exponential mixing for u_α in the sense of Definition 1.2, completing the proof of Theorem 2.

We begin with the exponential mixing of T. For α sufficiently large, the map T falls within the general class of piecewise hyperbolic maps studied in [35]. This is an easy consequence of the uniform hyperbolicity and singularity set structure of T discussed in Section 3. It follows from [35, Theorem 2.8] that if T^k is ergodic with respect to Lebesgue measure for every $k \in \mathbb{N}$, then T is mixing and enjoys exponential decay of correlations for C^1 observables. Note that we are using here that the measure $\bar{\mu}$ present in [35, Theorem 2.8] is simply Lebesgue measure in our setting, which follows immediately from the fact that Lebesgue measure is invariant for T. Thus, we just need to prove that T^k is ergodic, which we will show follows from Lemma 2.1.

LEMMA 2.7

For all α sufficiently large, T^k is ergodic with respect to Lebesgue measure for every $k \in \mathbb{N}$, and consequently there exist constants $c, C > 0$ such that

$$\left| \fint (f \circ T^n) g - \fint f \fint g \right| \le C e^{-cn} \|f\|_{C^1} \|g\|_{C^1}$$

for every $f, g \in C^1(\mathbb{T}^2)$.

Proof

The proof that T^k is ergodic relies only on the fact that Lemma 2.1 holds for $\xi = 0$ with T replaced by T^k, so we may take $k = 1$ without loss of generality. Let $\mathcal{H}_N$ be as defined in the proof of Theorem 4. Using Lemma 2.1 and following the idea of the proof of (43), we can show that there exists $\delta > 0$ such that for any $N \in \mathbb{N}$ and $f \in \mathcal{H}_N$ there exists $n \in \mathbb{N}$ for which

$$\left| \int_{\mathbb{T}^2} (f \circ T^n) f \right| \le (1 - \delta) \|f\|_{L^2}^2. \tag{47}$$

We omit the details for the sake of brevity.

Let $f \in L^2$ be mean zero and invariant for T. That is, $f \circ T = f$ almost everywhere. We need to show that $f = 0$. Suppose for contradiction that this is not the case. Then, by standard approximation arguments, for any $\epsilon > 0$ there exist $N \in \mathbb{N}$ and $\psi \in \mathcal{H}_N$ such that $\|f - \psi\|_{L^2} \le \epsilon \|f\|_{L^2}$. Since f is invariant for T, for any $n \in \mathbb{N}$ we get from the triangle inequality that

$$\|f\|_{L^2}^2 = \int_{\mathbb{T}^2} (f \circ T^n) f \le \left| \int_{\mathbb{T}^2} (\psi \circ T^n) \psi \right| + (3\epsilon^2 + 2\epsilon) \|f\|_{L^2}^2.$$

By (47), we can choose n such that

$$\|f\|_{L^2}^2 \le (1-\delta)(1+\epsilon)^2\|f\|_{L^2}^2 + (3\epsilon^2+2\epsilon)\|f\|_{L^2}^2.$$

Taking ϵ small enough so that $(1-\delta)(1+\epsilon)^2 + 3\epsilon^2 + 2\epsilon < 1$ yields a contradiction. □

It is now a fairly routine argument to upgrade to exponential mixing in continuous time.

Proof of Theorem 2

Fix $f, g \in C^1(\mathbb{T}^2)$, which we may assume without loss of generality are both mean zero. Choose $t > 0$, and let $\lfloor t \rfloor$ denote the first integer less than or equal to t. By Lemma 2.7, there exist $C, c > 0$ such that

$$\left|\int_{\mathbb{T}^2}(f\circ\phi_t^{-1})g\right| = \left|\int_{\mathbb{T}^2}(f\circ T^{\lfloor t\rfloor})(g\circ\phi_{t-\lfloor t\rfloor})\right| \le Ce^{-c\lfloor t\rfloor}\|f\|_{C^1}\|g\circ\phi_{t-\lfloor t\rfloor}\|_{C^1}.$$

Since u_α is uniformly Lipschitz, there exists a constant $C_\alpha > 0$ such that $\|g\circ\phi_{t-\lfloor t\rfloor}\|_{C^1} \le C_\alpha\|g\|_{C^1}$. It follows that

$$\left|\int_{\mathbb{T}^2}(f\circ\phi_t^{-1})g\right| \le (CC_\alpha e^c)e^{-ct}\|f\|_{C^1}\|g\|_{C^1},$$

completing the proof. □

3. Dynamics estimates

In this section, we obtain the main dynamics estimates necessary for the proof of Lemma 2.1. In Section 3.1 we record precisely the basic facts concerning the uniform hyperbolicity of T. Then, in Section 3.2 we prove the complexity lemma, which describes how the stretching of unstable curves under the iterates of T dominates the cutting across singularity curves.

3.1. *Uniform hyperbolicity*

Define the mappings T_1 and T_2 by

$$T_1(x,y) = \begin{pmatrix} x \\ y+\alpha|x-1/2| \end{pmatrix} \quad \text{mod } 1,$$

$$T_2(x,y) = \begin{pmatrix} x+\alpha|y-1/2| \\ y \end{pmatrix} \quad \text{mod } 1,$$

and observe that $T_2 \circ T_1 = T$. Assuming that $\alpha \ge 2$ is an even integer and writing $z = (x, y)$, it is easy to see that

$$T_1(z) = \begin{cases} \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} z \mod 1 & x \geq 1/2, \\ \begin{pmatrix} 1 & 0 \\ -\alpha & 1 \end{pmatrix} z \mod 1 & x < 1/2, \end{cases} \tag{48}$$

and similarly

$$T_2(z) = \begin{cases} \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} z \mod 1 & y \geq 1/2, \\ \begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix} z \mod 1 & y < 1/2. \end{cases} \tag{49}$$

Let

$$\mathcal{A} = \left\{ \begin{pmatrix} 1+\alpha^2 & \alpha \\ \alpha & 1 \end{pmatrix}, \begin{pmatrix} 1-\alpha^2 & \alpha \\ -\alpha & 1 \end{pmatrix}, \begin{pmatrix} 1-\alpha^2 & -\alpha \\ \alpha & 1 \end{pmatrix}, \begin{pmatrix} 1+\alpha^2 & -\alpha \\ -\alpha & 1 \end{pmatrix} \right\}$$

denote the collection of possible products BA, where A is one of the matrices in (48) and B is one of the matrices in (49). The map T is smooth away from the singularity set

$$\mathcal{S}^+ = \big(\{x = 0\} \cup \{x = 1/2\}\big) \cup T_1^{-1}\big(\{y = 1/2\} \cup \{y = 0\}\big), \tag{50}$$

which partitions $\mathbb{T}^2 \setminus \mathcal{S}^+$ into finitely many open, connected components such that on each one $Tz = Az \mod 1$ for some $A \in \mathcal{A}$. One can check by direct computation that there exists $C \geq 1$ such that for any $\alpha \geq 4$, every matrix $A \in \mathcal{A}$ has eigenvalues $\lambda_u = c_\alpha$, $\lambda_s = 1/c_\alpha$ for some $c_\alpha \in \mathbb{R}$ with $|c_\alpha| \geq \alpha^2/4$ and associated normalized eigenvectors

$$e_u = \begin{pmatrix} \cos\theta_u \\ \sin\theta_u \end{pmatrix}, \qquad e_s = \begin{pmatrix} \cos\theta_s \\ \sin\theta_s \end{pmatrix}$$

that satisfy $|\tan(\theta_u)| \leq C\alpha^{-1}$ and $|\tan(\theta_s)| \geq C^{-1}\alpha$. As mentioned in the introduction, this implies that T is uniformly hyperbolic for α large. We now describe this uniform hyperbolicity more precisely. Let $\mathcal{S}^- = T(\mathcal{S}^+)$ denote the singularity set for T^{-1}, and for $\delta_1 \in (0,1)$ let C_u and C_s be the cones defined in Section 1.2.2. Then, for a suitable choice of δ_1 and all α sufficiently large we have

$$(\nabla T)(z) C_u \subset C_u \qquad \text{and} \qquad (\nabla T^{-1})(z) C_s \subset C_s,$$

where the two inclusions hold for all $z \in \mathbb{T}^2 \setminus \mathcal{S}^+$ and all $z \in \mathbb{T}^2 \setminus \mathcal{S}^-$, respectively.

Moreover, we have that

$$\inf_{v\in C_u}\inf_{z\in\mathbb{T}^2\setminus\mathcal{S}^+}\frac{\|(\nabla T)(z)v\|}{\|v\|}\geq\delta_1\alpha^2,\qquad\inf_{v\in C_s}\inf_{z\in\mathbb{T}^2\setminus\mathcal{S}^-}\frac{\|(\nabla T^{-1})(z)v\|}{\|v\|}\geq\delta_1\alpha^2.$$

Since the uniform hyperbolicity of T was a simple consequence of the stable and unstable eigenvectors being approximately aligned everywhere in space, it is clear that the hyperbolic properties of T extend for free to the randomly kicked maps T_ξ^n. For $z_0\in\mathbb{R}^2$, let $\mathcal{S}_{z_0}^+=\mathcal{S}^+-\pi(z_0)$, where as before $\pi:\mathbb{R}^2\to\mathbb{T}^2$ denotes the natural projection. Then, for $\xi\in\Omega$ and $n\in\mathbb{N}$ define

$$\mathcal{S}_\xi^{+,n}=\bigcup_{k=1}^{n}T_\xi^{-(k-1)}(\mathcal{S}_{\xi_k}^+),\qquad\mathcal{S}_\xi^{-,n}=T_\xi^n(\mathcal{S}_\xi^{+,n}).$$

With these definitions, $\mathcal{S}_\xi^{+,n}$ and $\mathcal{S}_\xi^{-,n}$ denote the singularity sets of T_ξ^n and T_ξ^{-n}, respectively. The following lemma summarizes the hyperbolic properties of T_ξ^n and T_ξ^{-n} away from their singularity sets.

