

# LOWER BOUNDS ON THE MIX NORM OF PASSIVE SCALARS ADVECTED BY INCOMPRESSIBLE ENSTROPY-CONSTRAINED FLOWS.

GAUTAM IYER, ALEXANDER KISELEV, AND XIAOQIAN XU

ABSTRACT. Consider a diffusion-free passive scalar  $\theta$  being mixed by an incompressible flow  $u$  on the torus  $\mathbb{T}^d$ . Our aim is to study how well this scalar can be mixed under an enstrophy constraint on the advecting velocity field. Our main result shows that the mix-norm  $(\|\theta(t)\|_{H^{-1}})$  is bounded below by an exponential function of time. The exponential decay rate we obtain is not universal and depends on the size of the support of the initial data. We also perform numerical simulations and confirm that the numerically observed decay rate scales similarly to the rigorous lower bound, at least for a significant initial period of time. The main idea behind our proof is to use recent work of Crippa and DeLellis ('08) making progress towards the resolution of Bressan's rearrangement cost conjecture.

## 1. INTRODUCTION

The mixing of tracer particles by fluid flows is ubiquitous in nature, and have applications ranging from weather forecasting to food processing. An important question that has attracted attention recently is to study “how well” tracers can be mixed under a constraint on the advecting velocity field, and what is the optimal choice of the “best mixing” velocity field (see [24] for a recent review).

Our aim in this paper is to study how well passive tracers can be mixed under an *enstrophy constraint* on the advecting fluid. By passive, we mean that the tracers provide no feedback to the advecting velocity field. Further, we assume that diffusion of the tracer particles is weak and can be neglected on the relevant time scales. Mathematically, the density of such tracers (known as passive scalars) is modeled by the transport equation

$$(1.1) \quad \partial_t \theta(x, t) + u \cdot \nabla \theta = 0, \quad \theta(x, 0) = \theta_0(x).$$

To model stirring, the advecting velocity field  $u$  is assumed to be incompressible. For simplicity we study (1.1) with periodic boundary conditions (with period 1), mean zero initial data, and assume that all functions in question are smooth.

The first step is to quantify “how well” a passive scalar is mixed in our context. For *diffusive* passive scalars, the decay of the variance is a commonly used measure of mixing (see for instance [10, 14, 22, 25] and references there in). But for diffusion free scalars the variance is a conserved and does not change with time. Thus,

---

This material is based upon work partially supported by the National Science Foundation under grants DMS-1007914, DMS-1104415, DMS-1159133, DMS-1252912. GI acknowledges partial support from an Alfred P. Sloan research fellowship. AK acknowledges partial support from a Guggenheim fellowship. The authors also thank the Center for Nonlinear Analysis (NSF Grants No. DMS-0405343 and DMS-0635983), where part of this research was carried out.

following [18] we quantify mixing using the  $H^{-1}$ -Sobolev norm: the *smaller*  $\|\theta\|_{H^{-1}}$ , the better mixed the scalar  $\theta$  is.

The reason for using a negative Sobolev norm in this context has its roots in [14, 18, 20, 22]. The motivation is that if the flow generated by the velocity field is mixing in the ergodic theory sense, then any advected quantity (in particular  $\theta$ ) converges to 0 weakly in  $L^2$  as  $t \rightarrow \infty$ . This can be shown to imply that  $\|\theta(\cdot, t)\|_{H^s} \rightarrow 0$  for all  $s < 0$ , and conversely, if  $\|\theta(\cdot, t)\|_{H^s} \rightarrow 0$  for some  $s < 0$  then  $\theta(x, t)$  converges weakly to zero. Thus any negative Sobolev norm of  $\theta$  can in principle be used to quantify its mixing properties. In two dimensions the choice of using the  $H^{-1}$  norm in particular was suggested by Lin et. al. [18] as it scales like the area dominant unmixed regions; a natural length scale associated with the system. We will work with the same Sobolev norm in any dimension  $d$ ; the ratio of  $H^{-1}$  norm to  $L^2$  norm has a dimension of length, and since  $L^2$  norm of  $\theta(x, t)$  is conserved, the  $H^{-1}$  norm provides a natural length scale associated with the mixing process.

The questions we study in this paper are motivated by recent work of Lin et. al. [18]. In [18], the authors address two questions on the two dimensional torus:

- The time decay of  $\|\theta\|_{H^{-1}}$ , given the *fixed energy* constraint  $\|u(t)\|_{L^2} = U$ .
- The time decay of  $\|\theta(t)\|_{H^{-1}}$  given a *fixed enstrophy* constraint of the form  $\|\nabla u(t)\|_{L^2} = F$ .

In the first case the authors prove a lower bound for  $\|\theta(\cdot, t)\|_{H^{-1}(\mathbb{T}^2)}$  that is linear with negative slope. This suggests that it may be possible to “mix perfectly in finite time”; namely choose  $u$  in a manner that drives  $\|\theta(\cdot, t)\|_{H^{-1}}$  to zero in finite time. This was followed by an explicit example in [19] exhibiting finite time perfect mixing, under a finite energy constraint. This example uses an elegant “slice and dice” construction, which requires the advecting velocity field to develop finer and finer scales. Thus, while their example maintains a fixed energy constraint, the enstrophy ( $\|\nabla u\|_{L^2}$ ) explodes. Together with the numerical analysis in [18, 19] this suggests that finite time perfect mixing by an enstrophy constrained incompressible flow might be impossible. Our main theorem settles this affirmatively.

**Theorem 1.1.** *Let  $u$  be a smooth (time dependent) incompressible periodic vector field on the  $d$ -dimensional torus, and let  $\theta$  solve (1.1) with periodic boundary conditions and  $L^\infty$  initial data  $\theta_0$ . For any  $p > 1$  and  $\lambda \in (0, 1)$  there exists a length scale  $r_0 = r_0(\theta_0, \lambda)$ , an explicit constant  $\varepsilon_0 = \varepsilon_0(\lambda, d)$ , and a constant  $c = c(d, p)$  such that*

$$(1.2) \quad \|\theta(t)\|_{H^{-1}} \geq \varepsilon_0 r_0^{d/2+1} \|\theta_0\|_{L^\infty} \exp\left(\frac{-c}{m(A_\lambda)^{1/p}} \int_0^t \|\nabla u(s)\|_{L^p} ds\right).$$

Here  $A_\lambda$  is the super level set  $\{\theta_0 > \lambda \|\theta_0\|_{L^\infty}\}$ .

