Diffusion limited mixing rates in passive scalar advection

Christian Seis

Institut für Analysis und Numerik, Westfälische Wilhelms-Universität Münster, Orléans-Ring 10, 48149 Münster, Germany.
E-mail address: seis@wwu.de

Abstract: We are concerned with flow enhanced mixing of passive scalars in the presence of diffusion. Under the assumption that the velocity gradient is suitably integrable, we provide upper bounds on the exponential rates of mixing and of enhanced dissipation. Our results suggest that there is a crossover from advection dominated to diffusion dominated mixing, and we observe a slow down in the exponential decay rates by (some power of) a logarithm of the diffusivity.

1 Introduction

The theory of fluid mixing has become an active area of research in the applied mathematics community in recent years. Mixing refers to the homogenization process of a heterogenously distributed physical quantity and can be driven by diffusion or the result of advection by a straining fluid flow. Each of these transport mechanisms has a different action on the mixture. While diffusion balances local differences in concentration, which results thus in a decay in the concentration intensity, advection creates finer and finer filaments and acts thus on the scale of fluctuations.

A dominant feature of chaotic or turbulent fluid motions is the continual transfer of concentration towards smaller and smaller length scales. The progressing filamentation is accompanied by the creation of concentration gradients, which are subsequently diminished by diffusion. The action of advection is thus limited to length scales above a certain critical size, below which the smoothing effect of diffusion dominates and prevents any further spatial refinements. It is the spatial scale at which advection and diffusion achieve a balance. It was first identified by Batchelor and is now commonly referred to as the Batchelor scale [3]. When advection becomes ineffective for mixing, the large time behavior is governed by the reduction in the concentration intensity.

The mathematical model describing such mixing processes is the advection-diffusion equation

\[ \partial_t \theta^\kappa + u \cdot \nabla \theta^\kappa = \kappa \Delta \theta^\kappa, \]  

(1)
that we consider, for simplicity, in the periodic box $T^d = [0,1]^d$. Here $\theta^x$ is the physical quantity of interest and $u$ is the divergence-free velocity of the fluid. The constant $\kappa$ is the diffusivity and can be interpreted as the inverse of the Pécelet number in the non-dimensionalized setting. We shall always suppose that the diffusivity is small but finite, $\kappa \ll 1$. The above equation is linear as it is assumed that the observed quantity has no feedback on the fluid flow itself — think, for instance, of dye in water. Such quantities are in the literature often referred to as passive scalars.

In this paper, we are concerned with the question of how fast mixing can happen if the fluid velocity field belongs to the class $L^s(\mathbb{R};W^{1,p}(\mathbb{R}^d))$ for some $p \in (1,\infty]$ and $s \in [1,\infty]$. More precisely, we shall always assume that $u$ is arbitrarily given such that $\|\nabla u\|_{L^s_tL^p_x}$ is bounded by 1. Note that different values of $s$ correspond to different application scenarios. For instance, if $s = \infty$ (and $p = 2$), the fluid velocity could origin from an industrial stirring process, in which $\|\nabla u\|_{L^2_x}$ is the amount of work an agent spends per time unit to maintain stirring. The case $s = 2$ is natural in applications in which the energy balance law for the Navier–Stokes equations guarantee a control over $\|\nabla u\|_{L^2_x}$ even in the presence of a rough volume force.

In order to ensure the well-posedness of distributional solutions for the non-diffusive model, cf. [18, 30], we shall moreover suppose that the initial concentration distribution $\theta_0$ belongs to the space $L^q(\mathbb{R}^d)$ for some $q \in [1,\infty]$ such that $1/p + 1/q \leq 1$. For convenience, we also assume that the initial configuration has zero mean, which is propagated by (1).

In the typical scenario, when a chaotic or turbulent flow is applied to a distribution of order one scale and intensity, advection is initially much more efficient as a mixing mechanism than molecular diffusion. In this early mixing stage, the situation is thus comparable to advection in the non-diffusive setting. Here, the underlying mathematical model is the transport equation

$$\partial_t \theta + u \cdot \nabla \theta = 0, \quad (2)$$

for the passive scalar $\theta$.

