

# Universal Mixers in All Dimensions

Tarek M. Elgindi and Andrej Zlatoš

September 14, 2019

## Abstract

We construct universal mixers, incompressible flows that mix arbitrarily well general solutions to the corresponding transport equation, in all dimensions. This mixing is exponential in time (i.e., essentially optimal) for any initial condition with at least some regularity, and we also show that a uniform mixing rate for all initial conditions cannot be achieved. The flows are time periodic and uniformly-in-time bounded in spaces  $W^{s,p}$  for a range of  $(s, p)$  that includes points with  $s > 1$  and  $p > 2$ .

## 1 Introduction

The problem of mixing via incompressible flows is classical and rich with connections to several branches of analysis including PDE, geometric measure theory, ergodic theory, and topological dynamics. When diffusion is absent or negligible over the relevant time scales, one can model the process of mixing by the transport equation

$$\rho_t + u \cdot \nabla \rho = 0, \tag{1.1}$$

with a fluid velocity  $u : Q_d \times \mathbb{R}^+ \rightarrow \mathbb{R}^d$  and  $Q_d$  being some  $d$ -dimensional physical domain. Incompressibility of the advecting fluid requires  $u$  to be divergence-free and we also assume the no-flow boundary condition for  $u$ , that is, the fluid does not cross the boundary of the domain and satisfies  $u \cdot n = 0$  on  $\partial Q_d \times \mathbb{R}^+$ . Since we are interested in the study of mixing in the bulk of the domain and not in effects of rough boundaries, we will simply assume that  $Q_d := (0, 1)^d$  is either the unit cube in  $\mathbb{R}^d$  or it is  $\mathbb{T}^d$  (in the latter case opposite sides of  $(0, 1)^d$  are identified, so  $\partial \mathbb{T}^d = \emptyset$  and the boundary conditions instead become periodic).

The function  $\rho : Q_d \times \mathbb{R}^+ \rightarrow \mathbb{R}$  represents the concentration of the mixed quantity with a given initial value  $\rho(\cdot, 0)$ , which we can allow to be negative on account of (1.1) being invariant with respect to addition of constants. It will be convenient to take  $\rho$  to be mean-zero, so we will always assume that  $\int_{Q_d} \rho(x, 0) dx = 0$ . A special case is when  $\rho(\cdot, 0)$  is the characteristic function of a subset of  $Q_d$  (minus a constant), modeling the mixing of two fluids. The central question now is how well is it possible to mix a given initial condition via divergence-free flows satisfying some physically relevant constraints, which flows are most efficient at this process, and what is their dependence on the initial condition.

---

<sup>1</sup>Department of Mathematics, UC San Diego. E-mail: telgindi@ucsd.edu.

<sup>2</sup>Department of Mathematics, UC San Diego. E-mail: zlatos@ucsd.edu.

In order to study mixing efficiency of flows, one needs to start with a quantitative definition of how well mixed the advected scalar  $\rho$  is at any given time (see the review [31] for various options). While diffusive mixing results in the time-decay of  $L^p$  norms of  $\rho$  [8, 12, 29, 32] (recall that  $\rho$  is mean zero), these norms are all conserved for (1.1). Instead, we need to use measures of mixing that capture the small-scale variations of  $\rho$ . The following two definitions, one geometric in flavor and the other functional-analytic, have been in use recently (see also the discussion at the end of this introduction for a relation to the dynamical systems point of view). In them, when  $Q_d \subseteq \mathbb{R}^d$  (i.e., not  $\mathbb{T}^d$ ), then we extend the function  $f : Q_d \rightarrow \mathbb{R}$  on the rest of  $\mathbb{R}^d$  by zero; and as always  $\int_A f dx dy = |A|^{-1} \int_A f dx dy$ .

**Definition 1.1.** Let  $f \in L^\infty(Q_d)$  be mean-zero on  $Q_d$ .

(i) We say that  $f$  is  $\kappa$ -mixed to scale  $\varepsilon$ , with  $\kappa, \varepsilon \in (0, 1)$ , if for each  $y \in Q_d$ ,

$$\left| \int_{B_\varepsilon(y)} f(x) dx \right| \leq \kappa \|f\|_\infty.$$

The smallest such  $\varepsilon$  is the ( $\kappa$ -dependent) *geometric mixing scale* of  $f$ .

(ii) The *functional mixing scale* of  $f$  is  $\|f\|_{\dot{H}^{-1/2}}^2 \|f\|_\infty^{-2}$ .

*Remarks.* 1. The definition in (i) is from [33], which was in turn motivated by [6], where the special case  $d = 2$ ,  $\kappa = \frac{1}{3}$ ,  $f(Q_2) = \{-1, 1\}$  was considered. In it, the “worst-mixed” region determines the mixing scale, but deviations of size roughly  $\kappa$  are tolerated.

2. The definition in (ii), which is motivated by [22] (where  $\mathbb{T}^d$  was considered instead of  $Q_d$  but their analysis easily extend to  $Q_d$ ), does not tolerate deviations but averages the degree of “mixedness” of  $f$  over all of  $Q_d$ . To see the latter, we note that (30) in [22] shows for mean-zero functions equivalence of the  $\dot{H}^{-1/2}$ -norm and the *mix-norm*

$$\Phi(f) := \left[ \int_{Q_d \times (0,1)} \left( \int_{B_r(y)} f(x) dx \right)^2 dy dr \right]^{1/2}. \quad (1.2)$$

3. Other  $\dot{H}^{-s}$ -norms of  $f$  have been used to quantify mixing, particularly the  $\dot{H}^{-1}$ -norm [2, 17, 19, 21, 27], with the functional mixing scale being  $\|f\|_{\dot{H}^{-1}} \|f\|_\infty^{-1}$  (Wasserstein distance of  $f_+$  and  $f_-$  has also been used [5, 25, 27, 28]). The latter may sometimes be more convenient than the  $\dot{H}^{-1/2}$ -norm and also is directly related to mixing-enhanced diffusion rates when diffusion is present [8, 10], but it lacks the useful connection to the mix-norm. We use the  $\dot{H}^{-1/2}$ -norm here but note that our mixing results for it also hold for the  $\dot{H}^{-1}$ -norm because the former controls the latter. The only case when this is not obvious is non-existence of a universal mixing rate in Theorem 1(ii), but that proof can be easily adjusted to accommodate the  $\dot{H}^{-1}$ -norm (see Remark 2 after Lemma 2.5).

If  $\rho$  solves (1.1) with mean-zero  $\rho(\cdot, 0) \in L^\infty(Q_d)$  and divergence free  $u : Q_d \times \mathbb{R}^+ \rightarrow \mathbb{R}^d$ , we will say that the flow  $\kappa$ -mixes  $\rho(\cdot, 0)$  to scale  $\varepsilon$  by time  $t$  whenever  $\rho(\cdot, t)$  is  $\kappa$ -mixed to scale  $\varepsilon$ . In [6, 7], Bressan conjectured that if  $\rho(\cdot, 0) = \chi_{(0,1/2) \times (0,1)} - \chi_{(1/2,1) \times (0,1)}$  on  $Q_2$  is  $\frac{1}{3}$ -mixed to some scale  $\varepsilon \ll 1$  in time  $t$  by a divergence-free  $u$ , then

$$\int_0^t \|\nabla u(\cdot, \tau)\|_1 d\tau \geq C |\log \varepsilon|$$

(with some  $\varepsilon$ -independent  $C < \infty$ ). After an appropriate change of the time variable, this is equivalent to existence of  $C < \infty$  such that if a divergence-free  $u$  satisfies

$$\sup_{t>0} \|\nabla u(\cdot, t)\|_1 \leq 1 \tag{1.3}$$

and it  $\frac{1}{3}$ -mixes  $\rho(\cdot, 0)$  to some scale  $\varepsilon \ll 1$  in time  $t$ , then  $t \geq C|\log \varepsilon|$ . That is, given the constraint (1.3), the mixing scale cannot decrease super-exponentially in time (in two dimensions).

This *rearrangement cost conjecture* of Bressan is still open, but Crippa and De Lellis proved its version with any  $\kappa > 0$  and (1.3) replaced by

$$\sup_{t>0} \|\nabla u(\cdot, t)\|_p \leq 1 \tag{1.4}$$

with arbitrary  $p > 1$  [11]. This result also extends to the  $\dot{H}^{-s}$ -norm-based definition of mixing via the relationship of the  $\dot{H}^{-1/2}$ -norm and the mix-norm [17, 27].

This motivates the natural question of whether for any mean zero  $\rho(\cdot, 0) \in L^\infty(Q_2)$  in two dimensions, there is a flow that satisfies (1.4) and achieves the (qualitatively optimal for  $p > 1$ ) exponential-in-time decay of the mixing scale of the solution  $\rho$ . This was recently answered in the affirmative by Yao and Zlatoš [33] for any such  $\rho(\cdot, 0)$  and all  $p \in [1, \frac{3+\sqrt{5}}{2})$ , while for  $p \geq \frac{3+\sqrt{5}}{2}$  they proved existence of flows achieving mixing scale rates  $O(e^{-t^{\nu_p}/C_p})$  with  $\nu_p > \frac{3}{4}$  (geometric only). Alberti, Crippa, and Mazzucato [1, 2] independently showed that exponential decay can be obtained for all  $p \geq 1$ , but their result only applies to special initial data  $\rho(\cdot, 0)$  (characteristic functions of some regular sets, up to constants).

In both these works, the mixing flows intricately depend on the initial data, which may not always be very practical. It is therefore natural to ask whether this dependence can be removed and universally mixing flows exist. This was done in questions (C) and (D) of the following list of five fundamental questions from [2] concerning mixing by flows.

- (A) Given any mean-zero  $\rho(\cdot, 0) \in L^\infty(Q_2)$ , is there a divergence-free  $u$  satisfying (1.4) such that the mixing scale (functional or geometric) of the solution  $\rho$  decays to zero as  $t \rightarrow \infty$ ?
- (B) If the answer to (A) is affirmative, can  $u$  be chosen so that the mixing scale decay is exponential?
- (C) Is there a divergence-free  $u$  on  $Q_2$  satisfying (1.4) such that the mixing scale (functional or geometric) of the solution  $\rho$  decays to zero as  $t \rightarrow \infty$  for any mean-zero  $\rho(\cdot, 0) \in L^\infty(Q_2)$ ?
- (D) If the answer to (A) is affirmative, can  $u$  be chosen so that the mixing scale decay is exponential?
- (E) Is there  $s > 1$ , a mean-zero  $\rho(\cdot, 0) \in L^\infty(Q_2)$ , and a divergence-free  $u$  satisfying

$$\sup_{t>0} \|u(\cdot, t)\|_{W^{s,p}} \leq 1 \tag{1.5}$$

such that the mixing scale (functional or geometric) of the solution  $\rho$  decays to zero exponentially as  $t \rightarrow \infty$ ?

Additionally, it is natural to state the universal mixing version of (E):

(F) If the answer to (E) is affirmative, can a single such  $u$  be chosen for all mean-zero  $\rho(\cdot, 0) \in L^\infty(Q_2)$ ?

The flows from [2, 33] are self-similar in nature with exponentially-in-time decreasing turbulent scales, which appears to be a serious obstacle when it comes to (E) and (F). Moreover, both [2, 33] only consider the two-dimensional case  $d = 2$  and the construction in [33] for general initial data does not appear to extend easily to  $d > 2$  (the trivial extension of [2, 33] involving initial data depending on only two coordinates is obviously possible). Hence an obvious seventh question is

(G) What are the answers to (A)–(F) in dimensions  $d \geq 3$ ?

The results from [33] answer (A) and (B) in the affirmative for  $p \in [1, \frac{3+\sqrt{5}}{2})$ , as well as (A) in the affirmative for  $p \geq \frac{3+\sqrt{5}}{2}$  and the geometric mixing scale only. (Additionally, [2] proves exponential mixing for all  $p \geq 1$  and some special initial data  $\rho(\cdot, 0)$ .) In this paper we answer (C)–(F) for  $p \in [1, \frac{3+\sqrt{5}}{2})$ , as well as the multi-dimensional version (G) of (A)–(F) (except of (B) for very rough  $\rho(\cdot, 0)$ ). For these  $p$  and in all dimensions  $d \geq 2$ , the answers to (C) and (E) are affirmative, while the answers to (D) and (F) are “no but morally yes”. More precisely, we construct universally mixing flows in all dimensions that also yield exponential mixing for all mean-zero initial conditions  $\rho(\cdot, 0)$  with at least some regularity (belonging to  $H^\sigma(Q_2)$  for some  $\sigma > 0$ , which includes all characteristic functions of regular sets), while we also show that no flow has a fully universal mixing rate (exponential or otherwise). Moreover, the flows we construct will be periodic in time, with no emergence of small-scale structures as  $t \rightarrow \infty$ . We also note that all the answers are the same for both the geometric (with any  $\kappa > 0$ ) and functional mixing.

The following definition encapsulates our goals.

- Definition 1.2.** (i) A divergence-free  $u : Q_d \times \mathbb{R}^+ \rightarrow \mathbb{R}^d$  is called a *universal mixer* (in the functional sense or geometric sense for some fixed  $\kappa \in (0, 1)$ ) if for any  $\rho(\cdot, 0) \in L^\infty(Q_d)$  the mixing scale (functional or geometric with the given  $\kappa$ ) of the solution to (1.1) converges to 0 as  $t \rightarrow \infty$ .
- (ii) A universal mixer  $u$  (in either sense from (i)) has *mixing rate*  $\lambda : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ , with  $\lambda$  decreasing and  $\lim_{t \rightarrow \infty} \lambda(t) = 0$ , if for each  $\rho(\cdot, 0) \in L^\infty(Q_d)$  there is  $\tau < \infty$  such that the mixing scale of  $\rho(\cdot, t)$  (in the relevant sense) is at most  $\lambda(t - \tau)$  for all  $t > \tau$ .
- (iii) A universal mixer  $u$  (in either sense from (i)) is a *universal exponential mixer* if there is  $\gamma > 0$  such that  $u$  has mixing rate  $\lambda(t) = e^{-\gamma t}$ .
- (iv) A universal mixer  $u$  (in either sense from (i)) is an *almost-universal exponential mixer* if for each  $\sigma > 0$  there is  $\gamma_\sigma > 0$  such that for each  $\rho(\cdot, 0) \in L^\infty(Q_d) \cap H^\sigma(Q_d)$  there is  $\tau < \infty$  such that the mixing scale of  $\rho(\cdot, t)$  (in the relevant sense) is at most  $e^{-\gamma_\sigma(t-\tau)}$  for all  $t > \tau$ .
- (v) If  $u$  is a universal mixer in the geometric sense for each  $\kappa \in (0, 1)$ , we say that  $u$  is a *universal mixer in the geometric sense*. If  $u$  is an (almost-)universal exponential mixer in the geometric sense for each  $\kappa \in (0, 1)$ , with the relevant exponential mixing rates independent of  $\kappa$ , then we say that  $u$  is an *(almost-)universal exponential mixer in the geometric sense*.