In particular, if the instantaneous enstrophy constraint  $\|\nabla u\|_{L^2} \leq F$  is enforced, then  $\|\theta(t)\|_{H^{-1}}$  decays at most exponentially with time.

Before commenting on the  $r_0$  and  $m(A_\lambda)$  dependence, we briefly mention some applications. There are many physical situations where  $\int_0^t \|\nabla u(s)\|_{L^2} ds$  is well controlled. Some examples are when  $u$  satisfies the incompressible Navier-Stokes equations with  $L^2$  forcing [9, 13], the 2D incompressible Euler equations [3] or a variety of active scalar equations including the critical surface quasi-geostrophic equation [7, 8, 11, 17]. In each of these situations the passive scalars can not be mixed perfectly in finite time. More precisely, a lower bound for the  $H^{-1}$ -norm

of the scalar density can be read off using (1.2) and the appropriate control on  $\|\nabla u\|_{L^2}$ .

We also mention that the proof of this theorem is not based on energy methods. Instead, the main idea is to relate the notion of “mixed to scale  $\delta$ ” to the  $H^{-1}$  norm, and use recent progress by Crippa and DeLellis [12] towards Bressan’s rearrangement cost conjecture [5]. Some of these ideas were already suggested in [19].

We defer the proof of Theorem 1.1 to Section 2, and pause to analyze the dependence of the bound in (1.2) on  $r_0$  and  $m(A_\lambda)$ .

The length scale  $r_0$  is morally scale at which the super level set  $A_\lambda$  is “unmixed”; a notion that is made precise later. Our proof, however, imposes a slightly stronger condition: namely, our proof will show that  $r_0$  can be any length scale such that “most” of the super level set  $A_\lambda$  occupies “most” of the union of disjoint balls of radius *at least*  $r_0$ . While we are presently unable to estimate  $r_0$  in terms of a tangible norm of  $\theta_0$ , we remark that we at least expect a connection between  $r_0$  and the ratio of the measure of  $A_\lambda$  to the perimeter of  $A_\lambda$  (see [23] for a related notion).

On the other hand, we point out that the pre-factor in (1.2) can be improved at the expense of the decay rate. To see this, suppose for some  $\kappa \in [0, 1/2]$  there exists  $N$  disjoint balls of radius at least  $r_1$  such that the fraction of each of these balls occupied by  $A_\lambda$  is *at least*  $1 - \kappa$ . Then our proof will show that (1.2) in Theorem 1.1 can be replaced by

$$(1.2') \quad \|\theta(t)\|_{H^{-1}} \geq \varepsilon_0 r_1^{d/2+1} \|\theta_0\|_{L^\infty} \exp\left(\frac{-c}{(Nr_1^d)^{1/p}} \int_0^t \|\nabla u(s)\|_{L^p} ds\right).$$

In this case, if  $\theta_0 \in C^1$ , the mean value theorem will guarantee that we can choose  $N = 1$  and  $r_1 > \frac{\|\theta_0\|_{L^\infty}}{C\|\nabla\theta_0\|_{L^\infty}}$  for a purely dimensional constant  $C$ .

Next we turn to the exponential decay rate. The dependence of this on  $m(A_\lambda)$  is natural. To see this, suppose momentarily that  $\theta_0$  only takes on the values  $\pm 1$  or 0 representing two insoluble, immiscible fluids which are injected into a large fluid container. Physical intuition suggests that the less the amount of fluid that is injected, the faster one can mix it. Indeed, this is reflected in (1.2) as in this case  $m(A_\lambda) = \frac{1}{2}m(\text{supp}(\theta_0))$ ; so the smaller the support of the initial data, the worse the lower bound (1.2) is. We mention that a bound similar to (1.2) was proved in [21] using optimal transport and ideas from [4]. In [21], however, the author only considers bounded variation “binary phase” initial data, where the two phases occupy the entire region; consequently the result does not capture the dependence of the decay rate on the initial data.

We do not know if the estimate of the exponential decay provided by bound (1.2) is optimal, however, numerical simulations suggest that it may be not far off. A good candidate for the velocity field that might achieve the optimal lower bound was presented in [18] using a steepest descent method (equation (4.1), below). Due to the nonlinear nature of this formula, it is hard to rigorously prove upper bounds; but all our numerical simulations in Section 4 seem to indicate an exponential lower bound bound with a rate constant that is in a good qualitative agreement with (1.2).

However, our numerical simulations show that even if we start with initial data that is localized to a small region, it gradually spreads during mixing. The incompressibility constraint will of course guarantee that the measure of the support of the solution is conserved in time. But since the enstrophy constraint forbids abrupt changes in the velocity field, the “region occupied” by the initial data tends to spread and is likely to eventually “fill” the entire torus (see Figure 1).

A very interesting question is whether there will eventually be a transition in dynamics where the factor in the exponential decay of the mix norm depends only on the volume of the entire domain, and not the size of the region occupied by the initial data. Our attempts to get insight into this question numerically were inconclusive, as we ran out of resolution before observing such a regime change. What we address here, however, is an interesting link between universality of the exponential lower bound and mixing in domains with boundaries.

It has been observed formally [15,16] that presence of walls with no slip conditions inhibits mixing; hyperbolic flows which usually lead to exponential decay of various mixing measures lead to only algebraic mixing rates in presence of walls. Here, we provide an elementary and rigorous argument showing that universality of the exponential lower bound in mixing with an enstrophy constraint would lead to an algebraic in time lower bound if the initial data is compactly supported away from the boundary and the advecting velocity field vanishes at the boundary. Agreement with earlier heuristic arguments is intriguing; but it is not clear to us if one can expect such result to be true in full generality. It would be very interesting to know whether the efficient mixing by an incompressible enstrophy constrained flow spreads the initial data over the entire ambient volume and results in the slowdown of the exponential decay. We plan to further investigate this issue in the future.