LEMMA 3.1

Let $\delta_1\in(0,1)$ and $\alpha\gg1$ be as above. Fix $\xi\in\Omega$ and $n\in\mathbb{N}$. For any $z\in\mathbb{T}^2\setminus\mathcal{S}_\xi^{+,n}$, $(\nabla T_\xi^n)(z)\in\mathbb{R}^{2\times2}$ is hyperbolic with normalized eigenvectors $e_u\in C_u$, $e_s\in C_s$ and associated eigenvalues λ_u, λ_s which satisfy

$$|\lambda_u|\geq(\delta_1\alpha^2)^n,\qquad|\lambda_s|\leq(\delta_1\alpha^2)^{-n}.$$

Moreover, there holds

$$\inf_{v\in C_u}\inf_{z\in\mathbb{T}^2\setminus\mathcal{S}_\xi^{+,n}}\frac{\|(\nabla T_\xi^n)(z)v\|}{\|v\|}\geq(\delta_1\alpha^2)^n,$$

$$\inf_{v\in C_s}\inf_{z\in\mathbb{T}^2\setminus\mathcal{S}_\xi^{-,n}}\frac{\|(\nabla T_\xi^{-n})(z)v\|}{\|v\|}\geq(\delta_1\alpha^2)^n.$$

Remark 3.1

A key aspect of the preceding lemma is that all of the relevant bounds are *independent* of $\xi\in\Omega$.

3.2. *Complexity lemma*

Recall from Section 1.2.3 that $\mathcal{W}_u$ denotes the collection of line segments on $\mathbb{T}^2$ that are tangent to some $v\in C_u$. Since T is piecewise linear, the strict invariance of the unstable cone implies that if $W\in\mathcal{W}_u$, then for any $\xi\in\Omega$ we can write $T_\xi(W)$ as a finite, disjoint union of elements in $\mathcal{W}_u$. The main result of this section is the complexity bound of Lemma 1.1 from the introduction, restated precisely below. In what follows, $|\gamma|$ denotes the length of a piecewise smooth curve γ on $\mathbb{T}^2$.

LEMMA 3.2
There is a constant $C \geq 1$ *so that if* α *is sufficiently large, then for every* $W \in \mathcal{W}_u$ *with* $|W| \leq 1/2$, $\xi \in \Omega$, *and* $n \geq 1 + \log(1/|W|) + \log(8\alpha)$ *there is a finite collection* $\{W_i\}_{i\in I} \subseteq \mathcal{W}_u$ *such that*

$$T_\xi^n(W) = \bigcup_{i\in I} W_i$$

and

$$\sum_{|W_i|\geq 5/4} |T_\xi^{-n} W_i| \geq \Big(1 - \frac{C}{\alpha}\Big)|W|. \tag{51}$$

Moreover, W_{i_1} *and* W_{i_2} *are disjoint when* $i_1 \neq i_2$, *and for each* $i \in I$ *the interior of* W_i *with respect to the subspace topology of* $T_\xi^n(W)$ *is contained in* $\mathbb{T}^2 \setminus \mathcal{S}_\xi^{-,n}$.

Remark 3.2
The lower bound in (51) is the precise quantification of the fact that the line segments in $T_\xi^n(W)$ which span across the entire torus dominate for large α. The complexity estimate in this form says that the fraction of the initial segment W that is mapped to long lines is very close to 1. Phrasing (51) in this way is crucial to the area formula argument of Section 4.

Remark 3.3
Let $\mathcal{W}_s$ be defined in the same way as $\mathcal{W}_u$, but with C_u replaced by the stable cone C_s. Then, a decomposition of $T_\xi^{-n}(W)$ for $W \in \mathcal{W}_s$ with $|W| \leq 1/2$ exactly analogous to the one in Lemma 3.2 holds.

As mentioned earlier, our proof of Lemma 3.2 follows closely ideas from [4] and [34]. We utilize the fact that T is piecewise linear and the structure of the singularity set $\mathcal{S}^+$ to obtain the complexity estimate in the form of (51), which is slightly different from the corresponding complexity lemmas in [4] and [34].

3.2.1. Structure of the singularity sets

We begin by noting that from (50), we can deduce the following facts about $\mathcal{S}^+$, which is plotted in Figure 1:

- $\mathcal{S}^+$ is a union of $\{x = 0\}$, $\{x = \frac{1}{2}\}$ and finitely many lines of slope $\pm\frac{1}{\alpha}$ that partition $\mathbb{T}^2 \setminus \mathcal{S}^+$ into finitely many connected components $\mathcal{M}^+ = \{M_i^+\}$.

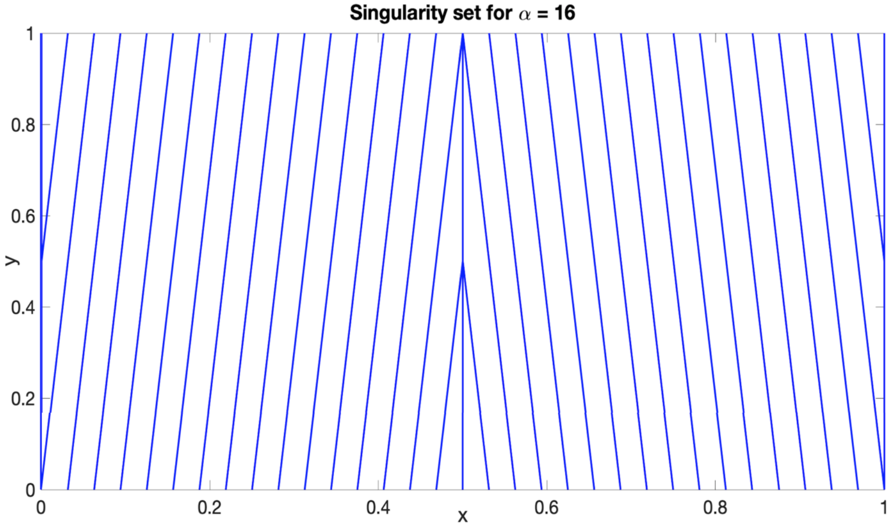

Figure 1. (Color online.) Plot of the singularity set $\mathcal{S}^+$ on $\mathbb{T}^2$ in the case where $\alpha = 16$.

- The lines with slope $\frac{1}{\alpha}$ are contained in the region $0 \leq x \leq \frac{1}{2}$ and are horizontally spaced by $\frac{1}{2\alpha}$. The lines with slope $-\frac{1}{\alpha}$ are similarly in $\frac{1}{2} \leq x \leq 1$.
- The maximal number of singularity curves that intersect at a point is three.
- If $W \in \mathcal{W}_u$ with $|W| \leq \alpha^{-1}/4$ and α is sufficiently large, then W can have nonempty intersection with at most four distinct elements of $\mathcal{M}^+$.

Recall that for $z_0 \in \mathbb{R}^2$, $\mathcal{S}_{z_0}^+$ denotes the singularity of T_{z_0}. As the set $\mathcal{S}_{z_0}^+$ is obtained simply by translating $\mathcal{S}^+$, it has similar properties. In particular, the last bullet remains true for $\mathcal{S}_{z_0}^+$ for any $z_0 \in \mathbb{R}^2$ with $\mathcal{M}^+$ replaced by its translate $\mathcal{M}_{z_0}^+ = \{M_{z_0,i}^+\}$.

3.2.2. Generations of lines in $\mathcal{W}_u$

Given $\xi = (\xi_1, \xi_2, \dots) \in \Omega$ and $W \in \mathcal{W}_u$ with $|W| \leq \alpha^{-1}/4$, we define subsequent generations $\mathcal{G}_n(\xi, W)$ of lines in $\mathcal{W}_u$ with length less than or equal to $\alpha^{-1}/4$ obtained by mapping iteratively through T_{ξ_1}, T_{ξ_2}, and so on. First, let $\{M_{\xi_1,i_j}^+\}_{j\in J} \subseteq \mathcal{M}_{\xi_1}^+$ denote the distinct elements of $\mathcal{M}_{\xi_1}^+$ with which W has nonempty intersection and note that as described above $\#J \leq 4$. Let $\tilde{W}_j = T_{\xi_1}(W \cap \overline{M_{\xi_1,i_j}^+})$, and then obtain $W_j \in \mathcal{W}_u$ from $\tilde{W}_j$ by modifying just the endpoints in such a way that $\{W_j\}_{j\in J} \subseteq \mathcal{W}_u$ consists of disjoint segments. We then define $\mathcal{G}(\xi_1, W) \subseteq \mathcal{W}_u$ as follows. If $|W_j| \leq \alpha^{-1}/4$, then we declare $W_j \in \mathcal{G}(\xi_1, W)$. If $|W_j| > \alpha^{-1}/4$, then we subdivide W_j into disjoint elements $\{W_{ij}\} \subseteq \mathcal{W}_u$ with $|W_{ij}| \in (\alpha^{-1}/8, \alpha^{-1}/4)$ and declare $W_{ij} \in \mathcal{G}(\xi_1, W)$ for each i. Set $\mathcal{G}_1(\xi, W) = \mathcal{G}(\xi_1, W)$, and then for $n \geq 2$ define

$$\mathcal{G}_n(\xi, W) = \bigcup_{V \in \mathcal{G}_{n-1}(\xi, W)} \mathcal{G}(\xi_n, V).$$

It is clear that for every $n \in \mathbb{N}$ there holds

$$T_\xi^n(W) = \bigcup_{W_i \in \mathcal{G}_n(\xi, W)} W_i$$

and that two line segments $W_{i_1}, W_{i_2} \in \mathcal{G}_n(\xi, W)$ are disjoint when $i_1 \neq i_2$. Moreover, note that by construction the interior of each $V \in \mathcal{G}_n(\xi, W)$ with respect to the subspace topology of $T_\xi^n(W)$ is contained in $\mathbb{T}^2 \setminus \mathcal{S}_\xi^{-,n}$. In other words, the inverse map T_ξ^{-n} is smooth on a neighborhood of the interior of V. Similarly, $T_{\theta^{n-k}\xi}^{-k}$ is smooth on a neighborhood of the interior of V for every $k = 1, 2, \ldots, n$, where we recall that $\theta : \Omega \to \Omega$ denotes the shift map.