Here are our main results, for  $Q_d$  being either the unit cube or the torus  $\mathbb{T}^d$ .

**Theorem 1.** (i) For any  $d \geq 2$ , there is a divergence-free time-periodic vector field  $u$  on  $Q_d$  satisfying no-flow (or periodic) boundary conditions such that  $u \in L^\infty([0, \infty); W^{s,p}(Q_d))$  (or  $u \in L^\infty([0, \infty); W^{s,p}(\mathbb{T}^d))$ ) for any  $s < \frac{1+\sqrt{5}}{2}$  and  $p \in [1, \frac{2}{2s+1-\sqrt{5}})$ , and  $u$  is a universal mixer and an almost-universal exponential mixer in both the geometric and functional sense.

(ii) For any  $d \geq 2$ , there is no divergence-free universal mixer on  $Q_d$  satisfying either no-flow or periodic boundary conditions that has a mixing rate, in either the functional sense or the geometric sense with some  $\kappa \in (0, 1)$ . In particular, there are no universal exponential mixers in any dimension.

*Remarks.* 1. Hence  $(s, p)$  with  $s > 1$  and  $p > 2$  are included in (i). As mentioned above, prior mixing results involving at least  $W^{1,1}$  flows required two dimensions  $d = 2$ , as well as  $u$  that depended on the mixed function and only belonged to spaces  $L^\infty([0, \infty); W^{1,p}(Q_2))$ . There are, however, results involving less regular universal mixers such as the class of linked twist maps (see [30] and the discussion below for more details).

2. The one-period flow map of our flow in (i) on  $Q_2$  with no-flow boundary conditions is the folded Baker's map, and the other cases are also related to it (see Section 2). Thus it is exponentially mixing in the dynamical systems sense, and topologically conjugate to the Smale Horseshoe map.

3. Our  $u$  in (i) will in fact only take finitely many (two when  $d = 2$ ) distinct values as a function of time. This is virtually the simplest possible time-dependence, and we do not know whether time-independent universal mixers exist on  $Q_d$  (they do not for  $d = 2$ , due to divergence-free vector fields on  $Q_2$  having a stream function).

4. While the flows we construct in (i) are discontinuous in time, they can be made smooth in time by a simple re-parametrization described in [33]. In space, these flows are Hölder continuous, and smooth away from a finite number of hyperplanes.

5. (i) also shows that replacing the  $W^{1,p}$ -norm of  $u$  in the constraint (1.4) by the  $W^{s,p}$ -norm for  $(s, p)$  as in (i) does not qualitatively improve the exponential upper bound on mixing efficiency of flows (i.e., exponential lower bound on the mixing scale) from [11, 17, 27].

6. We also remark that it follows from (i) that mean-zero solutions to the corresponding advection-diffusion equation

$$\rho_t + u \cdot \nabla \rho = \nu \Delta \rho$$

with  $\nu \in (0, \frac{1}{2})$  satisfy

$$\|\rho(\cdot, t)\|_2 \leq e^{-c\nu^{0.62}t} \|\rho(\cdot, 0)\|_2$$

for all  $t > \nu^{-0.62}$ , with  $c$  a universal constant (see Theorem 4.4 in [10], or [16]). That is, their diffusive time scale is at most  $O(\nu^{-0.62})$ , which is much shorter than the time scale  $O(\nu^{-1})$  for the heat equation (without advection) when  $0 < \nu \ll 1$ . However, this still does not appear to be optimal.

We do not know whether universal mixers or almost universal exponential mixers with a higher regularity than ours exist. One candidate for an efficient universal mixer in two dimensions (on  $\mathbb{T}^2$ ) is almost every realization of the random vector field taking values  $(\sin(2\pi x_2 + \omega_n), 0)$  and  $(0, b \sin(2\pi x_1 + \omega_n))$  (with  $b \in \mathbb{R}$  a constant) on time intervals  $(n-1, n-\frac{1}{2}]$  and  $(n-\frac{1}{2}, n]$ , respectively, for  $n \in \mathbb{N}$  and with  $\omega_n$  independent random variables uniformly distributed over

$\mathbb{T}$ . While these alternatively horizontal and vertical shear flows, considered by Pierrehumbert in [26], appear to have very good mixing properties, we are not aware of their rigorous proofs.

**Discussion of related dynamical systems results.** The study of mixing maps and flows has a rich history. While there are a plethora of *maps* that are known to be good mixers, there are only a few examples of *flows* that mix well (see [9]). One important class of examples are Anosov flows, introduced in [3]. A flow  $\Phi_t : M \rightarrow M$  on a compact Riemannian manifold  $M$  is an Anosov flow if at each  $x \in M$ , the tangent space  $TM_x$  can be decomposed into three subspaces, one contracting, one expanding, and one that is 1-dimensional and corresponds to the direction of the flow. It was shown in important works of Dolgopyat [13] and Liverani [20] that all smooth enough Anosov flows are exponentially mixing in the sense of the decay of correlations (which implies exponential mixing in the sense of Definition 1.1). Anosov flows are known to exist in a number of settings, the most concrete of which seems to be as geodesic flows on certain negatively curved Riemannian manifolds of dimensions  $d \geq 3$ . A very interesting open problem is to construct an incompressible velocity field in the flat geometry of  $\mathbb{T}^d$  for  $d \geq 3$  that is smooth uniformly in time, and whose flow is an exponential mixer. Theorem 1(i) provides a Hölder continuous time-periodic example on both  $\mathbb{T}^d$  and  $(0, 1)^d$  for  $d \geq 2$ . (We note that A. Katok constructed mixing flows on all smooth two dimensional manifolds [18], but their mixing rates are only logarithmic or algebraic in time, depending on the allowed regularity of the flows [14].)

It is also easy to show that many linked twist maps lead to exponential mixers on  $\mathbb{T}^2$  (see [30]). For example, Arnold’s cat map can be realized as the composition of maps  $(x, y) \mapsto (x + y, y)$  and  $(x, y) \mapsto (x, x + y)$ , which are both flow maps of divergence-free velocity fields (exponential mixing in this case is not difficult to establish). However, these velocity fields are discontinuous on  $\mathbb{T}^2$ , and although they belong to the space  $BV$ , they do not belong to  $W^{s,p}$  for any  $s, p \geq 1$ . To the best of our knowledge, prior to this work there were no known exponential mixers on  $\mathbb{T}^d$  or  $(0, 1)^d$  (for any  $d \geq 2$ ) with regularity higher than  $BV$  (see [15] for even more irregular examples). Allowing the velocity fields to be discontinuous also trivially allows for flow maps with rigid and “cut-and-paste” motions, which can be easily combined to produce exponential mixers. The requirement of continuity, let alone the constraint (1.5) when  $s, p > 1$ , adds non-trivial difficulties.

It is also important to emphasize that none of the above examples, whether on negatively curved manifolds or even on  $\mathbb{T}^d$ , answers questions about mixing in physically relevant real world settings, that is, bounded domains in  $\mathbb{R}^2$  and  $\mathbb{R}^3$ . Theorem 1(i) and the (initial-data-dependent) two-dimensional flows from [33] are the only exponentially mixing flows in the deterministic setting that we are currently aware of (for the stochastic setting see the paper [4], which was written several months after the present paper).

**Acknowledgements.** AZ acknowledges partial support by NSF grants DMS-1652284 and DMS-1656269. TME acknowledges partial support by NSF grant DMS-1817134. We also thank Dmitri Dolgopyat for pointing us to [15, 18], Gautam Iyer for mentioning to us [26], Amir Mohammadi and Sheldon Newhouse for useful discussions, and the anonymous referee for helpful comments that improved the exposition of this paper.

## 2 Discrete time mixing

Since our flow will be time-periodic, a crucial step will be the analysis of its flow map at integer multiples of its period. Of course, this is just the sequence of powers of the flow map for a single period. We note that Definition 1.2 naturally extends to the case of the initial datum  $\rho(\cdot, 0)$  and the solution  $\rho$  being replaced by, respectively,  $f \in L^\infty(Q_d)$  and the sequence of functions  $\{f \circ T^{-n}\}_{n \geq 1}$  obtained by repeatedly applying some measure-preserving bijection  $T : Q_d \rightarrow Q_d$  to  $f$ . Or, more generally, with  $\{f \circ T_n^{-1}\}_{n \geq 1}$  instead of  $\{f \circ T^{-n}\}_{n \geq 1}$ , where  $T_n : Q_d \rightarrow Q_d$  is some sequence of measure-preserving bijections. In those cases we will call  $T$  resp.  $\{T_n\}_{n \geq 1}$  (almost-)universal (exponential) mixers on  $Q_d$ .

We start with the two-dimensional case, in which we will use the notation  $(x, y) \in Q_2$ , and afterwards extend our analysis to all dimensions. We note that the maps considered here will only work as flow maps in the no-flow case  $Q_d = (0, 1)^d$ . Adjustments needed for the periodic case  $\mathbb{T}^d$  will be performed in the next section.

### 2.1 Construction of a discrete time universal mixer $T : Q_2 \rightarrow Q_2$

There are various maps on the square  $Q_2 = (0, 1)^2$  that are known to have good mixing properties, one such example being *Baker's map*. The classical Baker's map  $B : Q_2 \rightarrow Q_2$  is obtained by cutting the square vertically in two halves and mapping these affinely onto the upper and lower halves of  $Q_2$ , with no rotation. Unfortunately, it seems that divergence-free velocity fields whose flow map at some time is  $B$  are no more regular than BV. It turns out, however, that a closely related map, for which one of the rectangles is rotated by  $180^\circ$  (the *folded Baker's map* in Figure 1 below), has very similar mixing properties and can be realized via incompressible flows with higher regularity. We analyze it now, and will construct the relevant velocity field in the next section.

**Definition 2.1.** Let  $Q_2 := (0, 1)^2$ , and

$$Q'_2 := Q_2 \setminus \{(x, y) \in Q_2 \mid 2^k x \in \mathbb{Z} \text{ or } 2^k y \in \mathbb{Z} \text{ for some } k \in \mathbb{N}\}.$$

Define  $T : Q'_2 \rightarrow Q'_2$  by

$$T(x, y) = \begin{cases} -(2x, \frac{y}{2}) + (1, \frac{1}{2}) & x \in (0, \frac{1}{2}), \\ (2x, \frac{y}{2}) + (-1, \frac{1}{2}) & x \in (\frac{1}{2}, 1). \end{cases} \quad (2.1)$$

For  $k \in \mathbb{N} \cup \{0\}$  and  $j \in \mathbb{Z} \cap [0, 2^k)$ , let

$$H_k^j := \left[ (0, 1) \times \left( \frac{j}{2^k}, \frac{j+1}{2^k} \right) \right] \cap Q'_2,$$

$$V_k^j := \left[ \left( \frac{j}{2^k}, \frac{j+1}{2^k} \right) \times (0, 1) \right] \cap Q'_2$$

be the *horizontal and vertical dyadic strips of width  $2^{-k}$* , respectively. Finally, let

$$G_{k,k'} := \left\{ H_k^j \cap V_{k'}^{j'} \mid j \in \mathbb{Z} \cap [0, 2^k) \text{ and } j' \in \mathbb{Z} \cap [0, 2^{k'}) \right\}$$

be the collection of all *dyadic rectangles* of size  $2^{-k'} \times 2^{-k}$  and let  $G_k := G_{k,k}$  be the collection of all *dyadic squares* of size  $2^{-k} \times 2^{-k}$ .

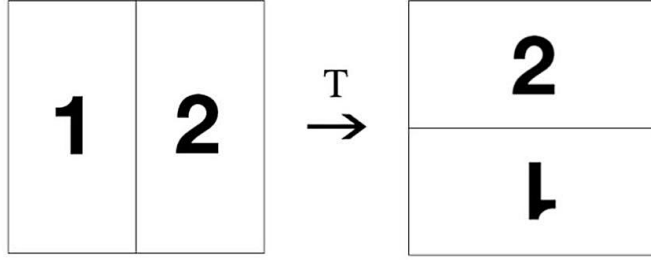


Figure 1: Folded Baker's map  $T$ .

The restriction of the above definitions to  $Q'_2$  is due to the fact that the map  $T$  is not defined on the line  $\{x = \frac{1}{2}\}$  and more generally,  $T^l$  is not defined on  $Q_2 \cap \{2^l x \in \mathbb{Z}\}$  for any  $l \in \mathbb{N}$ . Similarly,  $Q_2 \cap \{2^l y \in \mathbb{Z}\}$  is not in the image of  $T^l$ . However,  $T$  is a bijection on the full-measure set  $Q'_2$ , and restriction to  $Q'_2$  will avoid some technicalities.

**Lemma 2.2.** *Let  $k, k', l \in \mathbb{N}$ , and let  $j \in \mathbb{Z} \cap [0, 2^k)$  and  $j' \in \mathbb{Z} \cap [0, 2^{k'})$ .*

(i) *If  $i \in \mathbb{Z} \cap [0, 2^l)$ , then  $T^l(H_k^j) \cap H_i^i$  is a single horizontal dyadic strip of width  $2^{-(k+l)}$ .*

(ii) *If  $l \leq k'$ , then  $T^l(H_k^j \cap V_{k'}^{j'})$  is a single horizontal dyadic rectangle from  $G_{k+l, k'-l}$ .*

*Remarks.* 1. That is, any dyadic rectangle  $H_k^j \cap V_{k'}^{j'}$  is being doubled horizontally and halved vertically by repeated applications of  $T$  until it becomes a horizontal dyadic strip. Continued applications of  $T$  then always double the number of horizontal dyadic strips and halve their widths, while these strips become fairly regularly distributed throughout  $Q'_2$ . One can write down recursive formulas for  $T^l(H_k^j)$  and  $T^l(H_k^j \cap V_{k'}^{j'})$ , but we will not need these.

2. Baker's map  $B$ , which coincides with  $T$  for  $x \in (\frac{1}{2}, 1)$  but  $B(x, y) = (2x, \frac{y}{2})$  for  $x \in (0, \frac{1}{2})$ , has the same properties and the distances between the strips constituting  $B^l(H_k^j)$  are the same.

*Proof.* Both statements are immediate by induction on  $l$ , using that the bijections  $T : V_1^0 \rightarrow H_1^0$  and  $T : V_1^1 \rightarrow H_1^1$  are both of the form  $(x, y) \mapsto \pm(2x, \frac{y}{2}) + (\mp 1, \frac{1}{2})$ .  $\square$

This now directly implies that  $T$  is mixing in the classical sense. For the sake of completeness, let us extend  $T$  (bijectively) to all of  $Q_2$  by (2.1) for  $(x, y) \in Q_2 \setminus [\{\frac{1}{2}\} \times (0, 1)]$  and by  $T(x, y) := (y, x)$  for  $(x, y) \in \{\frac{1}{2}\} \times (0, 1)$ .