**Notational convention, and plan of this paper.** We will assume throughout the paper that  $d \geq 2$  is the dimension, and  $\mathbb{T}^d$  is the  $d$ -dimensional torus, with side length 1. All periodic functions are assumed to be 1-periodic, and we use  $m$  to denote the Lebesgue measure on  $\mathbb{T}^d$ . We will use  $\|f\|_{H^s}$  to denote the *homogeneous* Sobolev norms.

This paper is organized as follows: In Section 2 we describe the notion of  $\delta$ -mixed data, and prove Theorem 1.1, modulo a few Lemmas. In Section 3 we prove the required lemmas. In Section 4 we present numerics suggesting that the bound stated in Theorem 1.1 is not far from optimal, at least for an initial period of time. Finally, we conclude this paper with a scaling argument showing that an exponential lower bound on  $\|\theta\|_{H^{-1}}$  with rate independent of  $\theta_0$  will imply a stronger *algebraic* lower bound for mixing with flows satisfying no-slip boundary condition.

## 2. REARRANGEMENT COSTS AND THE PROOF OF THE MAIN THEOREM.

We devote this section to the proof of Theorem 1.1. The idea behind the proof is as follows. First, if  $\|\theta\|_{H^{-1}}$  is small enough, then its super-level sets are mixed to certain scales (Lemma 2.3 below). Second, any flow that starts with an “unmixed” set and mixes it to scale  $\delta$  has to do a minimum amount of work [5,12]. Putting these together yields Theorem 1.1.

We begin by describing the notion of “mixed to scale  $\delta$ ”, and relate this to the  $H^{-1}$  Sobolev norm.

**Definition 2.1.** Let  $\kappa \in (0, \frac{1}{2})$  be fixed. For  $\delta > 0$ , we say a set  $A \subseteq \mathbb{T}^d$  is  $\delta$ -semi-mixed if

$$\frac{m(A \cap B(x, \delta))}{m(B(x, \delta))} \leq 1 - \kappa \quad \text{for every } x \in \mathbb{T}^d.$$

If additionally  $A^c$  is also  $\delta$ -semi-mixed, then we say  $A$  is  $\delta$ -mixed (or mixed to scale  $\delta$ ).

*Remark 2.2.* The parameters  $\delta$  and  $\kappa$  measures the scale and “accuracy” respectively. The key parameter here is the scale  $\delta$ , and the accuracy parameter  $\kappa \in (0, 1/2)$  only plays an auxiliary role. Given a specific initial distribution to mix,  $\kappa$  can be chosen to optimize the bound.

Note that the notion of a set being mixed here is the same as that of Bressan [5]. A set being semi-mixed is of course a weaker notion.

One relation between  $\delta$ -semi-mixed and negative Sobolev norms is as follows.

**Lemma 2.3.** Let  $\lambda \in (0, 1]$  and  $\theta \in L^\infty(\mathbb{T}^d)$ . Then for any integer  $n > 0$ ,  $\kappa \in (0, \frac{\lambda}{1+\lambda})$  there exists an explicit constant  $c_0 = c_0(d, \kappa, \lambda, n)$  such that

$$\|\theta\|_{H^{-n}} \leq \frac{\|\theta\|_{L^\infty} \delta^{n+d/2}}{c_0} \implies A_\lambda \text{ is } \delta\text{-semi-mixed}.$$

Here  $A_\lambda$  is the super level set defined by  $A_\lambda \stackrel{\text{def}}{=} \{\theta > \lambda \|\theta\|_{L^\infty}\}$ .

Our interest in this Lemma is mainly when  $n = 1$ . Note that while Lemma 2.3 guarantees the super level sets  $A_\lambda$  are  $\delta$ -semi-mixed, they need not be  $\delta$ -mixed. Indeed if  $A_\lambda$  is very small, its complement won't be  $\delta$ -semi-mixed. Also, we remark that the converse of Lemma 2.3 need not be true. For example the function

$$f(x) = \sin(2\pi x) + 10 \sin(2\pi n x)$$

has  $\|f\|_{H^{-1}(\mathbb{T}^1)} = O(1)$ , and the super level set  $\{f > 5\}$  is certainly semi-mixed to scale  $1/n$  (see also [18]).

The proof of Lemma 2.3 follows from a duality and scaling argument. For clarity of presentation we postpone the proof to Section 3. Returning to Theorem 1.1, the main ingredient in its proof is a lower bound on the “amount of work” required to mix a set to fine scales. This notion goes back to a conjecture of Bressan for which a \$500 prize was announced [6].

**Conjecture 2.4** (Bressan '03 [5]). Let  $H$  to be the left half of the torus, and  $\Psi$  be the flow generated by an incompressible vector field  $u$ . If after time  $T$  the image of  $H$  under the flow  $\Psi$  is  $\delta$ -mixed, then there exists a constant  $C$  such that

$$(2.1) \quad \int_0^T \|\nabla u(\cdot, t)\|_{L^1} dt \geq \frac{|\ln \delta|}{C}.$$

We refer the reader to [5] for the motivation of the lower bound (2.1) and further discussion. To the best of our knowledge, this conjecture is still open. However, Crippa and De Lellis [12] made significant progress towards the resolution of this conjecture.

**Theorem 2.5** (Crippa, De Lellis '08 [12]). Using the same notation as in Conjecture 2.4, for any  $p > 1$  there exists a constant  $C_p$  such that

$$(2.2) \quad \int_0^T \|\nabla u(\cdot, t)\|_{L^p} dt \geq \frac{|\ln \delta|}{C_p}.$$

For our purposes we will need two extensions of Theorem 2.5. First, we will need to start with sets other than the half torus. Second, we will need lower bounds for the work done to *semi-mix* sets to small scales. Note that in order for a flow to  $\delta$ -mix a set  $A$ , it has to both  $\delta$ -semi-mix  $A$  and  $\delta$ -semi-mix  $A^c$ . Generically each of these steps should cost comparable amounts, and hence a semi-mixed version of Theorem 2.5 should follow using techniques in [12]. We state this as our next lemma.