For $W \in \mathcal{W}_u$ and $\mathcal{G}_n(\xi, W)$ constructed as above, let $L_n(\xi, W)$ denote the elements $W_i \in \mathcal{G}_n(\xi, W)$ with $|W_i| \geq \alpha^{-1}/8$, and let $S_n(\xi, W) = \mathcal{G}_n(\xi, W) \setminus L_n(\xi, W)$. Here the notation is "L" for "long" and "S" for "short." We define also $I_n(\xi, W)$ to be the elements $W_i \in S_n(\xi, W)$ such that for every $k = 1, 2, \ldots, n-1$ there exists $V_k \in S_{n-k}(\xi, W)$ such that $W_i \in S_k(\theta^{n-k}\xi, V_k)$. Stated roughly in simple terms, $I_n(\xi, W)$ is the collection of elements in $S_n(\xi, W)$ that have always been contained in a short element. Lastly, for $\xi \in \Omega$, $k \in \mathbb{N}$, and a collection E of disjoint elements of $\mathcal{W}_u$ we write

$$|E| = \sum_{W_i \in E} |W_i| \qquad \text{and} \qquad T_\xi^{\pm k} E = \{T_\xi^{\pm k} W : W \in E\}.$$

3.2.3. *Proof of Lemma 3.2*

The first step in the proof of Lemma 3.2 is a lemma which says roughly that for an initial $W \in \mathcal{W}_u$ with $|W| \approx \alpha^{-1}$, the fraction of its initial length which corresponds to lines which never become long up to time n decays exponentially in n.

LEMMA 3.3

Let $\epsilon \in (0, 2)$. There exists $\alpha_0(\epsilon) \geq 0$ so that if $\alpha \geq \alpha_0$, then for every $\xi \in \Omega$, $W \in \mathcal{W}_u$ with $|W| \in (\alpha^{-1}/8, \alpha^{-1}/4]$, and $n \in \mathbb{N}$ there holds

$$\left|T_\xi^{-n} I_n(\xi, W)\right| \leq \alpha^{-(2-\epsilon)n} |W|.$$

Proof

For any $|W| \leq \alpha^{-1}/4$, we have $\# I_1(\xi, W) \leq 4$. It is therefore easy to see by induction and the fact that

$$I_n(\xi, W) = \bigcup_{V \in I_{n-1}(\xi, W)} I_1(\theta^{n-1}\xi, V)$$

that

$$\#I_n(\xi, W) \leq 4^n \tag{52}$$

for every $n \in \mathbb{N}$. On the other hand, by the hyperbolicity described in Lemma 3.1, for any $V \in I_n(\xi, W) \subseteq S_n(\xi, W)$ we have that

$$|T_\xi^{-n} V| \leq (\delta_1 \alpha^2)^{-n} |V| \leq (\delta_1 \alpha^2)^{-n} (\alpha^{-1}/8) \leq (\delta_1 \alpha^2)^{-n} |W|. \tag{53}$$

Combining (52) and (53) yields

$$\frac{|T_\xi^{-n} I_n(\xi, W)|}{|W|} \leq \frac{4^n}{(\delta_1 \alpha^2)^n}. \qquad \square$$

Lemma 3.3 was a simple consequence of how the expansion rate $\delta_1 \alpha^2$ dominates the number of singularity curves that a sufficiently small line segment can cross. We now use Lemma 3.3 applied at each iteration to obtain bounds on $|T_\xi^{-n} S_n(\xi, W)|$ and $|T_\xi^{-n} L_n(\xi, W)|$, which is ultimately what is required to prove Lemma 3.2.

LEMMA 3.4

If α is sufficiently large, then for any $\xi \in \Omega$, $W \in \mathcal{W}_u$ with $|W| \leq \alpha^{-1}/4$, and $n \geq \log(1/|W|)$ there holds

$$\left|T_\xi^{-n} L_n(\xi, W)\right| \geq \left(1 - \frac{1}{\alpha}\right)|W|.$$

Proof

Fix $\xi \in \Omega$ and $W \in \mathcal{W}_u$. Since

$$\left|T_\xi^{-n} L_n(\xi, W)\right| + \left|T_\xi^{-n} S_n(\xi, W)\right| = |W|,$$

the lemma is equivalent to proving that

$$\left|T_\xi^{-n} S_n(\xi, W)\right| \leq \frac{1}{\alpha}|W|. \tag{54}$$

To estimate the left-hand side of (54), following [4] and [34], we group the elements of $S_n(\xi, W)$ in terms of their most recent long ancestor. In particular, given any $W_i \in S_n(\xi, W)$, either $W_i \in I_n(\xi, W)$ or there is a unique $k \in \{1, \dots, n-1\}$ such that $W_i \in I_k(\theta^{n-k}\xi, V)$ for some $V \in L_{n-k}(\xi, W)$. Thus, we can write

$$\begin{aligned}\frac{|T_\xi^{-n} S_n(\xi, W)|}{|W|} = &\frac{|T_\xi^{-n} I_n(\xi, W)|}{|W|} \\ &+ \frac{1}{|W|} \sum_{k=1}^{n-1} \sum_{V_i \in L_{n-k}(\xi, W)} \left|T_\xi^{-n} I_k(\theta^{n-k}\xi, V_i)\right|. \end{aligned} \tag{55}$$

For the first term, observe that by the proof of Lemma 3.3 we have

$$\frac{|T_\xi^{-n} I_n(\xi, W)|}{|W|} \le \Big(\frac{4}{\alpha^2 \delta_1}\Big)^n \frac{1}{4\alpha |W|}.$$

If α is large enough, then the choice of n implies that

$$\frac{|T_\xi^{-n} I_n(\xi, W)|}{|W|} \le \frac{1}{2\alpha}.$$

For the second term in (55), we first use the trivial fact that

$$|W| \ge \sum_{V_i \in L_{n-k}(\xi, W)} |T_\xi^{-(n-k)} V_i|$$

for every k to obtain

$$\begin{aligned} \frac{1}{|W|} \sum_{k=1}^{n-1} \sum_{V_i \in L_{n-k}(\xi, W)} \big|T_\xi^{-n} I_k(\theta^{n-k}\xi, V_i)\big| \\ \le \sum_{k=1}^{n-1} \frac{\sum_{V_i \in L_{n-k}(\xi,W)} |T_\xi^{-n} I_k(\theta^{n-k}\xi, V_i)|}{\sum_{V_i \in L_{n-k}(\xi,W)} |T_\xi^{-(n-k)} V_i|}. \end{aligned}$$

Using also the elementary inequality

$$\frac{a_1 + a_2 + \cdots + a_N}{b_1 + b_2 + \cdots + b_N} \le \max_{1 \le j \le N} \frac{a_j}{b_j}$$

for two lists of positive numbers $\{a_i\}_{i=1}^N$ and $\{b_i\}_{i=1}^N$, we deduce

$$\frac{1}{|W|} \sum_{k=1}^{n-1} \sum_{V_i \in L_{n-k}(\xi, W)} \big|T_\xi^{-n} I_k(\theta^{n-k}\xi, V_i)\big| \le \sum_{k=1}^{n-1} \max_{V_i \in L_{n-k}(\xi,W)} \frac{|T_\xi^{-n} I_k(\theta^{n-k}\xi, V_i)|}{|T_\xi^{-(n-k)} V_i|}.$$

We will now rewrite the sum above so that we are able to apply Lemma 3.3. Fix $1 \le k \le n-1$ and $V_i \in L_{n-k}(\xi, W)$. Let $W_i \in I_k(\theta^{n-k}\xi, V_i)$, and observe that by the definition of T_ξ^n and θ we have

$$T_\xi^{-n} W_i = T_\xi^{-(n-k)}(T_{\theta^{n-k}\xi}^{-k} W_i).$$

Since $T_{\theta^{n-k}\xi}^{-k} W_i \subseteq V_i$ and $\nabla T_\xi^{-(n-k)}$ is constant on V_i, it follows that

$$\frac{|T_\xi^{-n} W_i|}{|T_\xi^{-(n-k)} V_i|} = \frac{|T_\xi^{-(n-k)}(T_{\theta^{n-k}\xi}^{-k} W_i)|}{|T_\xi^{-(n-k)} V_i|} = \frac{|T_{\theta^{n-k}\xi}^{-k} W_i|}{|V_i|}.$$