**Lemma 2.3.** *If  $A, B \subseteq Q_2$  are measurable, then*

$$\lim_{n \rightarrow \infty} |T^n(A) \cap B| = |A| \cdot |B|.$$

*Similarly, if  $f, g \in L^2(Q_2)$ , then*

$$\lim_{n \rightarrow \infty} \int_{Q_2} (f \circ T^{-n}) g \, dx dy = \int_{Q_2} f \, dx dy \int_{Q_2} g \, dx dy.$$



*Proof.* Assume first that  $A$  and  $B$  are each a disjoint union of dyadic squares  $H_k^j \cap V_k^{j'}$  (with a fixed  $k$ ). Applying first Lemma 2.2(ii) with  $l := k$  and then Lemma 2.2(i) with  $l := n - k$ , we see that for  $n \geq 2k$  we have

$$|T^n(H_k^j \cap V_k^{j'}) \cap (H_k^i \cap V_k^{i'})| = 2^{-4k} = |H_k^j \cap V_k^{j'}| \cdot |H_k^i \cap V_k^{i'}| \quad (2.2)$$

for any  $j, j', i, i' \in \mathbb{Z} \cap [0, 2^k)$  because the intersection on the left is a single dyadic rectangle of size  $2^{-k} \times 2^{-3k}$ . Hence  $|T^n(A) \cap B| = |A| \cdot |B|$  for all  $n \geq 2k$ . For general  $A$  and  $B$ , the first claim follows via approximation by disjoint unions of dyadic squares.

The second claim follows from the first via approximation of  $f$  and  $g$  by simple functions.  $\square$

This shows that  $T$  is a (discrete time) universal mixer.

**Lemma 2.4.**  *$T$  is a universal mixer on  $Q_2$  in both the geometric and functional senses.*

*Proof.* From the previous result we know that  $\{f \circ T^{-n}\}_{n \geq 1}$  converges to 0 weakly in  $L^2(Q_2)$  for any mean-zero  $f \in L^\infty(Q_2)$ , and hence strongly in  $\dot{H}^{-1/2}(Q_2)$ . This shows that  $T$  is a universal mixer in the functional sense. Lemma A.1 in [33], which is a bound on the geometric mixing scale in terms of the functional one, then proves the geometric sense claim.  $\square$

## 2.2 Non-existence of a uniform mixing rate in any dimension

We will now show that Lemma 2.4 is sharp in the sense that no sequence of measure preserving bijections on  $Q_d$  (for any  $d \geq 2$ ) with no-flow or periodic boundary conditions possesses a universal (functional or geometric) mixing rate.

**Lemma 2.5.** *Let  $\{T_n\}_{n \geq 1}$  be a sequence of measure-preserving bijections  $T_n : Q_d \rightarrow Q_d$  and let decreasing  $\{\lambda_n\}_{n \geq 1}$  satisfy  $\lim_{n \rightarrow \infty} \lambda_n = 0$ . Then there is a measurable set  $A \subseteq Q_d$  with  $|A| = \frac{1}{2}$  and a sequence  $n_j \rightarrow \infty$  as  $j \rightarrow \infty$  such that  $T_{n_j}(A)$  contains a ball with radius  $\lambda_{n_j}^{1/d}$  for each  $j \in \mathbb{N}$ . In particular, if  $f := \chi_A - \chi_{Q_d \setminus A}$ , then  $\int_{Q_d} f \, dx dy = 0$  but  $f \circ T_{n_j}^{-1}$  is not  $\kappa$ -mixed to scale  $\lambda_{n_j}$  for any  $\kappa \in (0, 1)$  and  $j \in \mathbb{N}$ , as well as  $\|f \circ T_{n_j}^{-1}\|_{\dot{H}^{-1/2}(Q_d)} > \lambda_{n_j}$  for all large enough  $j$ .*

*Remarks.* 1. One can also show that for any measurable  $B \subseteq Q_d$ , there is a measurable  $A \subseteq Q_d$  such that

$$\limsup_{n \rightarrow \infty} \frac{||T_n(A) \cap B| - |A| \cdot |B||}{\lambda_n} = \infty.$$

2. While the last claim does not imply  $\|f \circ T_{n_j}^{-1}\|_{\dot{H}^{-1}(Q_d)} > \lambda_{n_j}$ , the construction from the proof below can be easily adjusted to achieve this (using the definition of the  $\dot{H}^{-1}$ -norm).

*Proof.* Pick  $n_j$  so that  $V_d \sum_{j=1}^{\infty} \lambda_{n_j} \leq \frac{1}{2}$ , with  $V_d$  the volume of the unit ball in  $\mathbb{R}^d$ , and let  $A \subseteq Q_d$  be any measurable set with  $|A| = \frac{1}{2}$  containing  $\bigcup_{j \geq 1} T_{n_j}^{-1}(B_{\lambda_{n_j}^{1/d}}(\frac{1}{2}, \dots, \frac{1}{2}))$ . This proves the first two claims. The last claim follows from this and equivalence of the  $\dot{H}^{-1/2}$ -norm and the mix-norm (1.2), because  $f \circ T_{n_j}^{-1} \equiv 1$  on at least  $O(\lambda_{n_j})$  proportion of all the balls in  $Q_d$  with any fixed radius  $r \in (0, \frac{1}{2} \lambda_{n_j}^{1/d})$  (hence the mix norm of  $f \circ T_{n_j}^{-1}$  is at least  $O(\lambda_{n_j}^{(d+1)/2d})$ ).  $\square$

**Lemma 2.6.** *For any  $d \geq 2$ , there is no universal mixer  $\{T_n\}_{n \geq 1}$  with measure-preserving bijections  $T_n : Q_d \rightarrow Q_d$  that has some mixing rate, in either the functional sense or the geometric sense with some  $\kappa \in (0, 1)$ .*

*Proof.* Given any decreasing  $\{\lambda_n\}_{n \geq 1}$  with  $\lim_{n \rightarrow \infty} \lambda(n) = 0$ , pick a positive decreasing sequence  $\{\lambda'_n\}_{n \geq 1}$  that decays more slowly than  $\{\lambda(n+m)\}_{n \geq 1}$  for any  $m \in \mathbb{N}$ . Then apply Lemma 2.5.  $\square$

### 2.3 Almost-universal exponential mixing by $T : Q_2 \rightarrow Q_2$

Finally, part (iii) of the following lemma shows that  $T$  is an almost-universal exponential mixer in both the geometric and functional senses. Recall that the box-counting (or Minkowski) dimension of any  $E \subseteq Q_2$  is  $\dim_{\text{box}}(E) := \limsup_{k \rightarrow \infty} \frac{1}{k} \log_2 n_k(E)$ , where  $n_k(E)$  is the number of squares from  $G_k$  that have a non-empty intersection with  $E$ .

**Lemma 2.7.** *Assume that  $f \in L^\infty(Q_2)$  with  $\int_{Q_2} f \, dx dy = 0$ . For  $A \subseteq Q_2$ , let  $\mathcal{C}_r(A)$  be the set of  $(x, y), (x', y') \in A$  such that  $|(x, y) - (x', y')| < r$  and the closed straight segment connecting  $(x, y)$  and  $(x', y')$  is contained in  $A$ .*

(i) *If there is a set  $\Gamma \subseteq Q_2$  with box-counting dimension  $\gamma < 2$  and  $r > 0$  such that*

$$\sup_{(x,y),(x',y') \in \mathcal{C}_r(Q_2 \setminus \Gamma)} |f(x, y) - f(x', y')| < \kappa \|f\|_\infty,$$

*then there is  $m \in \mathbb{N}$  such that  $f \circ T^{-n}$  is  $\kappa$ -mixed to scale  $2^{-(n-m)/2}$  for each  $n \in \mathbb{N}$ .*

(ii) *If there is a set  $\Gamma \subseteq Q_2$  with box-counting dimension  $\gamma < 2$  and  $\alpha > 0$  such that*

$$\sup_{(x,y),(x',y') \in \mathcal{C}_1(Q_2 \setminus \Gamma)} \frac{|f(x, y) - f(x', y')|}{|(x, y) - (x', y')|^\alpha} < \infty,$$

*then the functional mixing scale of  $f \circ T^{-n}$  decreases exponentially at a rate that only depends on  $\min\{\alpha, 2 - \gamma\}$ .*

(iii) *If  $f \in H^\sigma(Q_2)$  for some  $\sigma \in (0, 1]$ , then the functional mixing scale of  $f \circ T^{-n}$  and its geometric mixing scale for any  $\kappa \in (0, 1)$  decrease exponentially at a rate that only depends on  $\sigma$ .*

*Remark.* The hypothesis in (ii) means that  $f$  is Hölder continuous on  $Q_2$  except possibly “across”  $\Gamma$ .

*Proof.* (i) Let the supremum above be  $\kappa' \|f\|_\infty$ , with  $\kappa' < \kappa$  and let  $\gamma' \in (\gamma, 2)$ . Fix any  $k \in \mathbb{N}$  so that  $2^{-k} < \frac{r}{2}$ . For  $j, j' \in \mathbb{Z} \cap [0, 2^k)$  let  $a_{j,j'}$  be the average of  $f$  over the dyadic square  $H_k^j \cap V_k^{j'}$ , and let  $g(x, y) := a_{j,j'}$  for  $(x, y) \in H_k^j \cap V_k^{j'}$ . Note that  $\sum_{j,j'=0}^{2^k-1} a_{j,j'} = 0$  because  $\int_{Q_2} f \, dx dy = 0$ .

Just as in the proof of Lemma 2.3, we find that the intersection of any dyadic square  $Q \in G_k$  with  $T^{2k}(Q')$  for any dyadic square  $Q' \in G_k$  is a single dyadic rectangle of size  $2^{-k} \times 2^{-3k}$  (and the latter are disjoint for distinct  $Q' \in G_k$ ). Hence

$$\left| \int_Q f \circ T^{-2k} \, dx dy \right| \leq \left| \int_Q g \circ T^{-2k} \, dx dy \right| + \left| \int_Q (f - g) \circ T^{-2k} \, dx dy \right| \leq 0 + \kappa' \|f\|_\infty + 2^{1+(\gamma'-2)k} \|f\|_\infty$$

because  $\sum_{j,j'=0}^{2^k-1} a_{j,j'} = 0$ , and  $|f - g|$  is bounded by  $2\|f\|_\infty$  and exceeds  $\kappa'\|f\|_\infty$  on fewer than  $2^{\gamma'k}$  squares  $Q' \in G_k$  when  $k$  is large enough (namely those whose closures contain points from  $\Gamma$ ; then the intersection of  $Q$  with  $T^{2k}$  applied to these “bad” squares has measure less than  $2^{(\gamma'-4)k}$ ). Since  $\kappa' < \kappa$  and  $\gamma' < 2$  we find that

$$\left| \int_Q f \circ T^{-2k} dx dy \right| \leq \frac{\kappa' + \kappa}{2} \|f\|_\infty$$

for all large  $k$  and any  $Q \in G_k$  (a similar analysis shows the same for  $T^{-(2k+1)}$  in place of  $T^{-2k}$ ). The claim now follows easily by taking a large enough  $m$  (depending on  $\kappa - \kappa'$ ) and  $k := \lfloor \frac{n}{2} \rfloor$  for each  $n \in \mathbb{N}$ , so that the absolute value of the average of  $f \circ T^{-n}$  over any disc of radius  $2^{-(n-m)/2}$  cannot exceed the maximum of the absolute values of the averages of  $f \circ T^{-n}$  over all the squares from  $G_k$  by more than  $\frac{\kappa - \kappa'}{2} \|f\|_\infty$  (for instance, any  $m \geq 2 \log_2 \frac{8}{\kappa - \kappa'}$  works because then the squares from  $G_k$  that are fully contained in this disc cover the disc with the same center and radius  $2^{-(n-m)/2} - 2^{1-n/2}$ ).

(ii) This is similar to (i), but with  $\kappa'\|f\|_\infty$  replaced by  $C2^{-\alpha(k-1)}$  due to the hypothesis (with  $C$  the supremum in the statement of (ii)). Hence we get

$$\left| \int_Q f \circ T^{-2k} dx dy \right| \leq 2^{1 - \min\{\alpha, 2 - \gamma\}k} (C + \|f\|_\infty)$$

for all large  $k$  and any  $Q \in G_k$  (and again the same holds with  $T^{-(2k+1)}$  in place of  $T^{-2k}$ ). The claim now easily follows from the equivalence of the  $\dot{H}^{-1/2}$ -norm and the mix-norm (1.2).

(iii) Let  $f = \sum_{n \in \mathbb{N}^2} a_n \phi_n$  with  $\phi_n$  trigonometric polynomials such that  $\|\phi_n\|_\infty \leq 1$  and  $\|\nabla \phi_n\|_\infty \leq C|n|$ , as well as

$$C^{-1} \sum_{n \in \mathbb{N}^2} |a_n|^2 |n|^{2r} \leq \|f\|_{W^{r,2}}^2 \leq C \sum_{n \in \mathbb{N}^2} |a_n|^2 |n|^{2r}$$

for some  $C \geq 1$  and  $r \in \{0, \sigma\}$ . If  $f_k := \sum_{|n| \leq 2^{k/3}} a_n \phi_n$ , then the argument in (i) shows for any square  $Q \in G_k$ ,

$$\left| \int_Q f_k \circ T^{-2k} dx dy \right| \leq 2^{1-k} \sum_{|n| \leq 2^{k/3}} |a_n| C |n| \leq C 2^{1-k} \left( \sum_{|n| \leq 2^{k/3}} |a_n|^2 \right)^{1/2} 2^{2k/3} \leq C^2 2^{1-k/3} \|f\|_{H^\sigma}. \quad (2.3)$$

Equivalence of the  $\dot{H}^{-1/2}$ -norm and the mix-norm (1.2), together with

$$\|(f - f_k) \circ T^{-2k}\|_{\dot{H}^{-1/2}} \leq \|(f - f_k) \circ T^{-2k}\|_{L^2} = \|f - f_k\|_{L^2} \leq C 2^{-k\sigma/3} \|f\|_{H^\sigma} \quad (2.4)$$

and the fact that both claims again hold with  $T^{-(2k+1)}$  in place of  $T^{-2k}$ , now yields the functional mixing scale claim. The geometric mixing scale claim follows from this and Lemma A.1 in [33].  $\square$

## 2.4 Universal mixing and almost-universal exponential mixing on $Q_d$

Let us now consider the case  $x = (x_1, \dots, x_d) \in Q_d$  with  $d \geq 3$  instead of  $(x, y) \in Q_2$ .