**Lemma 2.6.** *Let  $\Psi$  be the flow map of an incompressible vector field  $u$ . Let  $A \subset \mathbb{T}^d$  be any measurable set and let  $p > 1$ . There exist constants  $r_0 = r_0(A)$  and  $a = a(d, \kappa, p) > 0$ , such that if for some  $\delta < r_0/2$  and  $T > 0$  the set  $\Psi_T(A)$  is  $\delta$ -semi-mixed, then*

$$(2.3) \quad \int_0^T \|\nabla u(\cdot, t)\|_{L^p} dt \geq \frac{m(A)^{1/p}}{a} \left| \ln \frac{2\delta}{r_0} \right|.$$

Morally the constant  $r_0$  above should be a length scale at which set  $A$  is not semi-mixed. Our proof, however, uses a condition on  $r_0$  which is slightly stronger than only requiring that  $A$  is not semi-mixed to scale  $r_0$ . Namely, we will require “most” of  $A$  to occupy “most” of the union of disjoint balls of radius at least  $r_0$ . Deferring the proof of Lemma 2.6 to Section 3, we prove Theorem 1.1.

*Proof of Theorem 1.1.* Replacing  $\theta$  with  $\theta/\|\theta\|_{L^\infty}$ , we may without loss of generality assume  $\|\theta_0\|_{L^\infty} = 1$ . Fix  $0 < \lambda \leq 1$ , and  $\kappa \in (0, \frac{\lambda}{1+\lambda})$ . Let  $a$  be the constant from Lemma 2.6, and  $c_0$  the constant from Lemma 2.3 with  $n = 1$ . Choose

$$\delta = \left( c_0 \|\theta(t)\|_{H^{-1}} \right)^{\frac{2}{d+2}}.$$

Then certainly  $\|\theta(t)\|_{H^{-1}} \leq \delta^{d/2+1}/c_0$  and by Lemma 2.3 the super level set  $\{\theta(t) > \lambda\}$  is  $\delta$ -semi-mixed.

Now, since  $\theta$  satisfies the transport equation (1.1), we know  $\{\theta(t) > \lambda\} = \Psi_t(A_\lambda)$ , where  $\Psi$  is the flow of the vector field  $u$ . Thus, Lemma 2.6 now implies

$$\delta \geq \frac{r_0}{2} \exp\left(\frac{-a}{m(A_\lambda)^{1/p}} \int_0^t \|\nabla u\|_{L^p} dt\right).$$

Consequently

$$\|\theta(t)\|_{H^{-1}} = \frac{\delta^{d/2+1}}{c_0} \geq \frac{r_0^{d/2+1}}{c_0 2^{d/2+1}} \exp\left(\frac{-da}{m(A_\lambda)^{1/p}} \int_0^t \|\nabla u\|_{L^p} dt\right),$$

finishing the proof.  $\square$

### 3. PROOFS OF LEMMAS.

We devote this section to the proofs of Lemmas 2.3 and 2.6.

*Proof of Lemma 2.3.* Suppose for the sake of contradiction that  $A_\lambda$  is not  $\delta$ -semi-mixed. Then by definition, there exists  $x \in \mathbb{T}^d$  such that

$$(3.1) \quad m(A_\lambda \cap B(x, \delta)) \geq (1 - \kappa)m(B(x, \delta)) = (1 - \kappa)\pi(d)\delta^d.$$

Here  $\pi(d)$  is the volume of  $d$ -dimensional unit ball.

By duality

$$(3.2) \quad \|\theta\|_{H^{-n}} = \sup_{f \in H^n} \frac{1}{\|f\|_{H^n}} \left| \int_{\mathbb{T}^d} \theta(x) f(x) dx \right|.$$

We choose  $f \in H^n$  to be a function which is identically equal to 1 in  $B(x, \delta)$ , and which vanishes outside  $B(x, (1 + \varepsilon)\delta)$  for some small  $\varepsilon > 0$ . A direct calculation shows that we can arrange

$$\|f\|_{H^n} \leq c_1(d) \cdot \varepsilon^{-n+\frac{1}{2}} \cdot \delta^{-n+\frac{d}{2}},$$

for some (explicit) constant  $c_1$  depending only on the dimension.

On the other hand using (3.1) gives

$$(3.3) \quad \int_{\mathbb{T}^d} \theta(x) f(x) dx \geq \pi(d) \|\theta\|_{L^\infty} \delta^d ((1 - \kappa)\lambda - \kappa - c_2(d)\varepsilon),$$

for some (explicit) dimensional constant  $c_2(d)$ . Choosing  $\varepsilon = \frac{\lambda - (1 + \lambda)\kappa}{2c_2(d)}$  and using (3.2) we obtain

$$\|\theta\|_{H^{-n}} \geq \frac{\|\theta\|_{L^\infty} \delta^{-n+\frac{d}{2}}}{c_0(d, \kappa, \lambda, n)}$$

as desired.  $\square$

*Remark.* Observe  $c_0 = c'_0(d, n)(\lambda - (1 + \lambda)\kappa)^{n-\frac{1}{2}}$ .

Now we turn to Lemma 2.6. For this, we need a result from [12] which controls the Lipschitz constant of the Lagrangian map except on a set of small measure.

**Proposition 3.1** (Crippa DeLellis '08 [12]). *Let  $\Psi(t, x)$  be the flow map of the (incompressible) vector field  $u$ . For every  $p > 1$ ,  $\eta > 0$ , there exists a set  $E \subset \mathbb{T}^d$  and a constant  $c = c(d, p)$  such that  $m(E^c) \leq \eta$  and for any  $t \geq 0$  we have*

$$(3.4) \quad \text{Lip}(\Psi^{-1}(t, \cdot)|_{E^c}) \leq \exp\left(\frac{c}{\eta^{\frac{1}{p}}} \int_0^t \|\nabla u(s)\|_{L^p} ds\right).$$

Here

$$\text{Lip}(\Psi^{-1}(t, \cdot)|_{E^c}) \stackrel{\text{def}}{=} \sup_{\substack{x, y \in E^c \\ x \neq y}} \frac{|\Psi^{-1}(t, x) - \Psi^{-1}(t, y)|}{|x - y|}$$

is the Lipschitz constant of  $\Psi^{-1}$  on  $E^c$ .