Summing this estimate over $W_i \in I_k(\theta^{n-k}\xi, V_i)$ and assuming that α is large enough so that Lemma 3.3 holds with $\epsilon = 1/2$, we obtain

$$\max_{V_i \in L_{n-k}(\xi,W)} \frac{|T_\xi^{-n} I_k(\theta^{n-k}\xi, V_i)|}{|T_\xi^{-(n-k)} V_i|} \le \max_{V_i \in L_{n-k}(\xi,W)} \frac{|T_{\theta^{n-k}\xi}^{-k} I_k(\theta^{n-k}\xi, V_i)|}{|V_i|}$$
$$\le \alpha^{-3(n-k)/2}.$$

Combining our estimates thus far and assuming that α is sufficiently large, we have

$$\frac{|T_\xi^{-n} S_n(\xi, W)|}{|W|} \le \frac{1}{2\alpha} + \sum_{k=1}^{n-1} \alpha^{-3(n-k)/2} \le \frac{1}{2\alpha} + \frac{1}{\alpha^{3/2} - 1} \le \frac{1}{\alpha},$$

which completes the proof. □

We are now ready to give the proof of Lemma 3.2. Note that the difference between Lemma 3.2 and Lemma 3.4 is that the "long lines" of Lemma 3.2 have length $5/4$, while those in Lemma 3.4 have length on the order of α^{-1}. The remedy is just to modify the splitting given in the definition of the generations $\mathscr{G}_n$ at the last step. Indeed, the plan is to first apply Lemma 3.4 to obtain a set $\mathscr{G}_{n-1}(\xi, W)$ that consists primarily of lines with length approximately α^{-1}, and then to consider the elements of $T_{\xi_n}(\mathscr{G}_{n-1}(\xi, W))$. The idea here is that since the expansion rate scales like α^2 and $\alpha^2\alpha^{-1} = \alpha \gg 1$, most line segments in $T_{\xi_n}(\mathscr{G}_{n-1}(\xi, W))$ necessarily span across $\mathbb{T}^2$.

Proof of Lemma 3.2

Fix $\xi \in \Omega$ and $W \in \mathcal{W}_u$. It is clear that by subdividing W we may assume without loss of generality that $|W| \le \alpha^{-1}/4$. For $z \in \mathbb{R}^2$ and $V \in \mathcal{W}_u$, let $\tilde{\mathscr{G}}(z, V)$ be defined in the same way as $\mathscr{G}(z, V)$ before but with the segments of length strictly greater than $5/2$ subdivided into segments with length strictly between $5/4$ and $5/2$. We define the collection $\tilde{\mathscr{G}}_n(\xi, W) \subseteq \mathcal{W}_u$ by

$$\tilde{\mathscr{G}}_n(\xi, W) = \bigcup_{W_i \in \mathscr{G}_{n-1}(\xi,W)} \tilde{\mathscr{G}}(\xi_n, W_i).$$

Except for (51), $\tilde{\mathscr{G}}_n(\xi, W)$ is easily seen by construction to satisfy all of the properties claimed in Lemma 3.2. It thus remains to prove, writing $\tilde{\mathscr{G}}_n(\xi, W) = \{W_i\}_{i\in I}$, that

$$\frac{1}{|W|} \sum_{|W_i| \ge 5/4} |T_\xi^{-n} W_i| \ge 1 - \frac{C}{\alpha}$$

for some constant $C \ge 1$ that does not depend on α or ξ. First, since $n - 1 \ge \log \frac{1}{|W|}$, by Lemma 3.4 we have

$$\frac{1}{|W|}\sum_{|W_i|\geq 5/4}|T_\xi^{-n}W_i| \geq \Big(1-\frac{1}{\alpha}\Big)\frac{1}{|T_\xi^{-(n-1)}L_{n-1}(\xi,W)|}\sum_{|W_i|\geq 5/4}|T_\xi^{-n}W_i|.$$

To simplify notation, for $V_k \in L_{n-1}(\xi, W)$, define

$$E_k = \big\{W_i \in \tilde{\mathscr{G}}_n(\xi, W) : T_{\xi_n}^{-1}W_i \subseteq V_k, |W_i| \geq 5/4\big\}.$$

Arguing similar to the proof of Lemma 3.4, we deduce

$$\begin{aligned}\frac{1}{|T_\xi^{-(n-1)}L_{n-1}(\xi,W)|}\sum_{|W_i|\geq 5/4}|T_\xi^{-n}W_i| &\geq \frac{\sum_{V_k\in L_{n-1}(\xi,W)}\sum_{W_i\in E_k}|T_\xi^{-n}W_i|}{\sum_{V_k\in L_{n-1}(\xi,W)}|T_\xi^{-(n-1)}V_k|}\\ &\geq \min_{V_k\in L_{n-1}(\xi,W)}\frac{\sum_{W_i\in E_k}|T_\xi^{-n}W_i|}{|T_\xi^{-(n-1)}V_k|}\\ &= \min_{V_k\in L_{n-1}(\xi,W)}\frac{\sum_{W_i\in E_k}|T_{\xi_n}^{-1}W_i|}{|V_k|}.\end{aligned}$$

Using that $V_k \in L_{n-1}(\xi, W)$ can cross at most three singularity curves and the lower bound on the expansion rate given by Lemma 3.1, it is not hard to show that the construction of $\tilde{\mathscr{G}}(\xi_n, V_k)$ implies that

$$\frac{\sum_{W_i\in E_k}|T_{\xi_n}^{-1}W_i|}{|V_k|} \geq \frac{|V_k| - \frac{15}{2\delta_1\alpha^2}}{|V_k|} \geq 1 - \frac{60}{\delta_1\alpha}.$$

Combining the estimates above, we obtain

$$\frac{1}{|W|}\sum_{|W_i|\geq 5/4}|T_\xi^{-n}W_i| \geq \Big(1-\frac{1}{\alpha}\Big)\Big(1-\frac{60}{\delta_1\alpha}\Big),$$

which completes the proof. ☐

4. Proof of the geometric mixing lemma

In this section, we use the results of Section 3 to prove Lemma 2.1. As discussed earlier, we make use of the area formula from geometric measure theory, which we recall below.

LEMMA 4.1 (Area formula)
Let M be an m-dimensional C^1 Riemannian manifold, and let $\mathcal{H}^m$ be the m-dimensional Hausdorff measure on M induced by the Riemannian metric. If $\mathcal{H}^m(M) < \infty$, $F : M \to M$ is a Lipschitz continuous map, and $\varphi : M \to \mathbb{R}$ is measurable, then

$$\int_M J_F(z)\varphi(z)\mathcal{H}^m(\mathrm{d}z) = \int_M \Big(\sum_{z \in F^{-1}(z')} \varphi(z) \Big) \mathcal{H}^m(\mathrm{d}z'),$$

where J_F denotes the Jacobian of F calculated using the Riemannian metric.

Remark 4.1
The proof of the area formula can be found in [42, Theorem 3.2.3 or Theorem 3.2.5] for an m-dimensional subset of $\mathbb{R}^n$. As discussed in [42, Section 3.2.49], by localizing to charts which map $\mathbb{R}^m$ into M we can extend the $\mathbb{R}^m$ result to the smooth manifold setting. Additional discussion can be found in [15], [41], [63], and [64].

Remark 4.2
Since the function F in the area formula is Lipschitz, Rademacher's theorem (cf. [42, Theorem 3.1.6]) tells us that it is differentiable almost everywhere. Thus, the Jacobian J_F is defined almost everywhere by $J_F(z) = |\nabla F(z)|$, which is sufficient to define the integral in the area formula. If in addition $J_F(z) > 0$ for almost every z, then

$$\int_M \varphi(z)\mathcal{H}^m(\mathrm{d}z) = \int_M \Big(\sum_{z \in F^{-1}(z')} \frac{\varphi(z)}{J_F(z)} \Big) \mathcal{H}^m(\mathrm{d}z'). \tag{56}$$

This can be shown by subdividing $A = \{z \in M : J_F(z) > 0\}$ into the disjoint sets $A_n = \{z \in A : J_F(z) \in [\frac{1}{n+1}, \frac{1}{n}) \cup [n, n+1)\}$ with $n \in \{1, 2 \ldots\}$. On each A_n the area formula implies (56) when applied to the function φ/J_F. Adding up the integrals over each A_n implies the general result assuming that the left-hand side of (56) is well defined. In our particular setting, J_F will be uniformly bounded from above and below by positive constants, which makes (56) immediate from the area formula.

By making an appropriate choice of F, we apply Lemma 4.1 to express $m(T_\xi^{2n}(R) \cap Q)$ as an integral that amounts essentially to counting intersection points (weighted by an appropriate Jacobian factor) between horizontal lines mapped through T_ξ^n and vertical lines mapped through $T_{\theta^n \xi}^{-n}$.