**Definition 2.8.** Let  $Q_d := (0, 1)^d$ , and

$$Q'_d := Q_d \setminus \{x \in Q_d \mid 2^k x_i \in \mathbb{Z} \text{ for some } i \in \{1, \dots, d\} \text{ and } k \in \mathbb{N}\}.$$

For  $i \in \{1, \dots, d-1\}$ , let  $T_{d,i} : Q_d \rightarrow Q_d$  be given by

$$T_{d,i}(x) := (x_1, \dots, x_{i-1}, [T(x_i, x_d)]_1, x_{i+1}, \dots, x_{d-1}, [T(x_i, x_d)]_2),$$

where  $T$  is from Definition 2.1 and  $[T]_j$  is its  $j^{\text{th}}$  coordinate (with  $T$  extended to all of  $Q_2$  as before Lemma 2.3). Also let  $T_d : Q_d \rightarrow Q_d$  be given by

$$T_d := T_{d,d-1} \circ \dots \circ T_{d,1}.$$

For  $k \in \mathbb{N} \cup \{0\}$  and  $j_1, \dots, j_d \in \mathbb{Z} \cap [0, 2^k)$ , let

$$H_k^{j_d} := \left[ (0, 1)^{d-1} \times \left( \frac{j_d}{2^k}, \frac{j_d+1}{2^k} \right) \right] \cap Q'_d,$$

$$V_k^{j_1, \dots, j_{d-1}} := \left[ \left( \frac{j_1}{2^k}, \frac{j_1+1}{2^k} \right) \times \dots \times \left( \frac{j_{d-1}}{2^k}, \frac{j_{d-1}+1}{2^k} \right) \times (0, 1) \right] \cap Q'_d$$

be the *horizontal dyadic slabs* and *vertical dyadic strips* of width  $2^{-k}$ , respectively. Finally, let

$$G_{k,k'}^d := \left\{ H_k^{j_d} \cap V_{k'}^{j_1, \dots, j_{d-1}} \mid j_d \in \mathbb{Z} \cap [0, 2^k) \text{ and } j_1, \dots, j_{d-1} \in \mathbb{Z} \cap [0, 2^{k'}) \right\}$$

be the collection of all *dyadic boxes* of size  $2^{-k'} \times \dots \times 2^{-k'} \times 2^{-k}$  and let  $G_k^d := G_{k,k}^d$  be the collection of all *dyadic cubes* of size  $2^{-k} \times \dots \times 2^{-k}$ .

That is,  $T_{d,i}$  performs the transformation  $T$  in the  $x_i x_d$  plane while all other coordinates are preserved. Therefore, any dyadic box is being doubled in all  $d-1$  horizontal directions  $x_1, \dots, x_{d-1}$  and shrunk by a factor of  $2^{d-1}$  in the vertical direction  $x_d$  by each repeated application of  $T_d$  until it becomes a horizontal dyadic slab. Continued applications of  $T_d$  then always multiply the number of horizontal dyadic slabs by  $2^{d-1}$  and shrink their widths by a factor of  $2^{d-1}$ , while these slabs become fairly regularly distributed throughout  $Q_d$ . This immediately gives the following extension of Lemma 2.2.

**Lemma 2.9.** *Let  $k, k', l \in \mathbb{N}$ , and let  $j_d \in \mathbb{Z} \cap [0, 2^k)$  and  $j_1, \dots, j_{d-1} \in \mathbb{Z} \cap [0, 2^{k'})$ .*

(i) *If  $i \in \mathbb{Z} \cap [0, 2^{(d-1)l})$ , then  $T_d^l(H_k^{j_d}) \cap H_{(d-1)l}^i$  is a single horizontal dyadic slab of width  $2^{-(k+(d-1)l)}$ .*

(ii) *If  $l \leq k'$ , then  $T_d^l(H_k^{j_d} \cap V_{k'}^{j_1, \dots, j_{d-1}})$  is a single horizontal dyadic box from  $G_{k+(d-1)l, k'-l}^d$ .*

With this lemma in hand, the remaining mixing results in two dimensions and their proofs easily extend to any dimension.

**Lemma 2.10.** *Lemmas 2.3 and 2.4 hold for any  $d \geq 2$ , with  $Q_2$  and  $T$  replaced by  $Q_d$  and  $T_d$ . The same is true for Lemma 2.7 if we also replace  $\gamma < 2$  and  $2 - \gamma$  by  $\gamma < d$  and  $d - \gamma$  in (i,ii).*

### 3 Construction of the relevant velocity fields

Let us first consider the no-flow boundary conditions case. In this case we will construct time-periodic velocity fields realizing the maps  $T_d$  from the previous section as their flow maps at the time equal to a single period. Later we will show how to modify our construction when the boundary conditions are periodic.

Again we will start with the  $d = 2$  case. Note that the crucial map  $T$  is obtained by a  $90^\circ$  rotation of the right and left halves of  $Q_2$  in the clockwise and counter-clockwise directions, respectively, followed by a  $90^\circ$  counterclockwise rotation of  $Q_2$ . We therefore just need to find a divergence free field  $u$  that rotates a square by  $90^\circ$ , and then easily adjust it so it rotates rectangles. This was done by Yao and the second author in [33], but we will redo and slightly alter their construction here as we also want to show that  $u(\cdot, t) \in W^{s,p}(Q_2)$  for some  $s > 1$ . Additionally, since the rotating flow will be time-independent, we will omit  $t$  below.

#### 3.1 Rotating velocity fields on $Q_2$

For  $\alpha \geq 0$ , let us consider the stream function  $\psi_\alpha : Q_2 \rightarrow \mathbb{R}$  given by

$$\psi_\alpha(x, y) := 2^\alpha \frac{\sin(\pi x) \sin(\pi y)}{(\sin(\pi x) + \sin(\pi y))^\alpha} \in (0, 1]$$

It is easy to show that  $\log \psi_\alpha$  is concave (see Lemma 3.1(i) below), so the level sets of  $\psi_\alpha$  are curves which foliate  $Q_2$ . We also define the quantity

$$T_{\psi_\alpha}(r) := \int_{\{\psi_\alpha=r\}} \frac{1}{|\nabla \psi_\alpha|} d\sigma,$$

which is the time it takes a particle from the level curve  $\{\psi_\alpha = r\}$  to traverse this level curve if it moves with the (divergence-free) velocity  $v_\alpha := \nabla^\perp \psi_\alpha$ . As in [33], we will next find another function  $\psi^\alpha$  with the same level sets as  $\psi_\alpha$  but with the ‘‘period’’  $T_{\psi^\alpha}(r)$  independent of  $r$  (e.g., equal to 1). Four-fold rotational symmetry of  $\psi_\alpha$  around  $(\frac{1}{2}, \frac{1}{2})$  will then show that the flow  $v^\alpha = \nabla^\perp \psi^\alpha$  rotates  $Q_2$  by  $90^\circ$  after a quarter of the period  $T_{\psi^\alpha}$ .

Let us start with some properties of  $\psi_\alpha$ .

**Lemma 3.1.** *For any  $\alpha \in (0, 1)$  and  $k \in \mathbb{N} \cup \{0\}$  there is  $C_{\alpha,k} > 0$  such that the following hold.*

- (i)  $\log \psi_\alpha$  is concave, so the super-level sets of  $\psi_\alpha$  are convex.
- (ii)  $|D^k \psi_\alpha(x, y)| \leq C_{\alpha,k} (\sin(\pi x) + \sin(\pi y))^{2-\alpha-k}$  for each  $(x, y) \in Q_2$ .
- (iii)  $T_{\psi_\alpha} \in C_{\text{loc}}^2((0, 1))$ , and for  $k = 0, 1, 2$  and each  $r \in (0, 1)$  we have

$$\left| T_{\psi_\alpha}^{(k)}(r) \right| \leq C_{\alpha,k} (1-r)^{-k} (1+r^{\frac{\alpha}{2-\alpha}-k}).$$

*Proof.* (i) is a direct calculation (see the appendix) and (ii) is immediate from the definition.

(iii) Since  $\psi$  is smooth and  $|\nabla \psi|$  is non-zero away from  $\partial Q_2 \cup \{(\frac{1}{2}, \frac{1}{2})\}$ , we only need to consider  $r$  close to 0 and 1.

For all  $r$  close to 1 we have

$$\sqrt{1-r} \approx |\{\psi_\alpha = r\}| \approx |\nabla\psi_\alpha(x, y)| \approx \left| (x, y) - \left(\frac{1}{2}, \frac{1}{2}\right) \right| \quad \text{for all } (x, y) \in \{\psi_\alpha = r\}, \quad (3.1)$$

where  $A \approx B$  means that there exists a constant  $C$ , independent of  $(x, y)$  as well as of  $r$  near 1, such that  $\frac{1}{C}A \leq B \leq CA$ . Indeed, this follows from  $\psi_\alpha$  having a maximum at  $(\frac{1}{2}, \frac{1}{2})$  with  $D^2\psi_\alpha(\frac{1}{2}, \frac{1}{2})$  a non-zero multiple of the identity matrix, and from uniform bounds on the third derivatives of  $\psi_\alpha$  near  $(\frac{1}{2}, \frac{1}{2})$ . It implies, in particular, uniform boundedness of  $T_{\psi_\alpha}$  near 1, which yields the claim for  $k = 0$  and all  $r$  near 1.

Let us next study the derivatives of  $T_{\psi_\alpha}$  near 1. As in [33], let us write

$$T_{\psi_\alpha}(r) = \int_{\{\psi_\alpha=r\}} \frac{1}{|\nabla\psi_\alpha|} d\sigma = \int_{\{\psi_\alpha=r\}} (-n) \cdot \frac{\nabla\psi_\alpha}{|\nabla\psi_\alpha|^2} d\sigma = - \int_{\psi_\alpha>r} \nabla \cdot \left( \frac{\nabla\psi_\alpha}{|\nabla\psi_\alpha|^2} \right) dx dy.$$

Thus

$$T'_{\psi_\alpha}(r) = \int_{\{\psi_\alpha=r\}} \frac{1}{|\nabla\psi_\alpha|} \nabla \cdot \left( \frac{\nabla\psi_\alpha}{|\nabla\psi_\alpha|^2} \right) d\sigma.$$

From this, (3.1), and (ii) for  $k = 2$  we now obtain for all  $r$  near 1,

$$|T'_{\psi_\alpha}(r)| \leq C_\alpha(1-r),$$

with some constant  $C_\alpha$  (depending only on  $\alpha$ ). Similarly we obtain

$$T''_{\psi_\alpha}(r) = \int_{\{\psi_\alpha=r\}} \frac{1}{|\nabla\psi_\alpha|} \nabla \cdot \left( \frac{\nabla\psi_\alpha}{|\nabla\psi_\alpha|^2} \nabla \cdot \left( \frac{\nabla\psi_\alpha}{|\nabla\psi_\alpha|^2} \right) \right) d\sigma,$$

and then for all  $r$  near 1,

$$|T''_{\psi_\alpha}(r)| \leq C'_\alpha(1-r)^2,$$

with some constant  $C'_\alpha$ . Hence the claim for  $k = 1, 2$  and all  $r$  near 1 also follows.

Let us now consider  $r$  near 0. A simple computation gives the lower bound

$$|\nabla\psi_\alpha(x, y)| \geq c_\alpha(\sin(\pi x) + \sin(\pi y))^{1-\alpha}$$

for all such  $r$  and all  $(x, y)$  with  $\psi_\alpha(x, y) = r$ . In particular,

$$T_{\psi_\alpha}(r) \leq \int_{\{\psi_\alpha=r\}} \frac{d\sigma}{c_\alpha(\sin(\pi x) + \sin(\pi y))^{1-\alpha}}.$$

Since  $\psi_\alpha$  and  $\sin(\pi x) + \sin(\pi y)$  have all the symmetries of  $Q_2$ , it suffices to consider 8 times the last integral restricted to the part of the curve  $\{\psi_\alpha = r\}$  between the lines  $y = x$  and  $x = \frac{1}{2}$  (where  $\sin(\pi x) \geq \sin(\pi y)$ ). For  $r$  near 0 and  $x$  near 0 such that  $\psi_\alpha(x, x) = r$  we have  $\sin(\pi x)^{2-\alpha} = r$ , so that  $x \geq \frac{1}{10}r^{\frac{1}{2-\alpha}}$ . It follows that for some  $C_\alpha$  and all  $r$  near 0 we have

$$T_{\psi_\alpha}(r) \leq C_\alpha \int_{0.1r^{\frac{1}{2-\alpha}}}^{\frac{1}{2}} \frac{dx}{x^{1-\alpha}} \leq \frac{C_\alpha}{\alpha},$$

which yields the claim for  $k = 0$  and all  $r$  near 0.

We now proceed as in the case  $r$  near 1. From

$$T'_{\psi_\alpha}(r) = \int_{\{\psi_\alpha=s\}} \frac{1}{|\nabla\psi_\alpha|} \nabla \cdot \left( \frac{\nabla\psi_\alpha}{|\nabla\psi_\alpha|^2} \right) d\sigma$$

and (ii) we obtain for  $r$  near 0 (with some constants  $C_\alpha, C'_\alpha, C''_\alpha$ ),

$$|T'_{\psi_\alpha}(r)| \leq \int_{\{\psi_\alpha=r\}} C_\alpha \frac{(\sin(\pi x) + \sin(\pi y))^{-\alpha}}{|\nabla\psi_\alpha|^3} d\sigma \leq C'_\alpha \int_{0.1r}^{\frac{1}{2}} \frac{1}{(\sin(\pi x))^{3-2\alpha}} dx \leq C''_\alpha r^{\frac{\alpha}{2-\alpha}-1}.$$

Similarly, from the formula for  $T''_{\psi_\alpha}$ , lower bound on  $|\nabla\psi_\alpha|$ , and (ii) we obtain for  $r$  near 0,

$$|T''_{\psi_\alpha}(r)| \leq C_\alpha \int_{0.1r}^{\frac{1}{2}} \frac{dx}{(\sin(\pi x))^{5-3\alpha}} \leq C'_\alpha r^{\frac{\alpha}{2-\alpha}-2}.$$

Hence the claim for  $k = 1, 2$  and all  $r$  near 0 also follows, and we are done. We note that one can continue to higher derivatives and, in particular, obtain that  $T_{\psi_\alpha} \in C^\infty_{\text{loc}}((0, 1))$ .  $\square$

We now define the function  $\psi^\alpha$  with the same level sets as  $\psi_\alpha$  but with  $T_{\psi_\alpha}(r)$  independent of  $r$ . Because we want to address the question of its (fractional) regularity, let us first define the fractional Sobolev spaces (see, e.g., [23, 24]).

**Definition 3.2.** For  $\gamma \in (0, 1)$  and for  $f \in C^1(Q_2)$ , let

$$\Lambda^\gamma(f)(x, y) = \int_{Q_2} \frac{f(x, y) - f(x', y')}{|(x, y) - (x', y')|^{2+\gamma}} dx' dy'.$$

For  $p \geq 1$ , define the  $W^{\gamma,p}$  norm of  $f$  by  $\|f\|_{W^{\gamma,p}(Q_2)} := \|f\|_{L^p(Q_2)} + \|\Lambda^\gamma(f)\|_{L^p(Q_2)}$ . The space  $W^{\gamma,p}(Q_2)$  is the completion of  $C^1(Q_2)$  with respect to this norm. We also let  $W^{0,p}(Q_2) := L^p(Q_2)$  and for  $\gamma \geq 1$  we say that  $f \in W^{\gamma,p}(Q_2)$  if and only if  $f, Df, \dots, D^{\lfloor \gamma \rfloor} f \in W^{\gamma-\lfloor \gamma \rfloor,p}(Q_2)$ .