The proof of Proposition 3.1 is built upon the simple observation [2] that for a passive scalar  $\theta(x, t)$  and smooth advecting velocity  $u$  one has the inequality

$$(3.5) \quad \int \log_+ |\nabla \theta(t, \Psi(t, x))| dx \leq \int_0^t \int |\nabla u(t, \Psi(t, x))| dx.$$

This can be proved by an elementary calculation. In fact, even the point wise bound

$$D \log |\nabla \theta| \leq |\nabla u|$$

is true, where  $D = \partial_t + u \cdot \nabla$  is the material derivative. In the form (3.5), this inequality is not very useful. But it turns out that the more sophisticated maximal form of this inequality [2, 12] can be much more useful and is essentially what leads to Proposition 3.1. We refer the reader to [12] for the details of the proof.

We use Proposition 3.1 to prove Lemma 2.6 below.

*Proof of Lemma 2.6.* The main idea behind the proof is as follows: Suppose first  $r_0$  is some large scale at which the set  $A$  is “not semi-mixed”. Let  $T > 0$  be fixed and suppose  $\Psi_T(A)$  is  $\delta$ -semi-mixed for some  $\delta < r_0/2$ . Since  $\Psi_T(A)$  is  $\delta$ -semi-mixed, there should be many points  $\tilde{x} \in \Psi_T(A)$  and  $\tilde{y} \in \Psi_T(A)^c$  such that  $|\tilde{x} - \tilde{y}| < \delta$ . Since  $A$  is “not semi-mixed” to scale  $r_0$ , there should be many points  $\tilde{x}$  and  $\tilde{y}$  so

that we additionally have  $|\Psi_T^{-1}(\tilde{x}) - \Psi_T^{-1}(\tilde{y})| \geq r_0/2$ . This will force the Lipschitz constant of  $\Psi_T^{-1}$  to be at least  $r_0/(2\delta)$  on a set of large measure. Combined with Proposition 3.1 this will give the desired lower bound on  $\int_0^t \|\nabla u\|_{L^p}$ .

We now carry out the details of the above outline. The first step in the proof is to choose the length scale  $r_0$ . Let  $\varepsilon = \varepsilon(\kappa, d)$  be a small constant to be chosen later. We claim that there exists a natural number  $l$  and finitely many disjoint balls  $B(x_1, r_1), \dots, B(x_l, r_l)$  such that

$$(3.6) \quad m\left(\bigcup_{i=1}^l B(x_i, r_i)\right) \geq \frac{m(A)}{2 \cdot 3^d} \quad \text{and} \quad \frac{m(A \cap B(x_j, r_j))}{m(B(x_j, r_j))} > 1 - \varepsilon$$

for every  $j \in \{1, \dots, l\}$ .

To see this, note that the metric density of  $A$  is 1 almost surely in  $A$ . Thus, removing a set of measure 0 from  $A$  if necessary, we know that for every  $x \in A$  there exists an  $r \in (0, 1]$  such that

$$\frac{m(A \cap B(x, r))}{m(B(x, r))} > 1 - \varepsilon.$$

Now choose  $K \subset A$  compact with  $m(K) > m(A)/2$ . Since the above collection of balls is certainly a cover of  $K$ , we pass to a finite sub-cover. Applying Vitali's lemma to this sub-cover we obtain a disjoint sub-family  $\{B(x_i, r_i) \mid i = 1, \dots, l\}$  with  $m(\cup B(x_i, r_i)) \geq m(K)/3^d$ . This immediately implies (3.6). For convenience let  $B_i = B(x_i, r_i)$ , and choose  $r_0 = \min\{r_1, \dots, r_l\}$ .

Now let  $\eta > 0$  be another small parameter that will be chosen later. By Proposition 3.1 we know that there exists a set  $E$  with  $m(E) \leq \eta$  such that the inequality (3.4) holds. Define the set

$$(3.7) \quad F = \{x \in \mathbb{T}^d \mid \frac{m(B(x, \delta) \cap E)}{m(B(x, \delta))} > \frac{\kappa}{2}\}$$

Clearly  $F \subset \{M\chi_E > \kappa/2\}$ , where  $M\chi_E$  is the maximal function of  $\chi_E$ . Consequently,

$$m(F) \leq m(\{M\chi_E > \frac{\kappa}{2}\}) \leq \frac{2c_1}{\kappa} m(E)$$

for some explicit constant  $c_1 = c_1(d)$ . (It is well known that  $c_1 = 3^d$  will suffice.)

Since  $\Psi_T$  is measure preserving we know  $m(\Psi_T^{-1}(E \cup F)) \leq (1 + 2c_1/\kappa)\eta$ . Thus choosing

$$\eta = \frac{\kappa}{\kappa + 2c_1} \left( \frac{1}{4^d} - \varepsilon \right) \sum_{i=1}^l m(B_i)$$

will guarantee

$$m(\Psi_T^{-1}(E \cup F)) \leq \left( \frac{1}{4^d} - \varepsilon \right) \sum_{i=1}^l m(B_i).$$

This implies that for some  $i_0 \leq l$  we must have

$$(3.8) \quad m((B_{i_0} \cap A) - \Psi_T^{-1}(E \cup F)) \geq \left(1 - \frac{1}{4^d}\right) m(B_{i_0}).$$

By reordering, we may without loss of generality assume that  $i_0 = 1$ . Consequently, for

$$C_1 = \left\{ x \in (B_1 \cap A) - \Psi_T^{-1}(E \cup F) \mid d(x, B_1^c) > \frac{r_1}{2} \right\}.$$



equation (3.8) implies

$$m(C_1) \geq \left( \frac{1}{2^d} - \frac{1}{4^d} \right) m(B_1).$$

Now, from the collection of open balls  $\{B(\tilde{x}, \delta) \mid \tilde{x} \in \Psi_T(C_1)\}$  the Vitali covering lemma allows us to extract a finite disjoint collection  $B(\tilde{x}_1, \delta), \dots, B(\tilde{x}_n, \delta)$  such that

$$m\left(\bigcup_{i=1}^n B(\tilde{x}_i, \delta)\right) \geq \frac{m(C_1)}{5^d}.$$

Our goal is to find  $\tilde{y}$  such that  $\tilde{y} \in B(\tilde{x}_i, \delta) - E$  for some  $i$ , and  $|\Psi_T^{-1}\tilde{y} - \Psi_T^{-1}x| > r_1/2$ .