LEMMA 4.2
Fix $n, N \in \mathbb{N}$, $\xi \in \Omega$, and $Q, R \in \mathcal{R}_N$. Define the map $F_{n,\xi} : \mathbb{T}^2 \to \mathbb{T}^2$ by

$$F_{n,\xi} = (\Pi_x \circ T_{\theta^n \xi}^n, \Pi_y \circ T_\xi^{-n}), \tag{57}$$

where $\Pi_x, \Pi_y : \mathbb{T}^2 \to \mathbb{T}$ are the projections $\Pi_x(x, y) = x$ and $\Pi_y(x, y) = y$. Let $W_{y,u} = \Pi_x(R) \times \{y\}$ and $W_{x,s} = \{x\} \times \Pi_y(Q)$. Then,

$$m\big(T_\xi^{2n}(R) \cap Q\big) = \int_{\Pi_x(Q)} \int_{\Pi_y(R)} \Big(\sum_{z \in T_\xi^n(W_{y',u}) \cap T_{\theta^n \xi}^{-n}(W_{x',s})} \frac{1}{J_{F_{n,\xi}}(z)} \Big) \mathrm{d}y'\, \mathrm{d}x'.$$

Proof

Since T is area-preserving, we have

$$m\big(T_\xi^{2n}(R)\cap Q\big)=m\big(T_\xi^{n}(R)\cap T_{\theta^n\xi}^{-n}(Q)\big)=\int_{\mathbb{T}^2}\chi_R\big(T_\xi^{-n}(z)\big)\chi_Q\big(T_{\theta^n\xi}^{n}(z)\big)\,\mathrm{d}z.$$

Thus, applying Lemma 4.1 with $F_{n,\xi}$ as defined above, we can write

$$\begin{aligned}
&m\big(T_\xi^{2n}(R)\cap Q\big)\\
&\quad=\int_{\mathbb{T}^2}\Big(\sum_{z\in F_{n,\xi}^{-1}(z')}\frac{\chi_R(T_\xi^{-n}(z))\chi_Q(T_{\theta^n\xi}^{n}(z))}{J_{F_{n,\xi}}(z)}\Big)\,\mathrm{d}z'\\
&\quad=\int_{\Pi_x(Q)}\int_{\Pi_y(R)}\Big(\sum_{z\in F_{n,\xi}^{-1}(x',y')}\frac{\chi_{\Pi_x(R)}(\Pi_x\circ T_\xi^{-n}(z))\chi_{\Pi_y(Q)}(\Pi_y\circ T_{\theta^n\xi}(z))}{J_{F_{n,\xi}}(z)}\Big)\\
&\qquad\times\mathrm{d}y'\,\mathrm{d}x'.
\end{aligned}$$

In the second line above we have observed that the definition of $F_{n,\xi}$ implies that if $z'=(x',y')\notin\Pi_x(Q)\times\Pi_y(R)$, then at least one of $\chi_R(T_\xi^{-n}(z))$ or $\chi_Q(T_{\theta^n\xi}^{n}(z))$ vanishes for every $z\in F_{n,\xi}^{-1}(z')$. Moreover, for $(x',y')\in\Pi_x(Q)\times\Pi_y(R)$, the product of characteristic functions in the second line above is nonzero if and only if $z\in T_\xi^n(W_{y',u})\cap T_{\theta^n\xi}^{-n}(W_{x',s})$. The lemma follows. ☐

To estimate the sum in Lemma 4.2 we use the decompositions of $T_\xi^n(W_{y',u})$ and $T_{\theta^n\xi}^{-n}(W_{x',s})$ guaranteed by Lemma 3.2 and Remark 3.3. The idea is essentially that for $(W,V)\in T_\xi^n(W_{y',u})\times T_{\theta^n\xi}^{-n}(W_{x',s})$ with $|W|,|V|\geq 5/4$, we must have $\#(W\cap V)\geq 1$ for α large since the angle between vectors $v_u\in C_u$ and $v_s\in C_s$ converges to $\pm\pi/2$ as $\alpha\to\infty$. To properly employ (51) and obtain the correct quantitative estimate we use Lemma A.1, which relates $J_{F_{n,\xi}}(z)$ to the product of the arc-length Jacobians along $W_{y',u}$ and $W_{x',s}$.

LEMMA 4.3

For any $N\in\mathbb{N}$ and α sufficiently large, there exists $c>0$ so that for all $\xi\in\Omega$, horizontal and vertical segments $(W_u,W_s)\in\mathcal{W}_u\times\mathcal{W}_s$ of length 2^{-N}, and $n\geq 1+N\log(2)+\log(8\alpha)$ we have

$$\sum_{z\in T_\xi^n(W_u)\cap T_{\theta^n\xi}^{-n}(W_s)}\frac{1}{J_{F_{n,\xi}}(z)}\geq c|W_u||W_s|,\tag{58}$$

where $F_{n,\xi}$ is as before.

Remark 4.3

The upper bound mentioned in Remark 2.1 is obtained by proving an upper bound analogous to (58).

Proof

Let $G_n^u(W_u) \subseteq \mathcal{W}_u$ and $G_n^s(W_s) \subseteq \mathcal{W}_s$ denote the decompositions of $T_\xi^n(W_u)$ and $T_{\theta^n\xi}^{-n}(W_s)$ guaranteed by Lemma 3.2 and Remark 3.3. Define also $L_n(W_u)$ as the elements $W \in G_n^u(W_u)$ with $|W| \geq 5/4$ and similarly define $L_n(W_s)$. Clearly, we have

$$\sum_{z \in T_\xi^n(W_u) \cap T_{\theta^n\xi}^{-n}(W_s)} \frac{1}{J_{F_{n,\xi}}(z)} = \sum_{W \in G_u^n(W_u)} \sum_{V \in G_s^n(W_s)} \sum_{z \in W \cap V} \frac{1}{J_{F_{n,\xi}}(z)}.$$

Let $J_{u,n}(z) = |\nabla T_\xi^n(z) e_x|$ and $J_{s,n}(z) = |\nabla T_{\theta^n\xi}^{-n}(z) e_y|$, where e_x and e_y are the standard basis vectors of $\mathbb{R}^2$. By Lemma A.1 and trivially bounding the sum from below by excluding the short lines, we have

$$\begin{aligned}
&\sum_{W \in G_n^u(W_u)} \sum_{V \in G_n^s(W_s)} \sum_{z \in W \cap V} \frac{1}{J_{F_{n,\xi}}(z)} \\
&\quad \geq (1 - C\alpha^{-1}) \sum_{W \in L_n(W_u)} \sum_{V \in L_n(W_s)} \sum_{z \in W \cap V} \frac{1}{J_{u,n}(T_\xi^{-n}(z)) J_{s,n}(T_{\theta^n\xi}^n(z))} \\
&\quad = (1 - C\alpha^{-1}) \sum_{W \in L_n(W_u)} \sum_{V \in L_n(W_s)} \#(W \cap V) \times \frac{|T_\xi^{-n} W| |T_{\theta^n\xi}^n V|}{|W||V|},
\end{aligned}$$

where in the last equality we used that

$$J_{u,n}\big(T_\xi^{-n}(z)\big) |T_\xi^{-n} W| = |W| \qquad \text{and} \qquad J_{s,n}\big(T_{\theta^n\xi}^n(z)\big) |T_{\theta^n\xi}^n V| = |V|$$

for all $z \in W \cap V$, since ∇T_ξ^n is constant on $T_\xi^{-n} W$, which is a horizontal line (and similarly for the second equality). For W and V as in the last sum above we have $\#(W \cap V) \geq 1$ and $|W|, |V| \leq 5/2$, and hence by (51) we have

$$\begin{aligned}
&\sum_{W \in G_n^u(W_u)} \sum_{V \in G_n^s(W_s)} \sum_{z \in W \cap V} \frac{1}{J_{F_{n,\xi}}(z)} \\
&\quad \geq \frac{4}{25}(1 - C\alpha^{-1}) \sum_{W \in L_n(W_u)} |T_\xi^{-n} W| \sum_{V \in L_n(W_s)} |T_{\theta^n\xi}^n V| \\
&\quad \geq \frac{4}{25}(1 - C\alpha^{-1})^3 |W_u||W_s|,
\end{aligned}$$

which completes the proof. ☐

The proof of Lemma 2.1 now follows easily.

Proof of Lemma 2.1

Fix $N \in \mathbb{N}$. Let α be large enough so that Lemma 4.3 applies, and let $c > 0$ be as in (58). Taking $n \geq 1 + N\log(2) + \log(8\alpha)$, it follows by Lemmas 4.2 and 4.3 that for any $R, Q \in \mathcal{R}_N$, and $\xi \in \Omega$ we have

$$m\big(T_\xi^{2n}(R) \cap Q\big) \geq c \int_{\Pi_x(Q)} \int_{\Pi_y(R)} |W_{y',u}||W_{x',s}| \,\mathrm{d}y' \,\mathrm{d}x = c2^{-4N} = cm(R)m(Q),$$

which is the desired estimate. □

Appendix. Jacobian estimates

We now prove the technical Jacobian estimate needed in the proof of Lemma 2.1.

LEMMA A.1

Fix $\xi \in \Omega$ and $n \in \mathbb{N}$. Define $J_{u,n}(z) = |\nabla T_\xi^n(z)e_x|$ and $J_{s,n}(z) = |\nabla T_{\theta^n\xi}^{-n}(z)e_y|$, and let $F_{n,\xi}$ be as defined in (57). There exists a constant $C \geq 1$ that does not depend on ξ or n such that for all α sufficiently large there holds

$$\left|\frac{J_{u,n}(T_\xi^{-n}(z))J_{s,n}(T_{\theta^n\xi}^n(z))}{J_{F_{n,\xi}}(z)} - 1\right| + \left|\frac{J_{F_{n,\xi}}(z)}{J_{u,n}(T_\xi^{-n}(z))J_{s,n}(T_{\theta^n\xi}^n(z))} - 1\right| \leq \frac{C}{\alpha} \tag{59}$$

for every $z \in \mathbb{T}^2$ in the full measure set where the derivatives in (59) exist.