**Lemma 3.3.** For any  $\alpha \in (0, 1)$  there exists  $\psi^\alpha \in W^{1,\infty}(Q_2) \cap W^{2,\infty}_{\text{loc}}(Q_2)$  with the following properties.

- (i)  $\psi^\alpha$  has the same level sets as  $\psi_\alpha$ , with  $\psi^\alpha > 0$  on  $Q_2$  and  $\psi^\alpha = 0$  on  $\partial Q_2$ .
- (ii)  $T_{\psi^\alpha}(r) = 1$  for all  $r \in (0, \|\psi^\alpha\|_{L^\infty})$ .
- (iii) For any  $\gamma \in [0, 1]$  we have  $\psi^\alpha \in W^{2+\gamma,p}(Q_2)$  whenever  $1 \leq p < (\max\{\frac{\alpha}{2}, \frac{2-2\alpha}{2-\alpha}\} + \gamma)^{-1}$ .

*Remark.* Optimal regularity in (iii) is thus obtained when  $\frac{\alpha}{2} = \frac{2-2\alpha}{2-\alpha}$ , that is, by taking  $\alpha_* := 3 - \sqrt{5}$ . This yields  $\psi^{\alpha_*} \in W^{2+\gamma,p}(Q_2)$  for all  $\gamma \in [0, \frac{\sqrt{5}-1}{2})$  and  $p \in [1, \frac{2}{3-\sqrt{5}+2\gamma})$ .

*Proof.* Let

$$\psi^\alpha(x, y) := \int_0^{\psi_\alpha(x, y)} T_{\psi_\alpha}(r) dr, \tag{3.2}$$

from which (i) follows immediately. Since  $\nabla\psi^\alpha(x, y) = T_{\psi_\alpha}(\psi_\alpha(x, y))\nabla\psi_\alpha(x, y)$ , Lemma 3.1 yields  $\psi^\alpha \in W^{1,\infty}$  and we also obtain (ii). It remains to show (iii).

We see that

$$|D^2\psi^\alpha| \leq |D^2\psi_\alpha|T_{\psi_\alpha}(\psi_\alpha) + |\nabla\psi_\alpha|^2|T'_{\psi_\alpha}(\psi_\alpha)|,$$

and

$$|D^3\psi^\alpha| \leq |D^3\psi_\alpha|T_{\psi_\alpha}(\psi_\alpha) + 3|D^2\psi_\alpha||\nabla\psi_\alpha||T'_{\psi_\alpha}(\psi_\alpha)| + |\nabla\psi_\alpha|^3|T''_{\psi_\alpha}(\psi_\alpha)|.$$

Lemma 3.1 and (3.1) show that away from  $\partial Q_2$ , the only unbounded terms may be the second and third term in the estimate for  $D^3\psi^\alpha$ , and they are both bounded above by  $C_\alpha|(x, y) - (\frac{1}{2}, \frac{1}{2})|^{-1}$  there. Thus  $D^2\psi^\alpha$  is bounded away from  $\partial Q_2$  (proving  $\psi^\alpha \in W_{\text{loc}}^{2,\infty}$ ), while  $D^3\psi^\alpha$  is in  $L^p$  for any  $p < 2$  there. Sobolev embedding now shows that  $D^2\psi^\alpha \in W^{\gamma,p}$  away from  $\partial Q_2$  whenever  $p < \frac{2}{\gamma}$ .

It therefore suffices to consider the neighborhood of  $\partial Q_2$ . From Lemma 3.1 we see that there

$$|D^2\psi^\alpha(x, y)| \leq C_\alpha \left( |\sin(\pi x) + \sin(\pi y)|^{-\alpha} + |\psi_\alpha|^{\frac{2\alpha-2}{2-\alpha}} \right) \leq 2C_\alpha |\sin(\pi x) \sin(\pi y)|^{-\max\{\frac{\alpha}{2}, \frac{2-2\alpha}{2-\alpha}\}}$$

for some  $C_\alpha$ . Similarly, near  $\partial Q_2$  we obtain

$$|D^3\psi^\alpha(x, y)| \leq C_\alpha |\sin(\pi x) \sin(\pi y)|^{-\max\{\frac{\alpha+1}{2}, \frac{\alpha}{2} + \frac{2-2\alpha}{2-\alpha}, \frac{4-3\alpha}{2-\alpha}\}} = C_\alpha |\sin(\pi x) \sin(\pi y)|^{-\frac{4-3\alpha}{2-\alpha}}$$

because  $\alpha \in [0, 1]$ . Corollary 4.2 applied to  $D^2\psi^\alpha$  now yields

$$|\Lambda^\gamma D^2\psi^\alpha(x, y)| \leq C_{\alpha,\gamma,\epsilon} |\sin(\pi x) \sin(\pi y)|^{-\max\{\frac{\alpha}{2}, \frac{2-2\alpha}{2-\alpha}\} - \gamma - \epsilon}$$

for all  $\epsilon > 0$  and  $(x, y)$  near  $\partial Q_2$ , with some  $C_{\alpha,\gamma,\epsilon}$ . Direct integration and the estimate away from  $\partial Q_2$  (with the range from (iii) included in  $p < \frac{2}{\gamma}$ ) finish the proof of (iii).  $\square$

### 3.2 Velocity fields realizing $T$ and $T_d$ as their flow maps (no-flow case)

According to the remark after Lemma 3.3, let us take  $\alpha_* := 3 - \sqrt{5}$  and let  $\psi := \frac{1}{2}\psi^{\alpha_*}$ . Then Lemma 3.3 shows that the velocity field  $v := \nabla^\perp \psi = (-\psi_y, \psi_x)$  belongs to  $W^{s,p}(Q_2)$  for any  $s < \frac{1+\sqrt{5}}{2}$  and  $p \in [1, \frac{2}{2s+1-\sqrt{5}})$  and it rotates  $Q_2$  clockwise by  $90^\circ$  in time  $\frac{1}{2}$ . Similarly, if  $\phi(x, y) := \frac{1}{2}\text{sgn}(2x-1)\psi(\{2x\}, y)$  (with  $\{x\}$  the fractional part of  $x$ ) and  $w := \nabla^\perp \phi$ , then  $w$  “rotates” the left and right halves of  $Q_2$  counterclockwise and clockwise by  $90^\circ$  in time  $\frac{1}{2}$ , respectively (by “rotation” we mean the affine map that is a bijection on the rectangle and permutes its corners in the indicated direction). It therefore follows that

$$u(x, y, t) := \begin{cases} w(x, y) & \{t\} \in [0, \frac{1}{2}) \\ -v(x, y) & \{t\} \in [\frac{1}{2}, 1) \end{cases} \quad (3.3)$$

is time 1-periodic, satisfies the no-flow boundary condition, and its flow map at time 1 is  $T$  from (2.1). Moreover,  $w$  also belongs to  $W^{s,p}(Q_2)$  for any  $s < \frac{1+\sqrt{5}}{2}$  and  $p \in [1, \frac{2}{2s+1-\sqrt{5}})$ . Indeed,  $\phi$  is Lipschitz continuous across  $\{x = \frac{1}{2}\}$  since  $\psi$  vanishes at  $\{x = 1\}$ . Also,  $\phi_x$  is clearly continuous across  $\{x = \frac{1}{2}\}$  while the same is true for  $\phi_y$  because it vanishes there. This and  $\psi \in W^{2,p}(Q_2)$  now yield  $\phi \in W^{2,p}(Q_2)$ , and Corollary 4.2 applied to  $D^2\phi$  proves the claim.

Of course, the application of  $v$  does nothing for mixing and one could instead replace it by  $R^{-1}w(R(x, y))$  for times  $t$  with  $\{t\} \in [\frac{1}{2}, 1)$ , where  $R$  is the counterclockwise rotation of  $Q_2$  by



90°. Then the time- $\frac{1}{2}$  and time-1 flow maps of the new 1-periodic flow will be  $R^{-1}T$  and  $R^{-2}T^2$ , respectively, but an analog of Lemma 2.2 holds in this case and so do other results in Section 2.

It is now also clear how to construct the relevant velocity fields in higher dimensions. They will have time periods  $d-1$  and have the above  $u$  acting in only 2 variables (specifically, 1 and  $d$ , 2 and  $d, \dots, d-1$  and  $d$ ) on each time interval with integer endpoints. These time-periodic and time-piecewise-constant fields obviously again belong to  $W^{s,p}(Q_d)$  for all  $(s, p)$  from Theorem 1(i). Of course, we can then scale them in time and value by  $d-1$  to obtain time period 1.

### 3.3 Velocity fields on $\mathbb{T}^d$ (periodic case)

In the case of periodic boundary conditions we cannot use the flow  $u$  from the previous subsection because it is not  $W^{s,p}$  across  $\partial(0,1)^2$ . A way to fix this is to extend the stream functions  $\psi, \phi$  oddly in both  $x$  and  $y$  onto  $(-1,1)^2$  and then map them back onto  $Q_2$  (or rather  $\mathbb{T}^2$ ). So we let

$$\psi'(x, y) := \frac{1}{2}\psi(2x-1, 2y-1)$$

(and similarly for  $\phi'$ ) where  $\psi$  is the doubly-odd extension. If we now let  $v' := \nabla^\perp \psi'$  and  $w' := \nabla^\perp \phi'$ , then these fields again belong to  $W^{s,p}(\mathbb{T}^2)$  for any  $s < \frac{1+\sqrt{5}}{2}$  and  $p \in [1, \frac{2}{2s+1-\sqrt{5}})$  by the argument from the previous subsection. Now each of the four square cells from  $G_1$  (of side length  $\frac{1}{2}$ ) is left invariant by  $v'$  and  $w'$ . Therefore, a flow like (3.3) can mix exponentially quickly within each of these four cells, but there is no “exchange” between the cells.

We remedy this problem by instead letting

$$u'(x, y, t) := \begin{cases} (\frac{1}{2}, \frac{1}{2}) & \{t\} \in [0, \frac{1}{2}) \\ w'(x, y) & \{t\} \in [\frac{1}{2}, \frac{3}{4}) \\ -v'(x, y) & \{t\} \in [\frac{3}{4}, 1), \end{cases} \quad (3.4)$$

which satisfies periodic boundary conditions and at each time belongs to all the  $W^{s,p}(\mathbb{T}^2)$  spaces above. We denote by  $T'$  the time-1 flow map of  $u'$ . Let us also define

$$\tilde{T}(x, y) = \begin{cases} -(2x, \frac{y}{2}) + (1, 1) & x \in (0, \frac{1}{2}), \\ (2x, \frac{y}{2}) + (-1, 0) & x \in (\frac{1}{2}, 1), \end{cases}$$

a mapping closely related to  $T$  from (2.1), and recall that  $R^2$  is the 180° degree rotation of  $Q_2$ .

Consider now any dyadic rectangle  $H_k^j \cap V_{k'}^{j'} \in G_{k,k'}$  with  $k \geq 2$ . If  $k' \geq 2$ , then its translation by  $(\frac{1}{4}, \frac{1}{4})$  that occurs under the action of  $u'$  over time interval  $[0, \frac{1}{2})$  transforms  $H_k^j \cap V_{k'}^{j'}$  into another element of  $G_{k,k'}$  that is fully contained inside one of the four cells from  $G_1$ . If this cell is  $(0, \frac{1}{2}) \times (0, \frac{1}{2})$ ,  $(\frac{1}{2}, 1) \times (0, \frac{1}{2})$ ,  $(0, \frac{1}{2}) \times (\frac{1}{2}, 1)$ , or  $(\frac{1}{2}, 1) \times (0, \frac{1}{2})$ , respectively, then the above action of  $w'$  and  $v'$  over time interval  $[\frac{1}{2}, 1)$  on this cell is the same as the action of  $T$ ,  $R^2T$ ,  $\tilde{T}$ , or  $R^2\tilde{T}$  on  $Q_2$ , respectively, up to scaling by a factor of 2 (due to the factor  $\frac{1}{2}$  in the definition of  $\psi'$  and  $\phi'$ ) and a translation that transform this cell into  $Q_2$  (recall also that  $R$  is the counterclockwise rotation of  $Q_2$  by 90°). Therefore,  $T'(H_k^j \cap V_{k'}^{j'}) \in G_{k+1, k'-1}$ .

If instead  $k' = 1$ , then the translation by  $(\frac{1}{4}, \frac{1}{4})$  splits the dyadic rectangle (of length  $\frac{1}{2}$ ) between two horizontally adjacent cells from  $G_1$ , but a simple computation (specifically, that

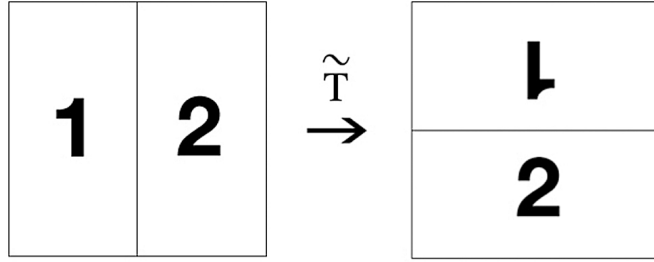


Figure 2: The mapping  $\tilde{T}$ .

$T(H_l^i \cap V_1^a) = R^2 T(H_l^i \cap V_1^{1-a}) \in G_{i+1,0}$  for  $a = 0, 1$  and any  $l, i$ , and that this also holds with  $\tilde{T}$  in place of  $T$ ) shows that again  $T'(H_k^j \cap V_{k'}^{j'}) \in G_{k+1,0}$ . It follows that Lemma 2.2(ii) continues to hold for  $T'$  in place of  $T$  as long as  $k \geq 2$ .

If now  $H_k^j \in G_{k,0}$  with  $k \geq 2$ , then  $T'(H_k^j)$  consists of two horizontal dyadic strips of width  $2^{-(k+1)}$  located either in  $H_2^0$  and  $H_2^1$ , or in  $H_2^2$  and  $H_2^3$ . In either case, a simple computation (also using  $T(H_l^i \cap V_1^a) = R^2 T(H_l^i \cap V_1^{1-a}) \in G_{i+1,0}$ ) shows that  $(T')^2(H_k^j)$  will then consist of four horizontal dyadic strips of width  $2^{-(k+2)}$ , one contained in each of  $H_2^0$ ,  $H_2^1$ ,  $H_2^2$ , and  $H_2^3$  (which is now also a setup preserved by a translation by  $(\frac{1}{4}, \frac{1}{4})$ ). A repeated application of  $T'$  then shows that Lemma 2.2(i) continues to hold for  $T'$  in place of  $T$  as long as  $k, l \geq 2$ .

This modified Lemma 2.2 is of course sufficient for the rest of the analysis in Section 2. This includes the case  $d \geq 3$ , where we construct the relevant mapping  $T'_d$  using  $T'$  in the same way we constructed  $T_d$  using  $T$  (and then even both parts of Lemma 2.9 will hold for  $T'_d$  in place of  $T_d$  as long as  $k \geq 2$ ). Thus there is again a time 1-periodic and time-piecewise-constant vector field that belongs to  $W^{s,p}(\mathbb{T}^d)$  for all  $(s, p)$  from Theorem 1(i) and whose time-1 flow map is  $T'_d$ .