For convenience set  $\tilde{B}_i = B(\tilde{x}_i, \delta)$ . Since  $\Psi_T(A)$  is  $\delta$ -semi-mixed and  $\tilde{x}_i \notin F$  we have

$$(3.9) \quad m(\Psi_T(A) \cap \tilde{B}_i) \leq (1 - \kappa)m(\tilde{B}_i) \quad \text{and} \quad m(E \cap \tilde{B}_i) \leq \frac{\kappa}{2}m(\tilde{B}_i).$$

Also, since  $\Psi_T$  is measure preserving and by the definition of  $B_1$  we see

$$(3.10) \quad m\left(\bigcup_{i=1}^n \tilde{B}_i \cap \Psi_T(B_1 - A)\right) \leq m(B_1 - A) < \varepsilon m(B_1)$$

Using the fact that  $\{\tilde{B}_i\}$  are all disjoint, summing (3.9) and using (3.10) gives

$$\begin{aligned} m\left(\bigcup_{i=1}^n \tilde{B}_i \cap E \cap \Psi_T(B_1)\right) &< \left(1 - \frac{\kappa}{2}\right) \sum_{i=1}^n m(\tilde{B}_i) + \varepsilon m(B_1) \\ &\leq \left(1 - \frac{\kappa}{2} + \varepsilon 5^d \left(\frac{1}{2^d} - \frac{1}{4^d}\right)^{-1}\right) \sum_{i=1}^n m(\tilde{B}_i). \end{aligned}$$

Thus choosing

$$\varepsilon < \frac{\kappa}{2 \cdot 5^d} \left( \frac{1}{2^d} - \frac{1}{4^d} \right)$$

will guarantee

$$m\left(\bigcup_{i=1}^n \tilde{B}_i \cap E \cap \Psi_T(B_1)\right) < m\left(\bigcup_{i=1}^n \tilde{B}_i\right)$$

This in turn will guarantee that for some  $i$  we can find  $\tilde{y} \in \tilde{B}_i - E - \Psi_T(B_1)$ .

Now observe that

$$\tilde{y}, \tilde{x}_i \notin E, \quad |\tilde{y} - \tilde{x}_i| < \delta, \quad \text{and} \quad |\Psi_T^{-1}(\tilde{y}) - \Psi_T^{-1}(\tilde{x}_i)| > \frac{r_1}{2}.$$

The last inequality above follows because  $\Psi_T^{-1}(\tilde{x}_i) \in C_1$  and  $\Psi_T^{-1}(\tilde{y}) \notin B_1$ . This forces

$$\text{Lip}(\Psi_T^{-1}|_{E^c}) \geq \frac{|\Psi_T^{-1}(\tilde{y}) - \Psi_T^{-1}(\tilde{x}_i)|}{|\tilde{y} - \tilde{x}_i|} > \frac{r_1}{2\delta} \geq \frac{r_0}{2\delta}.$$

Now using (3.4), and letting  $a = a(d, \kappa, p)$  denote a constant that changes from line to line we obtain

$$(3.11) \quad \int_0^T \|\nabla u(t)\|_{L^p} dt \geq \frac{\eta^{\frac{1}{p}}}{a} \left| \log\left(\frac{r_0}{2\delta}\right) \right|.$$

Observe finally that

$$\eta = c_2 m\left(\bigcup_{i=1}^l B_i\right) \geq \frac{c_2 m(A)}{2 \cdot 3^d}$$

for some explicit constant  $c_2 = c_2(d, \kappa)$ . Consequently (3.11) reduces to

$$\int_0^T \|\nabla u(t)\|_{L^p} dt \geq \frac{m(A)^{\frac{1}{p}}}{a} |\log(\frac{r_0}{2\delta})|,$$

as desired.  $\square$

#### 4. NUMERICAL RESULTS

In this section we present numerical results illustrating how the exponential decay rate varies with the initial data. For numerical purposes we work on the 1-periodic torus. Given a parameter  $a$ , we define the initial data  $\theta_0 = \theta'_0 / \|\theta'_0\|_{L^2}$  where

$$\theta'_0(x, y) = \begin{cases} \sin(\frac{2\pi x}{a}) \sin(\frac{2\pi(y + \frac{a}{8})}{a}) & \text{for } 0 < x < \frac{a}{2} \text{ and } \frac{-a}{8} < y < \frac{a}{2} - \frac{a}{8} \\ \sin(\frac{2\pi x}{a}) \sin(\frac{2\pi(y - \frac{a}{8})}{a}) & \text{for } \frac{a}{2} < x < a \text{ and } \frac{a}{8} < y < \frac{a}{2} + \frac{a}{8} \\ 0 & \text{otherwise.} \end{cases}$$

A figure of this is shown in 1(a).

We do not know which velocity field achieves the lower bound (1.2). However the steepest descent method introduced in [18] provides us with a reasonable candidate. Explicitly, their formula gives

$$(4.1) \quad u = \frac{-\Delta^{-1} P(\theta \nabla^{-1} \theta)}{\|\nabla^{-1} P(\theta \nabla^{-1} \theta)\|_{L^2}},$$

where  $P$  is the Leray-Hodge projection onto divergence free vector fields. This can be derived by multiplying both sides of (1.1) by  $\Delta^{-1} \theta$  and integrating by parts.

Using a pseudo-spectral method<sup>1</sup> retaining 768 Fourier modes in each variable we perform a numerical simulation of (1.1) with the initial data obtained by varying the parameter  $a$  over the set  $\{6/12, 7/12, \dots, 11/12\}$ , and the velocity obtained dynamically using (4.1). Plots of our solutions at various times (for  $a = 11/12$ ) are shown in Figure 1.