Proof

By the elementary inequality

$$\left|\frac{a}{b} - 1\right| \leq \frac{|\frac{b}{a} - 1|}{1 - |\frac{b}{a} - 1|} \tag{60}$$

for nonzero $a, b \in \mathbb{R}$, it suffices to estimate only the first term on the left-hand side of (59). We will also make use of the related inequality

$$\left|\frac{c}{d} - 1\right| \leq \left|\frac{c}{a} - 1\right|\left(1 + \left|\frac{a}{d} - 1\right|\right) + \left|\frac{a}{d} - 1\right|, \tag{61}$$

which holds for any $a, b, c, d \in \mathbb{R}$ with a and d nonzero. In this proof, C will denote a positive constant, which may change from line to line, that does not depend on ξ, n, or α for α sufficiently large.

Let $z \in \mathbb{T}^2$ be such that all of the derivatives in (59) exist. We first claim that

$$\frac{|\partial_y(\Pi_x \circ T^n_{\theta^n\xi})(z)|}{|\partial_x(\Pi_x \circ T^n_{\theta^n\xi})(z)|} \le \frac{C}{\alpha} \tag{62}$$

for α sufficiently large. Indeed, let e_s and e_u denote the normalized stable and unstable eigenvectors for $\nabla T^n_{\theta^n\xi}(z)$, and let $\lambda \in \mathbb{R}$ be the unstable eigenvalue. If $\theta \in [0,\pi)$ denotes the angle between e_x and e_s, then since $e_s \in C_s$ we have $|\theta - \pi/2| \le C\alpha^{-1}$. Therefore,

$$|e_x \cdot e_s| = |\cos\theta| = \left|\sin(\theta - \pi/2)\right| \le C\alpha^{-1}$$

for α large. Using also that $e_u \in C_u$, we similarly have

$$|e_y \cdot e_u| \le C\alpha^{-1}, \qquad |e_x \cdot e_u| \ge 1 - C^{-1}\alpha^{-1}, \qquad \text{and} \qquad |e_y \cdot e_s| \ge 1 - C^{-1}\alpha^{-1}.$$

For α large, we can thus write

$$e_x = a_1 e_u + b_1 e_s \qquad \text{and} \qquad e_y = a_2 e_u + b_2 e_s$$

for coefficients satisfying $|a_1|, |b_2| \approx 1$ and $|a_2|, |b_1| \le C\alpha^{-1}$. Noting also that $|e_u \cdot e_s| \le C\alpha^{-1}$ we have

$$\begin{aligned}
\frac{|\partial_y(\Pi_x \circ T^n_{\theta^n\xi})(z)|}{|\partial_x(\Pi_x \circ T^n_{\theta^n\xi})(z)|} &= \frac{|e_x \cdot \nabla T^n_{\theta^n\xi}(z) e_y|}{|e_x \cdot \nabla T^n_{\theta^n\xi}(z) e_x|} \\
&= \frac{|a_1 a_2 \lambda + b_1 b_2 \lambda^{-1} + (e_u \cdot e_s)(a_1 b_2 \lambda^{-1} + a_2 b_1 \lambda)|}{|a_1^2 \lambda + b_1^2 \lambda^{-1} + (e_u \cdot e_s) a_1 b_1 (\lambda + \lambda^{-1})|} \\
&\le \frac{C\alpha^{-1}|\lambda|}{a_1^2|\lambda| - b_1^2|\lambda|^{-1} - 2|\lambda||a_1 b_1||e_u \cdot e_s|} \\
&\le \frac{C\alpha^{-1}|\lambda|}{a_1^2|\lambda| - C\alpha^{-1}|\lambda|} \\
&\le \frac{C}{\alpha}.
\end{aligned}$$

This proves (62). The same argument shows that

$$\begin{gathered}
\frac{|\partial_x(\Pi_y \circ T^{-n}_{\xi})(z)|}{|\partial_y(\Pi_y \circ T^{-n}_{\xi})(z)|} \le \frac{C}{\alpha}, \qquad \frac{|\partial_x(\Pi_y \circ T^n_{\theta^n\xi})(z)|}{|\partial_x(\Pi_x \circ T^n_{\theta^n\xi})(z)|} \le \frac{C}{\alpha}, \\
\text{and} \qquad \frac{|\partial_y(\Pi_x \circ T^{-n}_{\xi})(z)|}{|\partial_y(\Pi_y \circ T^{-n}_{\xi})(z)|} \le \frac{C}{\alpha}.
\end{gathered} \tag{63}$$

Combining (62) and the first estimate in (63), it is immediate from the formula for $J_{F_{n,\xi}}(z)$ that

$$\left|\frac{J_{F_{n,\xi}}(z)}{|\partial_x(\Pi_x \circ T^n_{\theta^n\xi})(z)||\partial_y(\Pi_y \circ T^{-n}_\xi)(z)|} - 1\right| \le \frac{C}{\alpha}. \tag{64}$$

It is also straightforward to prove from the second two bounds in (63) that

$$\left|\frac{|\nabla T^n_{\theta^n\xi}(z)e_x||\nabla T^{-n}_\xi(z)e_y|}{|\partial_x(\Pi_x \circ T^n_{\theta^n\xi})(z)||\partial_y(\Pi_y \circ T^{-n}_\xi)(z)|} - 1\right| \le \frac{C}{\alpha}. \tag{65}$$

Together, (60), (61), (64), and (65) imply

$$\left|\frac{J_{F_{n,\xi}}(z)}{|\nabla T^n_{\theta^n\xi}(z)e_x||\nabla T^{-n}_\xi(z)e_y|} - 1\right| \le \frac{C}{\alpha}. \tag{66}$$

From (60), (61), and (66), it remains only to show that

$$\left|\frac{J_{u,n}(T^{-n}_\xi(z))J_{s,n}(T^n_{\theta^n\xi}(z))}{|\nabla T^n_{\theta^n\xi}(z)e_x||\nabla T^{-n}_\xi(z)e_y|} - 1\right| \le \frac{C}{\alpha}. \tag{67}$$

We will just prove that

$$\left|\frac{J_{u,n}(T^{-n}_\xi(z))}{|\nabla T^{-n}_\xi(z)e_y|} - 1\right| \le \frac{C}{\alpha}, \tag{68}$$

as the same inequality for the other ratio follows similarly, and combining them to deduce (67) is straightforward. Define $\bar{z} = T^{-n}_\xi(z)$, and let $\bar{e}_s$ and $\bar{e}_u$ denote the stable and unstable eigenvectors of $\nabla T^n_\xi(\bar{z})$. Reasoning similar to the proof of (62) above and noting also that $\bar{e}_s$ is the unstable eigenvector for $\nabla T^{-n}_\xi(z)$ shows that

$$\left|\frac{J_{u,n}(\bar{z})}{|\nabla T^n_\xi(\bar{z})\bar{e}_u|} - 1\right| + \left|\frac{|\nabla T^{-n}_\xi(z)\bar{e}_s|}{|\nabla T^{-n}_\xi(z)e_y|} - 1\right| \le \frac{C}{\alpha}. \tag{69}$$

Now, since T^n_ξ is area-preserving, we have

$$\left|\nabla T^n_\xi(\bar{z})\bar{e}_u\right| = \left|\nabla T^n_\xi(\bar{z})\bar{e}_s\right|^{-1} = \left|\nabla T^{-n}_\xi(z)\bar{e}_s\right|. \tag{70}$$

Combining (69) with (70) and then applying (61) yields (68). ☐

Acknowledgments. The first author thanks Craig Chen for very helpful computations and numerics related to this problem in the 2021–2022 academic year. All of the authors thank Gautam Iyer for useful feedback on an early draft of the paper and informative and illuminating discussions around this topic over the years. Lastly, we thank each of the anonymous referees for their comments and insights. Their perspectives were useful and led to improved exposition and proofs.

The authors thank the National Science Foundation (NSF) for its support through Research Training Group (RTG) grant DMS-2038056. Elgindi's work was partially supported by NSF grants DMS-2043024 and DMS-212474 as well as an Alfred P. Sloan Fellowship.