### 3.4 Non-integer times and the proof of Theorem 1

Theorem 1(ii) immediately follows from Lemma 2.6, with  $T_n$  the flow map of the vector field in question at time  $n$ .

Theorem 1(i) for only integer times follows from the above constructions and from Lemmas 2.4 and 2.7(iii) (as well as their multidimensional analogs in Lemma 2.10 and their periodic boundary conditions analogs from the previous subsection).

We will next extend the claim in Theorem 1(i) of  $u$  being a universal mixer in both senses to all times (we only do it on  $Q_2$ , the other cases are analogous). This uses the fact that the time-independent flows  $w$  and  $-v$  from (3.3) are bounded, due to Lemma 3.1. This implies that their actions are continuous in time on  $L^2$ , that is, if  $\{\Phi_{t;z}\}_{t \geq 0}$  are their flow maps (with  $z \in \{v, w\}$ ) and  $g \in L^2(Q_2)$ , then both maps  $t \mapsto g \circ \Phi_{t;z}^{-1}$  belong to  $C([0, \infty), L^2(Q_2))$ . Let also  $\{\Phi_t\}_{t \in [0,1]}$  be the flow maps for the (time-dependent) flow  $u$  from (3.3). Note that  $\Phi_{-t} \neq \Phi_t^{-1}$

due to time dependence, and in fact we have

$$\Phi_t = \begin{cases} \Phi_{t;w} & t \in [0, \frac{1}{2}], \\ \Phi_{t-1/2;v} \circ \Phi_{1/2;w} & t \in (\frac{1}{2}, 1], \\ \Phi_{t;v} & t \in [-\frac{1}{2}, 0), \\ \Phi_{t+1/2;w} \circ \Phi_{-1/2;v} & t \in [-1, -\frac{1}{2}). \end{cases}$$

Let a bounded mean-zero  $\rho$  solve (1.1) with the flow  $u$  on  $Q_2$  from (3.3) and let  $f := \rho(\cdot, 0)$ . Then  $\rho(\cdot, n+s) = f \circ T^{-n} \circ \Phi_s^{-1}$  for each  $n \in \mathbb{N}$  and  $s \in [0, 1]$ . Given any  $g \in L^2(Q_2)$  and  $\varepsilon > 0$ , let  $N \in \mathbb{N}$  be such that for  $z \in \{v, w\}$  we have

$$\sup_{|s| \leq 1/N} \|g \circ \Phi_{s;z}^{-1} - g\|_{L^2} \leq \varepsilon.$$

Let  $g_j := g \circ \Phi_{-j/N}^{-1} \in L^2(Q_2)$  for  $j = 0, \dots, N-1$ . The weak- $L^2$  convergence in the proof of Lemma 2.4 shows that  $\{\rho(\cdot, \frac{m}{N})\}_{n \geq 0}$  also converges weakly to 0 in  $L^2(Q_2)$ , because if  $m = nN + j$  with  $j \in \{0, \dots, N-1\}$ , then incompressibility of  $u$  shows that

$$\int_{Q_2} \rho\left(\cdot, \frac{m}{N}\right) g dx = \int_{Q_2} \left(f \circ T^{-n} \circ \Phi_{j/N}^{-1}\right) g dx = \int_{Q_2} (f \circ T^{-n}) g_j dx \quad (\rightarrow 0 \text{ as } m \rightarrow \infty).$$

Now if  $t = \frac{m}{N} + s$  with  $s \in [0, \frac{1}{N})$ , then

$$\int_{Q_2} \rho(\cdot, t) g dx = \int_{Q_2} \rho\left(\cdot, \frac{m}{N}\right) (g \circ \Phi_{-s;z}^{-1}) dx,$$

where  $z = w$  if  $\lfloor \frac{m}{N} \rfloor < \frac{N}{2}$  and  $z = v$  otherwise (recall that  $N$  is even). But then

$$\left| \int_{Q_2} \rho(\cdot, t) g dx \right| \leq \left| \int_{Q_2} \rho\left(\cdot, \frac{m}{N}\right) g dx \right| + \left\| \rho\left(\cdot, \frac{m}{N}\right) \right\|_{L^2}^{1/2} \|g \circ \Phi_{-s;z}^{-1} - g\|_{L^2}^{1/2}.$$

The first term on the right-hand side converges to 0 as  $t \rightarrow \infty$  (because then  $m \rightarrow \infty$ ), while the second is at most  $\|\rho(\cdot, 0)\|_{L^2}^{1/2} \varepsilon^{1/2}$ . Since  $g \in L^2(Q_2)$  and  $\varepsilon > 0$  were arbitrary, asymptotic mixing (i.e.,  $u$  being a universal mixer) in the functional sense in Theorem 1(i) is proved. The same claim in the geometric sense again follows from this and Lemma A.1 in [33].

It remains to extend the exponential mixing claim (i.e., of  $u$  being an almost universal exponential mixer) in Theorem 1(i) from integer times to all times. We show this after stating the following lemma, which is of independent interest. It shows that the flow maps of our square-rotating flow for  $t \in [0, 1]$  (so until the square is rotated by  $90^\circ$ ) are uniformly Hölder continuous. (Note that since the flow is not Lipschitz, Hölder continuity of the flow maps is not obvious.) For the sake of convenience, we switch to the notation  $x = (x_1, x_2)$  on  $Q_2$  in the rest of this subsection.

**Lemma 3.4.** *Let  $\alpha \in (0, 1)$  and consider the velocity field  $u' := \nabla^\perp \psi^\alpha$  on  $Q_2$ , with  $\psi^\alpha$  from (3.2). Then the flow maps  $\Phi'(\cdot, t)$  for  $u'$ , given by*

$$\frac{d}{dt} \Phi'(x, t) = u'(\Phi'(x, t)) \quad \text{and} \quad \Phi'(x, 0) = x,$$

*are Hölder continuous uniformly in  $t \in [0, 1]$  (with some exponent  $\beta_\alpha > 0$  and constant  $C_\alpha$ ).*

Let us now extend the claim in Theorem 1(i) of  $u$  being an almost universal exponential mixer to all times. Consider any  $\sigma > 0$  and  $\rho(\cdot, 0) \in H^\sigma(Q_2)$ . Let a bounded mean-zero  $\rho$  solve (1.1) with the flow  $u$  on  $Q_2$  from (3.3) and let  $f := \rho(\cdot, 0)$ . Then,  $\rho(\cdot, n+t) = f \circ T^{-n} \circ \Phi_t^{-1}$  for  $n \in \mathbb{N}$  and  $t \in [0, \frac{1}{2}]$ , where now  $\Phi_t$  is the flow map at time  $t$  for the time-independent flow  $w$  from (3.3) (denoted  $\Phi_{t,w}$  above). By Lemma 3.4, these flow maps are uniformly Hölder continuous (with some exponent  $\beta > 0$  and constant  $C$ ) separately on the left and right halves of  $Q_2$ . Assume  $\beta \leq \frac{2}{3}$  without loss.

Let  $f_k$  be from the proof of Lemma 2.7(iii) and let  $Q' \in G_{\lfloor \delta k \rfloor}$  be arbitrary, with some  $\delta > 0$ . We will show that estimate (2.3) for squares  $Q \in G_k$  and the Hölder bound on  $\Phi_t$  together yield an analogous estimate for  $f_k \circ T^{-2k} \circ \Phi_t^{-1}$  on  $Q'$  with any  $t \in [0, \frac{1}{2}]$ . Indeed, assume without loss that  $Q'$  belongs to the left half  $L$  of  $Q_2$  (on which the flow maps are Hölder continuous) and let  $Q_t := \Phi_t^{-1}(Q') \subseteq L$ . Let  $P_t \subseteq Q_t$  be the union of all the squares from  $G_k$  that are fully contained in  $Q_t$ . The Hölder bound shows that if  $x \in Q_t \setminus P_t$ , then  $\text{dist}(\Phi_t(x), \partial Q') \leq C2^{\beta(1-k)}$ . That is,  $|Q_t \setminus P_t| \leq 4C2^{\beta-\lfloor \delta k \rfloor - \beta k}$ . Picking  $\delta \leq \frac{\beta}{2}$  shows that  $|Q_t \setminus P_t| \leq C_1 2^{-\beta k/2} |Q_t|$  with some constant  $C_1$ . This, (2.3), and  $\Phi_t$  preserving measure now show (with some constants  $C_2, C_3$ )

$$\left| \int_{Q'} f_k \circ T^{-2k} \circ \Phi_t^{-1} dx \right| = \left| \int_{Q_t} f_k \circ T^{-2k} dx \right| \leq C_2 2^{-k/3} \|f\|_{H^\sigma} + C_1 2^{-\beta k/2} \|f\|_{L^\infty} \leq C_3 2^{-\beta k/2} \|f\|_{H^\sigma}$$

(and the same bound again holds with  $T^{-(2k+1)}$  in place of  $T^{-2k}$ ). The claim in (2.4) continues to hold if we replace  $T^{-2k}$  by  $T^{-2k} \circ \Phi_t^{-1}$ , which then yields exponential decay of  $\|\rho(\cdot, t)\|_{\dot{H}^{-1/2}}$  when restricted to  $t \geq 0$  with  $\{t\} \in [0, \frac{1}{2}]$ , with a  $\sigma$ -dependent rate (this again uses equivalence of the  $\dot{H}^{-1/2}$ -norm and the mix norm). To handle the times with  $\{t\} \in (\frac{1}{2}, 1)$ , it suffices to notice that  $\Phi_{1/2}$  simply rotates the two halves of  $Q_2$  by  $90^\circ$  (so it is Lipschitz with constant 2), and the flow  $-v$  from (3.3) satisfies the same Hölder estimate as  $w$ . Hence the above argument applies again and almost universal exponential mixing in the functional sense on  $Q_2$  follows. The same claim in the geometric sense again follows from this and Lemma A.1 in [33]. The other cases of  $Q_d$  and  $\mathbb{T}^d$  are again analogous because the flows involved are essentially the same as on  $Q_2$ .

To finish the proof of Theorem 1(i), it now only remains to prove Lemma 3.4.

*Proof of Lemma 3.4.* Let us first consider  $u := \nabla^\perp \psi_\alpha$  instead, and prove the claim about the corresponding flow map  $\Phi$  and with  $t$  in any interval  $[0, \tau]$ . In fact, for the sake of convenience, let us instead consider the equivalent case of

$$u(x, y) := \nabla^\perp \tilde{\psi}_\alpha(x) := \nabla^\perp \frac{\sin x_1 \sin x_2}{(\sin x_1 + \sin x_2)^\alpha}$$

in the square  $\pi Q_2$ .

Since  $u$  is Lipschitz away from the corners, it suffices to consider the case of points  $x, y \in \pi Q_2$  such that both trajectories  $\{\Phi(x, s)\}_{s \in [0, t]}$  and  $\{\Phi(y, s)\}_{s \in [0, t]}$  remain within some distance  $0 < \delta \ll 1$  from a corner of  $\pi Q_2$ . By symmetry, this corner can be assumed to be the origin, so in particular  $x, y \in A_\delta := B_\delta(0) \cap \pi Q_2$ . We note that  $\delta$  will depend on  $\alpha$  (then any constants depending on  $\delta$  will really only depend on  $\alpha$ ) and will be such that  $\inf_{z \in (0, \delta)} (\frac{\sin z}{z} + \cos z)$  is sufficiently close to 2. We will split the argument into several cases.

**Case 1.** We will first consider points  $x, y \in A_\delta$  near the  $x_2$  axis, specifically, in the region  $S_\delta := A_\delta \cap \{x_1 \leq \delta x_2\}$ . Their trajectories will be therefore moving primarily upwards and slightly

to the left while they remain in  $A_\delta$  (and therefore they will also remain in  $S_\delta$  during this time), with the vertical component of  $u(x)$  being

$$u_2(x) = \cos x_1 \sin x_2 \frac{(1 - \alpha) \sin x_1 + \sin x_2}{(\sin x_1 + \sin x_2)^{1+\alpha}}.$$

We therefore have

$$u_2(x) \geq (1 - 4\delta)x_2^{1-\alpha} \quad \text{for all } x \in S_\delta \quad (3.5)$$

if  $\delta$  is small enough.

Similarly, estimating  $\nabla u(x)$  yields

$$\sup_{x \in S_\delta} \max\{|\nabla u_1(x)x_2^\alpha - (\alpha - 1, 0)|, |\nabla u_2(x)x_2^\alpha - (-2\alpha, 1 - \alpha)|\} \leq C\delta \quad (3.6)$$

if  $\delta$  is small enough (with a universal constant  $C \geq 1$ ), and therefore

$$\sup_{x \in S_\delta} \sup_{v \in \mathbb{R}^2} |v|^{-2} |v \cdot \nabla u(x)v x_2^\alpha - [(1 - \alpha)(v_2^2 - v_1^2) - 2\alpha v_1 v_2]| \leq C\delta$$

(with a new universal  $C \geq 1$ ). Since the function  $(1 - \alpha)(v_2^2 - v_1^2) - 2\alpha v_1 v_2$  is 0-homogeneous and strictly below 1, it follows that

$$|v \cdot \nabla u(x)v| \leq (1 - 5\delta)x_2^{-\alpha}|v|^2 \quad \text{for all } v \in \mathbb{R}^2 \text{ and } x \in S_\delta \quad (3.7)$$

if  $\delta$  is small enough.

**Subcase 1a.** Let us now assume that  $x, y \in S_\delta$  and  $x_2 \leq y_2 \leq 2x_2$ , and consider any time  $t$  such that  $\{\Phi(x, s)\}_{s \in [0, t]}, \{\Phi(y, s)\}_{s \in [0, t]} \subseteq A_\delta$  (and therefore also in  $S_\delta$ ). It follows from (3.5) that  $\Phi_2(x, s) \geq (\alpha(1 - 4\delta)s + x_2^\alpha)^{1/\alpha}$  for  $s \in [0, t]$ , and similarly for  $y$  in place of  $x$ , so  $x_2 \leq y_2$  shows that

$$\min\{\Phi_2(x, s), \Phi_2(y, s)\} \geq (\alpha(1 - 4\delta)s + x_2^\alpha)^{1/\alpha} \quad \text{for all } s \in [0, t].$$

This and (3.7) show that

$$\left| \frac{d}{ds} |\Phi(x, s) - \Phi(y, s)| \right| \leq (1 - 5\delta)(\alpha(1 - 4\delta)s + x_2^\alpha)^{-1} |\Phi(x, s) - \Phi(y, s)|, \quad (3.8)$$

and integration then yields

$$|\Phi(x, t) - \Phi(y, t)| \leq \left( \frac{\alpha(1 - 4\delta)t + x_2^\alpha}{x_2^\alpha} \right)^{(1-5\delta)/\alpha(1-4\delta)} |x - y|. \quad (3.9)$$

Since  $x_2 \geq \frac{1}{2}|x - y|$  due to  $x_2 \leq y_2 \leq 2x_2$ , and  $x_2^\alpha + t \leq 1$  if  $\delta$  is small, it follows that

$$|\Phi(x, t) - \Phi(y, t)| \leq 2|x - y|^{1 - \frac{1-5\delta}{1-4\delta}} \leq 2|x - y|^\delta. \quad (3.10)$$

This is the desired uniform Hölder bound for the  $x, y$  we considered here.