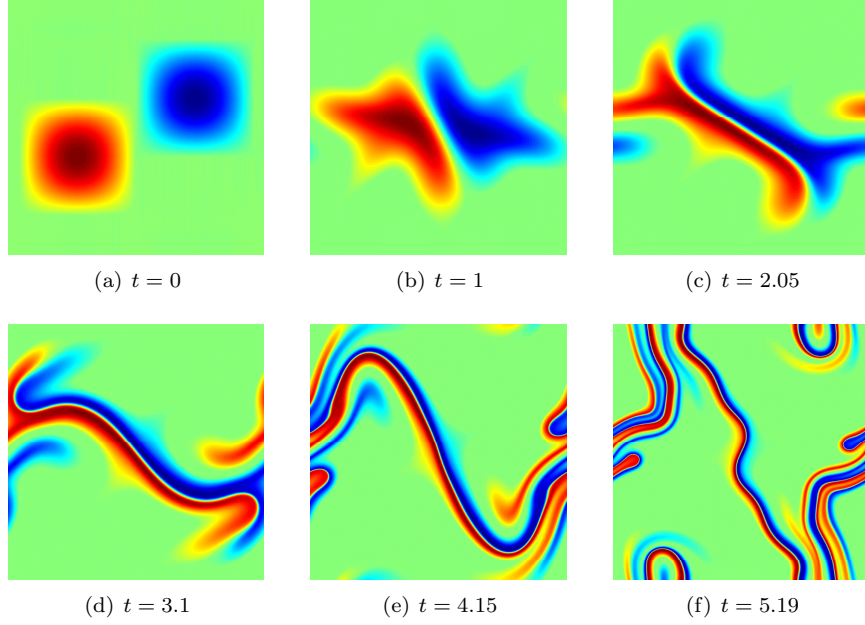
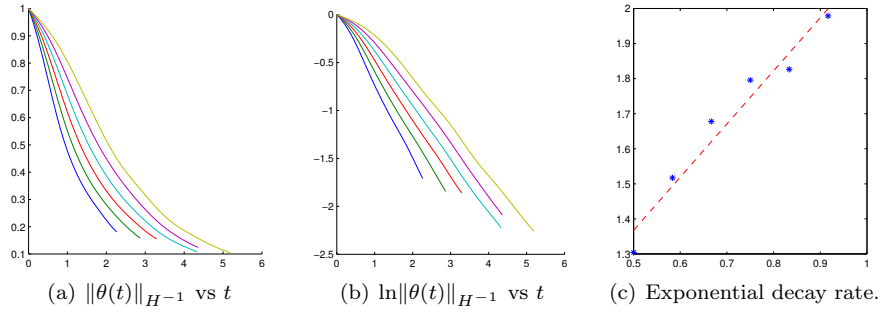
Figure 2(a) shows graphs of  $\|\theta(t)\|_{H^{-1}}$  vs  $t$  as the parameter  $a$  varies over the set  $\{6/12, \dots, 11/12\}$ . Figure 2(b) shows graphs of  $\ln \|\theta(t)\|_{H^{-1}}$  vs  $t$  for the same values of  $a$ . Following a short initial “settling down” period, the log plots in Figure 2(b) are essentially linear indicating an exponential in time decay of  $\|\theta_0\|_{H^{-1}}$ .

We fit each of the log plots in Figure 2(b) to a straight line, and plot the negative reciprocal of the slope vs  $a$  in Figure 2(c). Since  $m(\text{supp}(\theta_0)) = O(a^2)$ , Theorem 1.1 predicts this graph to be linear as a function of  $a$ . This is in good agreement with the observed numerics.

#### 5. A SCALING ARGUMENT AND UNIVERSAL DECAY RATES.

Physical intuition suggests that the exponential decay rate in (1.2) should have some dependence on the size of support of  $\theta_0$ . As we discussed in the introduction, the mixing process can spread around the compactly supported initial data. Whether this has to happen in the mixing process, and whether this leads to slow-down in the mixing rate are very interesting open questions. In this section we show that exponential in time lower bound on the decay of the mix norm with

<sup>1</sup>The code and more figures can be downloaded from [1].


 FIGURE 1. Solution plots at various times for  $a = 11/12$ .

 FIGURE 2. The mix norm of the scalar density (Figures (a) & (b)), and the negative reciprocal of the exponential decay rate vs  $a$  as  $a$  varies over  $\{6/12, \dots, 11/12\}$  (Figure (c)).

the rate in the exponential independent of the initial data would have interesting consequences for mixing in domains with no slip boundaries.

**Proposition 5.1.** *Let  $I = (0, \ell)^d$  be a cube in  $\mathbb{R}^d$ . Suppose that there exist  $k \in \mathbb{R}$ ,  $q \in [1, \infty]$  and  $c_0 > 0$  such that*

$$(5.1) \quad \|\theta(t)\|_{H^{-1}} \geq B(\theta_0) \exp\left(\frac{-c_0}{\ell^{d/p}} \int_0^t \|\nabla u\|_{L^p}\right)$$

for some  $p \in [1, d]$ , all incompressible  $u$  which vanish on  $\partial I$ , and all initial data  $\theta_0 \in C_c^\infty(I)$ . Assume that there exists  $\gamma \in \mathbb{R}$  such that the pre-factor  $B(\theta_0)$  satisfies

$$B(A\theta_0) = AB(\theta_0) \quad \text{and} \quad B(\theta_{0,a}) = a^{-\gamma}B(\theta_0),$$

where  $\theta_{0,a}(x) = \theta_0(x/a)$  for  $x \in (0, a)^d$  and  $\theta_{0,a}(x) = 0$  otherwise.

Then, for any mean zero  $\theta_0 \in C_c^\infty(I)$  and any smooth velocity field  $u$  such that

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t \|\nabla u\|_{L^p} < \infty, \quad \nabla \cdot u = 0, \quad \text{and} \quad u = 0 \text{ on } \partial I,$$

the decay of  $\|\theta(t)\|_{H^{-1}}$  as  $t \rightarrow \infty$  is bounded below by an algebraic function of time.