References

[1] G. ALBERTI, G. CRIPPA, and A. L. MAZZUCATO, *Exponential self-similar mixing by incompressible flows*, J. Amer. Math. Soc. **32** (2019), no. 2, 445–490. MR 3904158. DOI 10.1090/jams/913. *(1215, 1222, 1235)*

[2] D. ALBRITTON, R. BEEKIE, and M. NOVACK, *Enhanced dissipation and Hörmander's hypoellipticity*, J. Funct. Anal. **283** (2022), no. 3, art. ID 109522. MR 4413303. DOI 10.1016/j.jfa.2022.109522. *(1211, 1215)*

[3] S. ARMSTRONG and V. VICOL, *Anomalous diffusion by fractal homogenization*, Ann. PDE **11** (2025), no. 1, art. ID 2. MR 4846255. DOI 10.1007/s40818-024-00189-6. *(1212)*

[4] V. BALADI and M. F. DEMERS, *On the measure of maximal entropy for finite horizon Sinai billiard maps*, J. Amer. Math. Soc. **33** (2020), no. 2, 381–449. MR 4073865. DOI 10.1090/jams/939. *(1220, 1242, 1245)*

[5] V. BALADI, M. F. DEMERS, and C. LIVERANI, *Exponential decay of correlations for finite horizon Sinai billiard flows*, Invent. Math. **211** (2018), no. 1, 39–177. MR 3742756. DOI 10.1007/s00222-017-0745-1. *(1215, 1222)*

[6] L. BARREIRA and Y. PESIN, *Introduction to Smooth Ergodic Theory*, Grad. Stud. Math. **148**, Amer. Math. Soc., Providence, 2013. MR 3076414. DOI 10.1090/gsm/148. *(1222)*

[7] J. BEDROSSIAN, A. BLUMENTHAL, and S. PUNSHON-SMITH, *Almost-sure enhanced dissipation and uniform-in-diffusivity exponential mixing for advection-diffusion by stochastic Navier-Stokes*, Probab. Theory Related Fields **179** (2021), no. 3–4, 777–834. MR 4242626. DOI 10.1007/s00440-020-01010-8. *(1211, 1214, 1216)*

[8] ———, *Almost-sure exponential mixing of passive scalars by the stochastic Navier-Stokes equations*, Ann. Probab. **50** (2022), no. 1, 241–303. MR 4385127. DOI 10.1214/21-aop1533. *(1216)*

[9] ———, *Lagrangian chaos and scalar advection in stochastic fluid mechanics*, J. Eur. Math. Soc. (JEMS) **24** (2022), no. 6, 1893–1990. MR 4404792. DOI 10.4171/jems/1140. *(1216)*

[10] J. BEDROSSIAN and M. COTI ZELATI, *Enhanced dissipation, hypoellipticity, and anomalous small noise inviscid limits in shear flows*, Arch. Ration. Mech. Anal. **224** (2017), no. 3, 1161–1204. MR 3621820. DOI 10.1007/s00205-017-1099-y. *(1211, 1215)*

[11] J. BEDROSSIAN, P. GERMAIN, and N. MASMOUDI, *On the stability threshold for the 3D Couette flow in Sobolev regularity*, Ann. of Math. (2) **185** (2017), no. 2, 541–608. MR 3612004. DOI 10.4007/annals.2017.185.2.4. *(1215)*

[12] ———, *Dynamics near the subcritical transition of the 3D Couette flow, I: Below threshold case*, Mem. Amer. Math. Soc. **266** (2020), no. 1294. MR 4126259. DOI 10.1090/memo/1294. *(1215)*

[13] A. BLUMENTHAL, M. COTI ZELATI, and R. S. GVALANI, *Exponential mixing for random dynamical systems and an example of Pierrehumbert*, Ann. Probab. **51** (2023), no. 4, 1559–1601. MR 4597327. DOI 10.1214/23-aop1627. *(1216, 1217)*

[14] A. BRESSAN, *A lemma and a conjecture on the cost of rearrangements*, Rend. Semin. Mat. Univ. Padova **110** (2003), 97–102. MR 2033002. *(1215)*

[15] J. E. BROTHERS, *Integral geometry in homogeneous spaces*, Trans. Amer. Math. Soc. **124** (1966), no. 3, 480–517. MR 0202099. DOI 10.2307/1994389. *(1249)*

[16] E. BRUÉ and Q.-H. NGUYEN, *Advection diffusion equations with Sobolev velocity field*, Comm. Math. Phys. **383** (2021), no. 1, 465–487. MR 4236071. DOI 10.1007/s00220-021-03993-4. *(1212)*

[17] E. BRUÉ, M. COTI ZELATI, and E. MARCONI, *Enhanced dissipation for two-dimensional Hamiltonian flows*, Arch. Ration. Mech. Anal. **248** (2024), no. 5, art. ID 84. MR 4797689. DOI 10.1007/s00205-024-02034-3. *(1215)*

[18] L. A. BUNIMOVICH, Y. G. SINAI, and N. I. CHERNOV, *Statistical properties of two-dimensional hyperbolic billiards* (in Russian), Uspekhi Mat. Nauk **46** (1991), no. 4(280), 43–92; English translation in Russian Math. Surveys **46** (1991), no. 4, 47–106. MR 1138952. DOI 10.1070/RM1991v046n04ABEH002827. *(1215)*

[19] J. BURCZAK, L. SZÉKELYHIDI, JR., and B. WU, *Anomalous dissipation and Euler flows*, preprint, arXiv:2310.02934v2 [math.AP]. *(1212)*

[20] L.-T. CHENG, F. RAJASEKARAN, K. Y. J. WONG, and A. ZLATOŠ, *Numerical evidence of exponential mixing by alternating shear flows*, Commun. Math. Sci. **21** (2023), no. 2, 529–541. MR 4545838. DOI 10.4310/cms.2023.v21.n2.a10. *(1211, 1214, 1217)*

[21] N. CHERNOV, *Ergodic and statistical properties of piecewise linear hyperbolic automorphisms of the* 2*-torus*, J. Stat. Phys. **69** (1992), no. 1–2, 111–134. MR 1184772. DOI 10.1007/BF01053785. *(1215)*

[22] ———, *Decay of correlations and dispersing billiards*, J. Stat. Phys. **94** (1999), no. 3–4, 513–556. MR 1675363. DOI 10.1023/A:1004581304939. *(1215)*

[23] N. CHERNOV and H.-K. ZHANG, *Billiards with polynomial mixing rates*, Nonlinearity **18** (2005), no. 4, 1527–1553. MR 2150341. DOI 10.1088/0951-7715/18/4/006. *(1215, 1217)*

[24] M. COLOMBO, M. COTI ZELATI, and K. WIDMAYER, *Mixing and diffusion for rough shear flows*, Ars Inven. Anal. **2021**, no. 2. MR 4462470. DOI 10.15781/83fc-j334. *(1215)*

[25] M. COLOMBO, G. CRIPPA, and M. SORELLA, *Anomalous dissipation and lack of selection in the Obukhov-Corrsin theory of scalar turbulence*, Ann PDE **9** (2023), no. 2, art. ID 21. MR 4662772. DOI 10.1007/s40818-023-00162-9. *(1212)*

[26] P. CONSTANTIN, A. KISELEV, L. RYZHIK, and A. ZLATOŠ, *Diffusion and mixing in fluid flow*, Ann. of Math. (2) **168** (2008), no. 2, 643–674. MR 2434887. DOI 10.4007/annals.2008.168.643. *(1215, 1216)*

[27] W. COOPERMAN, *Exponential mixing by shear flows*, SIAM J. Math. Anal. **55** (2023), no. 6, 7513–7528. MR 4665036. DOI 10.1137/22M1513861. *(1215, 1216, 1217)*

[28] W. COOPERMAN, G. IYER, and S. SON, *A Harris theorem for enhanced dissipation, and an example of Pierrehumbert*, preprint, arXiv:2403.19858v2 [math.DS]. *(1214)*

[29] M. COTI ZELATI, M. G. DELGADINO, and T. M. ELGINDI, *On the relation between enhanced dissipation timescales and mixing rates*, Comm. Pure Appl. Math. **73** (2020), no. 6, 1205–1244. MR 4156602. DOI 10.1002/cpa.21831. *(1215, 1222, 1234)*

[30] M. COTI ZELATI and M. DOLCE, *Separation of time-scales in drift-diffusion equations on* $\mathbb{R}^2$, J. Math. Pures Appl. (9) **142** (2020), 58–75. MR 4149684. DOI 10.1016/j.matpur.2020.08.001. *(1211, 1215)*

[31] M. COTI ZELATI and T. D. DRIVAS, *A stochastic approach to enhanced diffusion*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **22** (2021), no. 2, 811–834. MR 4288672. *(1211, 1214)*

[32] M. COTI ZELATI, T. M. ELGINDI, and K. WIDMAYER, *Enhanced dissipation in the Navier-Stokes equations near the Poiseuille flow*, Comm. Math. Phys. **378** (2020), no. 2, 987–1010. MR 4134940. DOI 10.1007/s00220-020-03814-0. *(1215)*

[33] G. CRIPPA and C. DE LELLIS, *Estimates and regularity results for the DiPerna-Lions flow*, J. Reine Angew. Math. **616** (2008), 15–46. MR 2369485. DOI 10.1515/CRELLE.2008.016. *(1215)*

[34] M. F. DEMERS, *Uniqueness and exponential mixing for the measure of maximal entropy for piecewise hyperbolic maps*, Discrete Contin. Dyn. Syst. **41** (2021), no. 1, 217–256. MR 4182320. DOI 10.3934/dcds.2020217. *(1220, 1242, 1245)*

[35] M. F. DEMERS and C. LIVERANI, *Stability of statistical properties in two-dimensional piecewise hyperbolic maps*, Trans. Amer. Math. Soc. **360** (2008), no. 9, 4777–4814. MR 2403704. DOI 10.1090/S0002-9947-08-04464-4. *(1212, 1215, 1222, 1234, 1238)*

[36] W. DOEBLIN, *Sur les propriétés asymptotiques de mouvements régis par certains types de chaînes simples*, Numdam, n.p., 1938. MR 3533063. *(1224)*

[37] J. L. DOOB, *Asymptotic properties of Markoff transition probabilities*, Trans. Amer. Math. Soc. **63** (1948), no. 3, 393–421. MR 0025097. DOI 10.2307/1990566. *(1224)*

[38] T. D. DRIVAS, T. M. ELGINDI, G. IYER, and I.-J. JEONG, *Anomalous dissipation in passive scalar transport*, Arch. Ration. Mech. Anal. **243** (2022), no. 3, 1151–1180. MR 4381138. DOI 10.1007/s00205-021-01736-2. *(1212)*