**Subcase 1b.** If now  $x, y \in S_\delta$  are arbitrary (without loss we can assume  $x_2 \leq y_2$ ), let  $z^0 = x, z_1, \dots, z_n = y$  be points on the segment  $xy$  such that  $x_2 \leq z_2^1 \leq 2x_2$  and  $z_2^j = 2z_2^{j-1}$  for

all  $j > 1$ . Then telescoping (3.10) applied to couples of points  $\{(z^{j-1}, z^j)\}_{j=1}^n$  in place of  $(x, y)$  yields

$$|\Phi(x, t) - \Phi(y, t)| \leq \sum_{j=1}^{\infty} 2(2^{1-j}|x - y|)^{\delta} \leq 4\delta^{-1}|x - y|^{\delta} \quad (3.11)$$

as long as the trajectories remain in  $A_{\delta}$ . This finishes Case 1.

For later reference we note that  $|\Phi(x, \cdot) - \Phi(y, \cdot)|$  also cannot decrease too quickly in  $S_{\delta}$  if  $\delta$  is small enough. Specifically, let  $\delta \leq \frac{(1-\alpha)^2}{4C}$ , with  $C \geq 1$  from (3.6). This and (3.6) applied to all points on the segment  $\Phi(x, s)\Phi(y, s)$  show that if  $v(s) := \Phi(y, s) - \Phi(x, s)$  satisfies  $|v_2(t_0)| = \frac{4\alpha}{1-\alpha}|v_1(t_0)|$  at some  $t_0 \in [0, t]$ , then  $\frac{d}{ds}|v_1(t_0)| < 0$  and  $\frac{d}{ds}|v_2(t_0)| > 0$ . That is, we must have  $|v_2(s)| \geq \frac{4\alpha}{1-\alpha}|v_1(s)|$  for all  $s \in [t_0, t]$ , and (3.6) now shows that  $\frac{d}{ds}|v_2| > 0$  on  $[t_0, t]$ . This means that we only need to track  $|v(\cdot)|$  only on the longest interval  $(0, t_0] \subseteq [0, t]$  on which  $|v_2(\cdot)| \leq \frac{4\alpha}{1-\alpha}|v_1(\cdot)|$ . But then similarly to (3.8) we obtain

$$\left| \frac{d}{ds}|v_1(s)| \right| \leq \left( 1 - \alpha + C\delta + \frac{4\alpha}{1-\alpha}C\delta \right) (\alpha(1-4\delta)s + x_2^{\alpha})^{-1}|v_1(s)| =: C''(\alpha(1-4\delta)s + x_2^{\alpha})^{-1}|v_1(s)|$$

on  $(0, t_0]$ . Integrating this yields

$$|v_1(t)| \geq \left( \frac{x_2^{\alpha}}{\alpha(1-4\delta)t + x_2^{\alpha}} \right)^{C''/\alpha(1-4\delta)} |v_1(0)|$$

on  $(0, t_0]$ . Since again  $\alpha(1-4\delta)t + x_2^{\alpha} \leq 1$ , and also  $x_2 \geq cv_1(0)$  for some  $c = c(\alpha) > 0$  when  $|v_2(0)| \leq \frac{4\alpha}{1-\alpha}|v_1(0)|$  (because  $x, y \in S_{\delta}$  and  $\frac{1}{\delta} > \frac{4\alpha}{1-\alpha}$ ), it follows that

$$|v_1(t)| \geq c''|v_1(0)|^{(C''+1)/\alpha(1-4\delta)}.$$

on  $(0, t_0]$ , with some  $c'' = c''(\alpha)$ . Since we also have  $|v_2(\cdot)| \leq \frac{4\alpha}{1-\alpha}|v_1(\cdot)|$  there, we obtain

$$|\Phi(y, t) - \Phi(x, t)| \geq \gamma|x - y|^{1/\gamma}. \quad (3.12)$$

on  $(0, t_0]$  (and hence as long as the trajectories remain in  $A_{\delta}$ , by the previous argument) whenever  $x, y \in S_{\delta}$ , with some  $\gamma = \gamma(\alpha) > 0$ .

**Case 2.** Next we assume that  $x, y \in S'_{\delta} := A_{\delta} \cap \{x_2 \geq \delta x_1\}$ . We now want to pick small enough  $\varepsilon > 0$  (depending on  $\alpha$  and  $\delta$ , and hence on  $\alpha$ ) such that for any streamline  $\{\tilde{\psi}_{\alpha} = r\}$  that intersects  $A_{\delta}$  (in particular,  $r \leq \delta^{2-\alpha}$ ), the time it takes for a particle moving on this streamline (with velocity  $u$ ) to traverse the portion lying inside  $S'_{\delta} \setminus S_{\delta}$  is shorter than the time it takes for a particle moving on the streamline  $\{\tilde{\psi}_{\alpha} = r'\}$  for any  $r' \in [\frac{r}{2}, 2r]$  to traverse the portion lying inside  $\{\varepsilon\delta x_2 \leq x_1 \leq \delta x_2\}$ . Such  $\varepsilon$  exists because for the stream function  $\psi(x) = x_1 x_2 (x_1 + x_2)^{-\alpha}$  the ratio of these times only depends on  $\frac{r'}{r}$  (due to  $(2-\alpha)$ -homogeneity of  $\psi$ ) and converges to 0 as  $\varepsilon \rightarrow 0$  (uniformly in  $\frac{r'}{r} \in [\frac{1}{2}, 2]$ ), and the ratio of any derivative of  $\tilde{\psi}_{\alpha}$  and the same derivative of  $\psi$  is within  $[\frac{1}{2}, 2]$  on all of  $A_{\delta}$  if  $\delta$  is small enough. The reader may want to keep in mind that since the level sets of  $\psi$  are just scaled copies of each other, the same is true asymptotically for the level sets of  $\tilde{\psi}_{\alpha}$  near the origin.

**Subcase 2a.** The case  $x, y \in S_{\delta}$  was already covered. Let us now assume  $x, y \in S'_{\delta} \setminus S_{\delta}$ , as well as  $\tilde{\psi}_{\alpha}(x) \leq \tilde{\psi}_{\alpha}(y) \leq 2\tilde{\psi}_{\alpha}(x)$ . Let  $t_0 \geq 0$  be the first time when both trajectories  $\Phi(x, \cdot)$  and  $\Phi(y, \cdot)$

have left  $S'_\delta \setminus S_\delta$ . Because of our choice of  $\varepsilon$ , we must have  $\Phi(x, t_0), \Phi(y, t_0) \in \{\varepsilon\delta x_2 \leq x_1 \leq \delta x_2\}$ , so (3.11) shows that while both trajectories stay in  $A_\delta$ , we have

$$|\Phi(x, t) - \Phi(y, t)| \leq 4\delta^{-1}|\Phi(x, t_0) - \Phi(y, t_0)|^\delta \quad (3.13)$$

for  $t \geq t_0$ . To estimate the last term, we notice that while either trajectory is within  $S'_\delta \setminus S_\delta$ , its velocity is bounded below by  $c|x|^{1-\alpha}$  for some  $c = c(\alpha)$ , which means that  $t_0 \leq C_1|x|^\alpha$  for some  $C_1 = C_1(\alpha)$ . Since we also have the bound  $|\nabla u| \leq C_2|x|^{-\alpha}$  with  $C_2 = C_2(\alpha)$  on the portions of both trajectories lying in  $S_{\delta'} \cap \{x_1 \geq \varepsilon\delta x_2\}$  (recall that  $\delta, \varepsilon$  depend only on  $\alpha$ ), a simple integration shows that

$$|\Phi(x, t) - \Phi(y, t)| \leq e^{C_2|x|^{-\alpha}C_1|x|^\alpha}|x - y| = e^{C_2C_1}|x - y| \quad (3.14)$$

for  $t \in [0, t_0]$ . It follows that

$$|\Phi(x, t) - \Phi(y, t)| \leq C|x - y|^\delta \quad (3.15)$$

while both trajectories stay in  $A_\delta$ , with  $C = C(\alpha)$ .

**Subcase 2b.** Next assume  $x, y \in S'_\delta \setminus S_\delta$  (then without loss  $\tilde{\psi}_\alpha(x) \leq \tilde{\psi}_\alpha(y)$ ) and also that  $\tilde{\psi}_\alpha(x) < 2\tilde{\psi}_\alpha(y)$ . This last fact shows that the distance of any two points on the portions of the level sets of  $\tilde{\psi}_\alpha$  containing  $x$  and  $y$  that lie in  $S'_\delta \setminus S_\delta$  is at most  $C_3|x - y|$  for some  $C_3 = C_3(\alpha)$ . If now  $t_0$  is the first time when one of the trajectories exits  $S'_\delta \setminus S_\delta$ , we will have

$$|\Phi(x, t_0) - \Phi(y, t_0)| \leq C_3|x - y|,$$

which means we can finish this subcase by using the subcase in the next paragraph.

**Subcase 2c.** If now  $x \in S'_\delta \setminus S_\delta$  and  $y \in S_\delta$ , we let  $z$  be the point with  $\tilde{\psi}_\alpha(z) = \tilde{\psi}_\alpha(x)$  and  $z_1 = \delta z_2$ . Let also  $t_0$  be the first time when  $\Phi(x, \cdot)$  leaves  $S'_\delta \setminus S_\delta$ . Then the fact that the level sets of  $\tilde{\psi}_\alpha$  intersect the line  $\{x_1 = \delta x_2\}$  transversally (with angles uniformly bounded away from 0) shows that  $\max\{|x - z|, |y - z|\} \leq C_4|x - y|$  for some  $C_4 = C_4(\alpha)$ . If now  $t_0$  is the first time when  $\Phi(x, \cdot)$  leaves  $S'_\delta \setminus S_\delta$ , then the argument leading to (3.14) shows that

$$|\Phi(t_0, x) - \Phi(t_0, z)| \leq e^{C_2C_1}C_4|x - y|,$$

while (3.11) shows

$$|\Phi(t_0, y) - \Phi(t_0, z)| \leq 4\delta^{-1}|y - z|^\delta.$$

Adding these to estimate  $|\Phi(t_0, x) - \Phi(t_0, y)|$  and then applying (3.11) on the interval  $[t_0, t]$  yields

$$|\Phi(t, x) - \Phi(t, y)| \leq C|x - y|^{\delta^2} \quad (3.16)$$

while both trajectories stay in  $A_\delta$ , with a new  $C = C(\alpha)$ . Since this bound can absorb the bounds in (3.11) and (3.15) by adjusting  $C$ , this finishes Case 2.

**Case 3.** It remains to consider the case  $x \in A_\delta \setminus S'_\delta = A_\delta \cap \{x_2 < \delta x_1\}$ .

**Subcase 3a.** If  $y \in S'_\delta$ , let  $z$  be the point with  $\tilde{\psi}_\alpha(z) = \tilde{\psi}_\alpha(x)$  and  $z_2 = \delta z_1$ , and let  $t_0$  be the first time when  $\Phi(x, \cdot)$  enters  $S_\delta$  (and hence  $\Phi(x, t_0) = (z_2, z_1)$  due to symmetry of  $\tilde{\psi}_\alpha$  in its arguments). Just as in the last paragraph, it follows that  $\max\{|x - z|, |y - z|\} \leq C_4|x - y|$ , and therefore (3.16) yields

$$|\Phi(t_0, y) - \Phi(t_0, z)| \leq C|x - y|^{\delta^2},$$

with a new  $C = C(\alpha)$ . On the other hand, the same symmetry shows that  $\Phi(z, t_0) = (x_2, x_1)$ , so

$$|\Phi(t_0, x) - \Phi(t_0, z)| = |x - z| \leq C_4|x - y|.$$

Again, adding these to estimate  $|\Phi(t_0, x) - \Phi(t_0, y)|$  and then applying (3.11) on the interval  $[t_0, t]$  yields

$$|\Phi(t, x) - \Phi(t, y)| \leq C|x - y|^{\delta^3} \quad (3.17)$$

while both trajectories stay in  $A_\delta$ , with a new  $C = C(\alpha)$ .

**Subcase 3b.** Finally, if both  $x, y \in A_\delta \setminus S'_\delta$ , let  $t_0$  be the first time when one of  $\Phi(x, \cdot)$  and  $\Phi(y, \cdot)$  enters  $S'_\delta$  (assume it is  $\Phi(y, t_0)$ ). We can now apply the argument leading to (3.12) on the time interval  $[t_0, 0]$ , with time running in reverse and with the roles of the two coordinates reversed. This yields

$$|x - y| \geq \gamma|\Phi(y, t_0) - \Phi(x, t_0)|^{1/\gamma},$$

which together with (3.17) applied on the time interval  $[t_0, t]$  implies

$$|\Phi(t, x) - \Phi(t, y)| \leq C|x - y|^{\delta^3\gamma}$$

while both trajectories stay in  $A_\delta$ , with a new  $C = C(\alpha)$ . This finishes Case 3, and so also the proof for  $\Phi$ .

**Back to  $\Phi'$ .** We now note that  $\Phi'(x, t) = \Phi(x, P_\alpha(x)t)$ , where  $P_\alpha(x) := T_{\psi_\alpha}(\psi_\alpha(x))$ , and so

$$|\Phi'(x, t) - \Phi'(y, t)| \leq |\Phi(x, P_\alpha(x)t) - \Phi(y, P_\alpha(x)t)| + |\Phi(y, P_\alpha(x)t) - \Phi(y, P_\alpha(y)t)|$$

Boundedness of  $P_\alpha$  and the result for  $u'$  now show that the first term on the right-hand side is bounded by multiple of some power of  $|x - y|$ , uniformly in  $t \in [0, 1]$ . The same is true for the second term because  $u'$  is bounded and  $P_\alpha$  is Hölder continuous due to boundedness of  $\nabla\psi_\alpha$ , Lemma 3.1(iii), and the fact that  $\psi_\alpha$  has a quadratic critical point at the origin.  $\square$

## 4 Appendix: Fractional derivatives and concavity of $\log \psi_\alpha$

As above, we will again use the notation  $x = (x_1, x_2) \in \mathbb{R}^2$  in the next two results.

**Lemma 4.1.** *Let  $P_2 := (-1, 1)^2$  and  $\Gamma := [\{0\} \times (-1, 1)] \cup [(-1, 1) \times \{0\}]$ . If  $f \in C_{\text{loc}}^1(P_2 \setminus \Gamma)$  and there are  $K \geq 0$  and  $\beta \in [0, 1)$  such that*

$$|f(x_1, x_2)| \leq K|x_1x_2|^{-\beta},$$

$$|\nabla f(x_1, x_2)| \leq K|x_1x_2|^{-\beta-1}$$

for all  $(x_1, x_2) \in P_2$ , then for any  $\gamma \in (0, 1)$  and  $\epsilon > 0$  there is  $C_{\gamma, \beta, \epsilon} > 0$  such that

$$|\Lambda^\gamma f(x_1, x_2)| \leq C_{\gamma, \beta, \epsilon}K|x_1x_2|^{-\beta-\gamma-\epsilon}$$

for all  $(x_1, x_2) \in P_2$ .