*Proof.* We prove this using an elementary scaling argument. Without loss of generality, assume  $\ell = 1$ . Let  $\theta_0 \in C_c^\infty$  have zero mean and let  $f(t) = \|\theta(t)\|_{H^{-1}}$ . Our aim is to show that for  $t$  large,  $f$  decays algebraically with respect to  $t$ . Let  $a > 0$ , and define

$$I_a = (0, a)^d, \quad \eta(x, t) = \chi_{I_a} \theta(x/a, t/a), \quad v(x, t) = \chi_{I_a} u(x/a, t/a).$$

Then  $(\eta, v)$  is a solution of (1.1) on the *unscaled* cube  $I$ . Since  $\theta_0$  is compactly supported in  $I$  and  $u = 0$  on  $\partial I$ , we see  $\theta(t)$  remains compactly supported in  $I_a$  for all  $t > 0$ . By our assumption,

$$B(\eta_0) = a^{-\gamma}B(\theta_0) \quad \text{and} \quad \|\nabla v(t)\|_{L^p} = a^{\frac{d}{p}-1} \|\nabla u(t/a)\|_{L^p}.$$

Thus the assumed lower bound (5.1) gives

$$\begin{aligned} a^{d/2+1} f(t/a) = \|\eta(t)\|_{H^{-1}} &\geq B(\eta_0) \exp\left(-c_0 \int_0^t \|\nabla v\|_{L^p}\right) \\ &= a^{-\gamma} B(\theta_0) \exp\left(-c_0 a^{\frac{d}{p}-1} \int_0^t \|\nabla u(s/a)\|_{L^p} ds\right). \end{aligned}$$

The first equality above follows by duality and scaling.

Hence, taking  $t' = t/a$  gives

$$f(t') \geq a^{-N} B(\theta_0) \exp\left(-c_0 a^{d/p} \int_0^{t'} \|\nabla u(s)\|_{L^p} ds\right),$$

where  $N = d/2 + 1 + \gamma$ . This bound has to be true for every  $a > 0$ . Maximizing the right hand side in  $a$  (and changing  $t'$  to  $t$ ), we arrive at an algebraic lower bound

$$f(t) \geq C \left( \frac{N}{\int_0^t \|\nabla u\|_{L^p} ds} \right)^{-pN/d}. \quad \square$$

## 6. ACKNOWLEDGEMENTS.

The authors would like to thank Charlie Doering for introducing us to this problem and many helpful discussions.

## REFERENCES

- [1] <http://www.math.cmu.edu/~gautam/research/201208-mix-bounds>.
- [2] L. Ambrosio, M. Lecumberry, and S. Maniglia. Lipschitz regularity and approximate differentiability of the DiPerna-Lions flow. *Rend. Sem. Mat. Univ. Padova*, 114:29–50 (2006), 2005.
- [3] K. Bardos and È. S. Titi. Euler equations for an ideal incompressible fluid. *Uspekhi Mat. Nauk*, 62(3(375)):5–46, 2007.
- [4] Y. Brenier, F. Otto, and C. Seis. Upper bounds on coarsening rates in demixing binary viscous liquids. *SIAM J. Math. Anal.*, 43(1):114–134, 2011.
- [5] A. Bressan. A lemma and a conjecture on the cost of rearrangements. *Rend. Sem. Mat. Univ. Padova*, 110:97–102, 2003.
- [6] A. Bressan. Prize offered for the solution of a problem on mixing flows. <http://www.math.psu.edu/bressan/PSPDF/prize1.pdf>, 2006.
- [7] L. A. Caffarelli and A. Vasseur. Drift diffusion equations with fractional diffusion and the quasi-geostrophic equation. *Ann. of Math. (2)*, 171(3):1903–1930, 2010.
- [8] D. Chae, P. Constantin, D. Córdoba, F. Gancedo, and J. Wu. Generalized surface quasi-geostrophic equations with singular velocities. *Comm. Pure Appl. Math.*, 65(8):1037–1066, 2012.
- [9] P. Constantin and C. Foias. *Navier-Stokes equations*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1988.
- [10] P. Constantin, A. Kiselev, L. Ryzhik, and A. Zlatoš. Diffusion and mixing in fluid flow. *Ann. of Math. (2)*, 168(2):643–674, 2008.
- [11] P. Constantin and V. Vicol. Nonlinear maximum principles for dissipative linear nonlocal operators and applications. *Geom. Funct. Anal.*, 22(5):1289–1321, 2012.
- [12] G. Crippa and C. De Lellis. Regularity and compactness for the DiPerna-Lions flow. In *Hyperbolic problems: theory, numerics, applications*, pages 423–430. Springer, Berlin, 2008.
- [13] C. R. Doering and J. D. Gibbon. *Applied analysis of the Navier-Stokes equations*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 1995.
- [14] C. R. Doering and J.-L. Thiffeault. Multiscale mixing efficiencies for steady sources. *Phys. Rev. E*, 74(2, 2), AUG 2006.
- [15] E. Gouillart, O. Dauchot, B. Dubrulle, S. Roux, and J.-L. Thiffeault. Slow decay of concentration variance due to no-slip walls in chaotic mixing. *Phys. Rev. E (3)*, 78(2):026211, 16, 2008.
- [16] E. Gouillart, N. Kuncio, O. Dauchot, B. Dubrulle, S. Roux, and J.-L. Thiffeault. Walls inhibit chaotic mixing. *Phys. Rev. Lett.*, 99:114501, Sep 2007.
- [17] A. Kiselev, F. Nazarov, and A. Volberg. Global well-posedness for the critical 2D dissipative quasi-geostrophic equation. *Invent. Math.*, 167(3):445–453, 2007.
- [18] Z. Lin, J.-L. Thiffeault, and C. R. Doering. Optimal stirring strategies for passive scalar mixing. *J. Fluid Mech.*, 675:465–476, 2011.
- [19] 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), 2012.
- [20] G. Mathew, I. Mezić, and L. Petzold. A multiscale measure for mixing. *Phys. D*, 211(1-2):23–46, 2005.
- [21] C. Seis. Maximal mixing by incompressible fluid flows. 2013. Preprint.
- [22] 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.
- [23] D. Slepčev. Coarsening in nonlocal interfacial systems. *SIAM J. Math. Anal.*, 40(3):1029–1048, 2008.
- [24] J.-L. Thiffeault. Using multiscale norms to quantify mixing and transport. *Nonlinearity*, 25(2):R1–R44, 2012.
- [25] 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.

GAUTAM IYER, DEPARTMENT OF MATHEMATICS, CARNEGIE MELLON UNIVERSITY, PITTSBURGH,  
PA 15213

*E-mail address:* `gautam@math.cmu.edu`

ALEXANDER KISELEV, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN-MADISON, MADISON,  
WI 53706, USA

*E-mail address:* `kiselev@math.wisc.edu`

XIAOQIAN XU, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN-MADISON, MADISON,  
WI 53706, USA

*E-mail address:* `xxu@math.wisc.edu`