[39] T. M. ELGINDI and K. LISS, *Norm growth, non-uniqueness, and anomalous dissipation in passive scalars*, Arch. Ration. Mech. Anal. **248** (2024), no. 6, art. ID 120. MR 4829551. DOI 10.1007/s00205-024-02056-x. *(1212)*

[40] T. M. ELGINDI and A. ZLATOŠ, *Universal mixers in all dimensions*, Adv. Math. **356** (2019), no. 106807. MR 4008523. DOI 10.1016/j.aim.2019.106807. *(1213, 1215, 1217, 1222, 1235)*

[41] L. C. EVANS and R. F. GARIEPY, *Measure Theory and Fine Properties of Functions*, Textb. Math., CRC Press, Boca Raton, 2015. MR 3409135. *(1249)*

[42] H. FEDERER, *Geometric Measure Theory*, Grundlehren Math. Wiss. **153**, Springer, New York, 1969. MR 0257325. *(1249)*

[43] Y. FENG and G. IYER, *Dissipation enhancement by mixing*, Nonlinearity **32** (2019), no. 5, 1810–1851. MR 3942601. DOI 10.1088/1361-6544/ab0e56. *(1215, 1222, 1234)*

[44] Y. FENG, A. L. MAZZUCATO, and C. NOBILI, *Enhanced dissipation by circularly symmetric and parallel pipe flows*, Phys. D **445** (2023), no. 133640. MR 4530996. DOI 10.1016/j.physd.2022.133640. *(1215)*

[45] T. GALLAY and M. COTI ZELATI, *Enhanced dissipation and Taylor dispersion in higher-dimensional parallel shear flows*, J. Lond. Math. Soc. (2) **108** (2023), no. 4, 1358–1392. MR 4655267. DOI 10.1112/jlms.12782. *(1211, 1215)*

[46] S. HE, *Enhanced dissipation, hypoellipticity for passive scalar equations with fractional dissipation*, J. Funct. Anal. **282** (2022), no. 3, art. ID 109319. MR 4343756. DOI 10.1016/j.jfa.2021.109319. *(1211, 1215)*

[47] J. M. HILL, R. STURMAN, and M. C. T. WILSON, *Exponential mixing by orthogonal non-monotonic shears*, Phys. D **434** (2022), no. 133224. MR 4399244. DOI 10.1016/j.physd.2022.133224. *(1211, 1215, 1217)*

[48] ———, *Loss of exponential mixing in a non-monotonic toral map*, preprint, arXiv:2303.08515v2 [math.DS]. *(1211, 1214, 1217)*

[49] G. IYER, A. KISELEV, and X. XU, *Lower bounds on the mix norm of passive scalars advected by incompressible enstrophy-constrained flows*, Nonlinearity **27** (2014), no. 5, 973–985. MR 3207161. DOI 10.1088/0951-7715/27/5/973. *(1215)*

[50] G. IYER, E. LU, and J. NOLEN, *Using Bernoulli maps to accelerate mixing of a random walk on the torus*, Quart. Appl. Math. **82** (2024), no. 2, 359–390. MR 4720207. DOI 10.1090/qam/1668. *(1216)*

[51] G. IYER, X. XU, and A. ZLATOŠ, *Convection-induced singularity suppression in the Keller-Segel and other non-linear PDEs*, Trans. Amer. Math. Soc. **374** (2021), no. 9, 6039–6058. MR 4302154. DOI 10.1090/tran/8195. *(1215)*

[52] G. IYER and H. ZHOU, *Quantifying the dissipation enhancement of cellular flows*, SIAM J. Math. Anal. **55** (2023), no. 6, 6496–6516. MR 4662406. DOI 10.1137/22M1524576. *(1215, 1216)*

[53] A. KATOK and J.-M. STRELCYN, *Invariant Manifolds, Entropy and Billiards: Smooth Maps with Singularities*, Lecture Notes in Math. **1222**, Springer, Berlin, 1986. MR 0872698. DOI 10.1007/BFb0099031. *(1215, 1217)*

[54] L. KELVIN, *Stability of fluid motion: Rectilinear motion of viscous fluid between two parallel plates*, Phil. Mag. **24** (1887), no. 5, 188–196. *(1215)*

[55] A. KISELEV, R. SHTERENBERG, and A. ZLATOŠ, *Relaxation enhancement by time-periodic flows*, Indiana Univ. Math. J. **57** (2008), no. 5, 2137–2152. MR 2463964. DOI 10.1512/iumj.2008.57.3349. *(1215)*

[56] A. KOLMOGOROV, *Zufällige Bewegungen (zur Theorie der Brownschen Bewegung)*, Ann. of Math. (2) **35** (1934), no. 1, 116–117. MR 1503147. DOI 10.2307/1968123. *(1215)*

[57] D. A. LEVIN and Y. PERES, *Markov Chains and Mixing Times*, 2nd ed., Amer. Math. Soc., Providence, 2017. MR 3726904. DOI 10.1090/mbk/107. *(1224)*

[58] Z. LIN, J.-L. THIFFEAULT, and C. R. DOERING, *Optimal stirring strategies for passive scalar mixing*, J. Fluid Mech. **675** (2011), 465–476. MR 2801050. DOI 10.1017/S0022112011000292. *(1215)*

[59] C. LIVERANI, *Decay of correlations*, Ann. of Math. (2) **142** (1995), no. 2, 239–301. MR 1343323. DOI 10.2307/2118636. *(1215)*

[60] E. LUNASIN, Z. LIN, A. NOVIKOV, A. MAZZUCATO, and C. R. DOERING, *Optimal mixing and optimal stirring for fixed energy, fixed power, or fixed palenstrophy flows*, J. Math. Phys. **53** (2012), no. 11, art. ID 115611. MR 3026556. DOI 10.1063/1.4752098. *(1215)*

[61] N. MASMOUDI and W. ZHAO, *Stability threshold of two-dimensional Couette flow in Sobolev spaces*, Ann. Inst. H. Poincaré C Anal. Non Linéaire **39** (2022), no. 2, 245–325. MR 4412070. DOI 10.4171/aihpc/8. *(1215)*

[62] S. P. MEYN and R. L. TWEEDIE, *Markov Chains and Stochastic Stability*, Comm. Control Engrg. Ser., Springer, London, 1993. MR 1287609. DOI 10.1007/978-1-4471-3267-7. *(1224)*

[63] F. MORGAN, *Geometric Measure Theory: A Beginner's Guide*, 5th ed., Elsevier/Academic Press, Amsterdam, 2016. MR 3497381. *(1249)*

[64] L. I. NICOLAESCU, *The co-area formula*, preprint, 2011, http://www3.nd.edu/~lnicolae/Coarea.pdf. *(1249)*

[65] R. T. PIERREHUMBERT, *Tracer microstructure in the large-eddy dominated regime*, Chaos Solitons Fractals **4** (1994), no. 6, 1091–1110. DOI 10.1016/0960-0779(94)90139-2. *(1216)*

[66] Z. TAO and M. ZWORSKI, *Optimal enhanced dissipation for contact Anosov flows*, preprint, arXiv:2311.01000v2 [math.AP]. *(1214)*

[67] J. VUKADINOVIC, *The limit of vanishing diffusivity for passive scalars in Hamiltonian flows*, Arch. Ration. Mech. Anal. **242** (2021), no. 3, 1395–1444. MR 4334729. DOI 10.1007/s00205-021-01707-7. *(1215)*

[68] D. WEI, *Diffusion and mixing in fluid flow via the resolvent estimate*, Sci. China Math. **64** (2021), no. 3, 507–518. MR 4215997. DOI 10.1007/s11425-018-9461-8. *(1211, 1215)*

[69] D. WEI and Z. ZHANG, *Enhanced dissipation for the Kolmogorov flow via the hypocoercivity method*, Sci. China Math. **62** (2019), 1219–1232. MR 3951889. DOI 10.1007/s11425-018-9508-5. *(1215)*

[70] D. WEI, Z. ZHANG, and W. ZHAO, *Linear inviscid damping and enhanced dissipation for the Kolmogorov flow*, Adv. Math. **362** (2020), no. 106963. MR 4050586. DOI 10.1016/j.aim.2019.106963. *(1215)*

[71] L.-S. YOUNG, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. of Math. (2) **147** (1998), no. 3, 585–650. MR 1637655. DOI 10.2307/120960. *(1215)*

[72] C. ZILLINGER, *On geometric and analytic mixing scales: Comparability and convergence rates for transport problems*, Pure Appl. Anal. **1** (2019), no. 4, 543–570. MR 4026549. DOI 10.2140/paa.2019.1.543. *(1222)*

[73] A. ZLATOŠ, *Diffusion in fluid flow: Dissipation enhancement by flows in 2D*, Comm. Partial Differential Equations **35** (2010), no. 3, 496–534. MR 2748635. DOI 10.1080/03605300903362546. *(1215)*

[74] A. ZLATOŠ and Y. YAO, *Mixing and un-mixing by incompressible flows*, J. Eur. Math. Soc. (JEMS) **19** (2017), no. 7, 1911–1948. MR 3656475. DOI 10.4171/JEMS/709. *(1215, 1222, 1235)*

Elgindi
Department of Mathematics, Duke University, Durham, North Carolina, USA;
tarek.elgindi@duke.edu

Liss
Department of Mathematics, Duke University, Durham, North Carolina, USA;
kyle.liss@duke.edu

Mattingly
Department of Mathematics, Duke University, Durham, North Carolina, USA;
jonathan.mattingly@duke.edu