This lemma can be trivially extended to the case where  $f$  is smooth away from finitely many lines, with a controlled blow-up near those lines. In particular, we have the following corollary.



**Corollary 4.2.** *If  $f \in C_{\text{loc}}^1(P_2 \setminus \Gamma)$  and there are  $K \geq 0$  and  $\beta \in [0, 1]$  such that*

$$|f(x_1, x_2)| \leq K |\sin(\pi x_1) \sin(\pi x_2)|^{-\beta},$$

$$|\nabla f(x_1, x_2)| \leq K |\sin(\pi x_1) \sin(\pi x_2)|^{-\beta-1}$$

for all  $(x_1, x_2) \in P_2$ , then for any  $\gamma \in (0, 1)$  and  $\epsilon > 0$  there is  $C_{\gamma, \beta, \epsilon} > 0$  such that

$$|\Lambda^\gamma f(x_1, x_2)| \leq C_{\gamma, \beta, \epsilon} K |\sin(\pi x_1) \sin(\pi x_2)|^{-\beta-\gamma-\epsilon}$$

for all  $(x_1, x_2) \in P_2$ .

*Proof of Lemma 4.1.* Without loss of generality, assume that  $0 < x_1 \leq x_2$ . Then (with all integrals below having  $y \in P_2$ ),

$$\Lambda^\gamma f(x) = \int_{|x-y| < \frac{1}{2}x_1} \frac{f(x) - f(y)}{|x-y|^{2+\gamma}} dy + \int_{|x-y| \geq \frac{1}{2}x_1} \frac{f(x) - f(y)}{|x-y|^{2+\gamma}} dy.$$

To estimate the first integral, note that all  $y = (y_1, y_2)$  satisfying  $|x-y| < \frac{1}{2}x_1$  also satisfy  $y_1 \geq \frac{1}{2}x_1$  and  $y_2 \geq \frac{1}{2}x_2$ . Thus, if  $|x-y| < \frac{1}{2}x_1$ , we have

$$|f(x) - f(y)| = |f(x) - f(y)|^{1-\gamma-\epsilon} |f(x) - f(y)|^{\gamma+\epsilon} \leq 2K(x_1 x_2)^{-\beta-\gamma-\epsilon} |x-y|^{\gamma+\epsilon},$$

where in the last step we used the two assumed inequalities to estimate the two powers of  $|f(x) - f(y)|$ . We then get

$$\int_{|x-y| < \frac{1}{2}x_1} \frac{|f(x) - f(y)|}{|x-y|^{2+\gamma}} dy \leq 2K(x_1 x_2)^{-\beta-\gamma-\epsilon} \int_0^{x_1} r^{-1+\epsilon} dr \leq \frac{2K}{\epsilon} (x_1 x_2)^{-\beta-\gamma-\frac{\epsilon}{2}}.$$

It therefore suffices to consider the second integral. We split it into

$$\int_{\frac{1}{2}x_1 \leq |x-y| < \frac{1}{2}x_2} \frac{|f(x) - f(y)|}{|x-y|^{2+\gamma}} dy + \int_{|x-y| \geq \frac{1}{2}x_2} \frac{|f(x) - f(y)|}{|x-y|^{2+\gamma}} dy = I + II.$$

First we estimate  $II$  via

$$|II| \leq C_\gamma K (x_1 x_2)^{-\beta} \int_{\frac{1}{2}x_2}^3 r^{-1-\gamma} dr + K \int_{|x-y| \geq \frac{1}{2}x_2} \frac{|y_1 y_2|^{-\beta}}{|x-y|^{2+\gamma}} dy.$$

The first term is no more than  $C_\gamma K (x_1 x_2)^{-\beta} x_2^{-\gamma} \leq C_\gamma K (x_1 x_2)^{-\beta-\gamma}$  (with a new constant) and for the second we use Hölder's inequality to obtain

$$\begin{aligned} \int_{|x-y| \geq \frac{1}{2}x_2} \frac{|y_1 y_2|^{-\beta}}{|x-y|^{2+\gamma}} dy &\leq \left( \int_{P_2} |y_1 y_2|^{-(1-\delta)} dy \right)^{\frac{\beta}{1-\delta}} \left( \int_{|x-y| \geq \frac{1}{2}x_2} |x-y|^{-\frac{2+\gamma}{1-\frac{\beta}{1-\delta}}} dy \right)^{1-\frac{\beta}{1-\delta}} \\ &\leq C_\delta x_2^{-(2+\gamma)+2(1-\frac{\beta}{1-\delta})} = C_\delta x_2^{-\gamma-\frac{2\beta}{1-\delta}}, \end{aligned}$$

for any  $\delta \in (0, 1 - \beta)$ . We choose such  $\delta$  so that  $\frac{2\beta}{1-\delta} < 2\beta + \gamma$ , and then  $x_1 \leq x_2$  yields

$$|II| \leq C_{\gamma, \beta} K (x_1 x_2)^{-\beta-\gamma}.$$

We similarly estimate

$$|I| \leq C_\gamma K(x_1 x_2)^{-\beta} \int_{\frac{1}{2}x_1}^{\frac{1}{2}x_2} r^{-1-\gamma} dr + C_\beta K x_2^{-\beta} \int_{\frac{1}{2}x_1 \leq |x-y| < \frac{1}{2}x_2} \frac{|y_1|^{-\beta}}{|x-y|^{2+\gamma}} dy.$$

The first term is no more than  $C_\gamma K(x_1 x_2)^{-\beta} x_1^{-\gamma} \leq C_\gamma K(x_1 x_2)^{-\beta-\gamma}$  (with a new constant) and the second is no more than

$$C_\beta K x_2^{-\beta} \int_{|y_1| < \frac{1}{2}x_1} \frac{|y_1|^{-\beta}}{|x-y|^{2+\gamma}} dy + C_{\beta,\gamma} K(x_1 x_2)^{-\beta} \int_{\frac{1}{2}x_1}^{\frac{1}{2}x_2} r^{-1-\gamma} dr.$$

The second of these terms is again estimated by  $C_{\beta,\gamma} K(x_1 x_2)^{-\beta} x_1^{-\gamma} \leq C_{\beta,\gamma} K(x_1 x_2)^{-\beta-\gamma}$ , and the first by

$$C_{\beta,\gamma} K x_2^{-\beta} \int_{|y_1| < \frac{1}{2}x_1} \frac{|y_1|^{-\beta}}{(x_1^2 + (x_2 - y_2)^2)^{1+\frac{1}{2}\gamma}} dy \leq C_{\beta,\gamma} K x_2^{-\beta} x_1^{1-\beta} \int_0^\infty \frac{dz}{(x_1^2 + z^2)^{1+\frac{1}{2}\gamma}}.$$

This is no more than  $C_{\beta,\gamma} K x_2^{-\beta} x_1^{1-\beta} x_1^{-1-\gamma} \leq C_{\beta,\gamma} K(x_1 x_2)^{-\beta-\gamma}$ , concluding the proof.  $\square$

**Lemma 4.3.** *If  $\psi_\alpha$  is the function from Section 3, then  $\log \psi_\alpha$  is concave for any  $\alpha \in [0, 1]$ .*

*Proof.* To simplify notation, let us instead consider the function  $f_\alpha := \log(2^{-\alpha} \psi_\alpha(\frac{\cdot}{\pi}))$ , so that

$$f_\alpha(x, y) = \log(\sin(x)) + \log(\sin(y)) - \alpha \log(\sin(x) + \sin(y)).$$

We have

$$\begin{aligned} \partial_{xx} \log(\sin(x)) &= -\frac{1}{\sin^2(x)}, \\ \partial_{xx} \log(\sin(x) + \sin(y)) &= -\frac{1 + \sin(x) \sin(y)}{(\sin(x) + \sin(y))^2} \\ \partial_{yy} \log(\sin(x) + \sin(y)) &= -\frac{1 + \sin(x) \sin(y)}{(\sin(x) + \sin(y))^2} \\ \partial_{xy} \log(\sin(x) + \sin(y)) &= -\frac{\cos(x) \cos(y)}{(\sin(x) + \sin(y))^2} \end{aligned}$$

Now with  $a := \sin(x)$  and  $b := \sin(y)$  we obtain

$$\partial_{xx} f_\alpha = -\frac{1}{a^2} + \alpha \frac{1+ab}{(a+b)^2} \leq -\frac{1}{a^2} + \frac{1+ab}{(a+b)^2} \leq -\frac{b^2}{(a+b)^2 a^2} \leq 0$$

because  $a, b, \alpha \in [0, 1]$ . Symmetry yields  $\partial_{yy} f_\alpha \leq -\frac{a^2}{(a+b)^2 b^2} \leq 0$ , so

$$\text{tr } D^2 f_\alpha = \Delta f_\alpha \leq 0.$$

We also have

$$\det D^2 f_\alpha = \partial_{xx} f_\alpha \partial_{yy} f_\alpha - (\partial_{xy} f_\alpha)^2 \geq \frac{b^2}{(a+b)^2 a^2} \frac{a^2}{(a+b)^2 b^2} - \frac{(1-a^2)(1-b^2)}{(a+b)^4} \geq 0,$$

and concavity of  $f_\alpha$  follows.  $\square$

## References

- [1] G. Alberti, G. Crippa, and A.L. Mazzucato. Exponential self-similar mixing and loss of regularity for continuity equations. *C. R. Acad. Sci. Paris, Ser. I*, 352:901–906, 2014.
- [2] G. Alberti, G. Crippa, and A.L. Mazzucato. Exponential self-similar mixing by incompressible flows. *J. Amer. Math. Soc.*, 32:445–490, 2019.
- [3] D. V. Anosov. Geodesic flows on closed Riemannian manifolds with negative curvature. *Proc. Steklov Inst.*, 90, 1967.
- [4] J. Bedrossian, A. Blumenthal, and S. Punshon-Smith. Almost-sure exponential mixing of passive scalars by the stochastic Navier-Stokes equations. *Preprint*, arXiv:1905.03869.
- [5] Y. Brenier, F. Otto, and C. Seis. Upper bounds on the coarsening rates in demixing binary viscous fluids. *SIAM J. Math. Anal.*, 43:114–134, 2011.
- [6] A. Bressan. A lemma and a conjecture on the cost of rearrangements. *Rend. Sem. Mat. Univ. Padova*, 110:97–102, 2003.
- [7] A. Bressan. Prize offered for the solution of a problem on mixing flows. <http://www.math.psu.edu/bressan/PSPDF/prize1.pdf>, 2006.
- [8] P. Constantin, A. Kiselev, L. Ryzhik, and A. Zlatoš. Diffusion and mixing in fluid flow. *Ann. of Math. (2)*, 168(2):643–674, 2008.
- [9] I.P. Cornfeld, S.V. Fomin, and Y.G. Sinai. *Ergodic Theory*. Springer-Verlag, New York, 1982.
- [10] M. Coti-Zelati, M.G. Delgadino, and T.M. Elgindi. On the relation between enhanced dissipation time-scales and mixing rates. *Comm. Pure Appl. Math.*, to appear, arXiv:1806.03258.
- [11] G. Crippa and C. De Lellis. Estimates and regularity results for the DiPerna-Lions flow. *J. Reine Angew. Math.*, 616:15–46, 2008.
- [12] C. R. Doering and J.-L. Thiffeault. Multiscale mixing efficiencies for steady sources. *Phys. Rev. E*, 74 (2), 025301(R), August 2006.
- [13] D. Dolgopyat. On decay of correlations in Anosov flows. *Ann. of Math.* **147** (1998), 357-390.
- [14] D. Dolgopyat. Personal communication.
- [15] D. Dolgopyat, V. Kaloshin, and L. Korolov. Sample path properties of the stochastic flows. *Ann. of Prob.* **32** (2004), 1-27.
- [16] Y. Feng and G. Iyer. Dissipation Enhancement by Mixing. *Preprint*, arXiv:1806.03699.
- [17] G. Iyer, A. Kiselev, and X. Xu. Lower bounds on the mix norm of passive scalars advected by incompressible enstrophy-constrained flows. *Nonlinearity*, 27(5):973–985, 2014.
- [18] A. Katok. Bernoulli diffeomorphisms on surfaces. *Ann. of Math.* **110** (1979), 529-547.

- [19] Z. Lin, J. L. Thiffeault, and C. R. Doering. Optimal stirring strategies for passive scalar mixing. *J. Fluid Mech.*, 675:465–476, 2011.
- [20] C. Liverani. On contact Anosov flows. *Ann. of Math.* **159** (2004), 1275–1312.
- [21] 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(11):115611, 15, 2012.
- [22] G. Mathew, I. Mezić, and L. Petzold. A multiscale measure for mixing. *Physica D*, 211(1-2):23–46, 2005.
- [23] V. Maz’ya, *Sobolev Spaces. With Applications to Elliptic Partial Differential Equations*, 2nd, revised and augmented ed., Springer, Berlin, 2011.
- [24] E. Di Nezza, G. Palatucci, and E. Valdinoci. Hitchhiker’s guide to the fractional Sobolev spaces, *Bull. Sci. math.*, 136(5):521–573, 2012.
- [25] F. Otto, C. Seis, and D. Slepčev. Crossover of the coarsening rates in demixing of binary viscous liquids. *Commun. Math. Sci.* 11:441–464, 2013.
- [26] R. Pierrehumbert, *Tracer microstructure in the large-eddy dominated regime*, *Chaos, Solitons & Fractals* **4** (1994), 1091–1110.
- [27] C. Seis. Maximal mixing by incompressible fluid flows. *Nonlinearity*, 26(12):3279–3289, 2013.
- [28] D. Slepčev. Coarsening in nonlocal interfacial systems. *SIAM J. Math. Anal.*, 40(3):1029–1048, 2008.
- [29] T. A. Shaw, J.-L. Thiffeault, and C. R. Doering. Stirring up trouble: multi-scale mixing measures for steady scalar sources. *Phys. D*, 231(2):143–164, 2007.
- [30] J. Springham. Ergodic properties of linked-twist maps. *PhD Thesis*, University of Bristol, 2008.
- [31] J.-L. Thiffeault. Using multiscale norms to quantify mixing and transport. *Nonlinearity*, 25(2):R1–R44, 2012.
- [32] J.-L. Thiffeault, C. R. Doering, and J. D. Gibbon. A bound on mixing efficiency for the advection-diffusion equation. *J. Fluid Mech.*, 521:105–114, 2004.
- [33] Y. Yao and A. Zlatoš. Mixing and un-mixing by incompressible flows. *J. Eur. Math. Soc.*, 19(7): 1911–1948, 2017.