Estimates on mixing by advection were first derived in [16], at least, on the level of the Lagrangian flow, and then translated to the Eulerian setting in [7, 34, 22, 25]; see also [27] for the case of Lipschitz flows. A typical result of these papers indicates that mixing under the (generalized) enstrophy constraint $\|\nabla u\|_{L^s_tL^p_x} \leq 1$ cannot proceed faster than exponentially in time, more precisely,

$$e^{-A t} \lesssim e^{-A \int_0^t \|\nabla u\|_{L^p} dt} \lesssim \|\theta(t)\|_{H^{-1}}, \quad (3)$$

for some $A > 0$ and any $t > 0$. Here, mixing is measured in terms of the (homogeneous) $H^{-1}$ norm. There are several good reasons for choosing negative Sobolev norms (or related quantities such as Kantorovich–Rubinstein or Wasserstein distances borrowed from the theory of optimal transportation) to quantify mixing by

\footnote{Here and in the sequel, we write $A \lesssim B$ if there exists a uniform constant $C$ independent of $\kappa$ such that $A \leq CB$. Moreover, we write $A \sim B$ if $A \lesssim B$ and $B \lesssim A$. Finally $A \ll B$ if $A \leq CB$ for some sufficiently small $C$. We also use the convention that $\frac{\infty}{\infty} = 1$.}
advection: They metrize weak convergence \[26\], correspond to mixing in the sense of ergodic theory \[28\], and have the dimensions of (a positive power of) a length scale. The mixing estimate (3) can thus be interpreted as an estimate on the decay of the average variation length scale, or, as a lower bound on the rate of weak convergence towards the “perfectly mixed state” \(\theta \equiv 0\). Strong Lebesgue norms are, however, not suitable as they are preserved by the purely advective flow,

\[
\|\theta(t)\|_{L^q} = \|\theta_0\|_{L^q},
\]

for any \(t > 0\).

We remark that the lower bound in (3) in the case \(p = 1\) is still open \[9\]. On the positive side, recent flow constructions show that the lower bound in (3) is sharp \[17, 39, 1, 20, 5\].

In a certain sense, the mixing bound in (3) is the Eulerian (and Sobolev) version of the well-known estimate for the distance of particle trajectories in Lipschitz flows,

\[
- e^{-\int_0^t \|\nabla u\|_{L^\infty} \, dt} \leq \frac{|X(t,x) - X(t,y)|}{|x - y|} \leq e^{\int_0^t \|\nabla u\|_{L^\infty} \, dt},
\]

where \(X\) is the Lagrangian flow for the vector field \(u\), given as the solution of the ordinary differential equation \(\partial_t X(t,x) = u(t, X(t,x))\) with \(X(0,x) = x\). This point of view has been further elaborated on in \[35, 36\]. An estimate analogous to (5) in the case of Sobolev vector fields was established earlier in \[16\].

We conclude the discussion of (3) with the qualitative observation that the lower bound excludes the possibility of perfect mixing in finite time. This, however, is not at all surprising because (2) is linear, time-reversible, and well-posed \[18\]. Moreover, if \(s = 1\), the vector field is decaying very quickly in time so that no mixing occurs, even in the large time limit.

Our first result indicates that early stage mixing in the diffusive setting (1) is indeed dominated by advection and is thus characterized by the reduction of length scales rather than the decay in intensity. In fact, our result shows that mixing in the presence of diffusion occurs initially at the same rate as in the purely advective case (2). Moreover, it is established that the energy \(\|\theta^\kappa\|_{L^2}\) remains of order one, similar to the conservative case in (4). Unfortunately, the validity of this first result is proved for Lipschitz flows only.

**Theorem 1.** Let \(\theta_0\) be a mean-zero initial configuration satisfying \(\|\theta_0\|_{L^2} \sim 1\) and \(\|\nabla \theta_0\|_{L^2} \sim 1\). Suppose, moreover, that \(u\) is a Lipschitz vector field, satisfying \(\|\nabla u\|_{L^s,W^\kappa} \leq 1\). Let \(T_{\kappa,s}\) be given by

\[
T_{\kappa,s} = \begin{cases} 
\log \frac{s}{s-1} \frac{1}{\kappa} & \text{if } s > 1, \\
\frac{1}{\kappa} & \text{if } s = 1.
\end{cases}
\]

If there is a positive constant \(A\) such that

\[
\|\theta(t)\|_{H^{-1}} \lesssim e^{-A t^{\frac{s-1}{s}}},
\]

3
for any $t \ll T_{\kappa,s}$, then also
\[ \|\theta^\kappa(t)\|_{\dot{H}^{-1}} \lesssim e^{-\frac{At}{s}}. \] (7)

for any $t \ll T_{\kappa,s}$. Similarly, if
\[ e^{-\frac{At}{s}} \lesssim \|\theta(t)\|_{\dot{H}^{-1}}, \] (8)

for any $t \ll T_{\kappa,s}$, then also
\[ e^{-\frac{At}{s}} \lesssim \|\theta^\kappa(t)\|_{\dot{H}^{-1}}, \] (9)

for any $t \ll T_{\kappa,s}$. Moreover, it holds that
\[ 1 \lesssim \|\theta^\kappa(t)\|_{L^2}, \] (10)

for any $t \ll T_{\kappa,s}$.

The theorem also holds true if the $\dot{H}^{-1}$ norm in (6)–(9) is replaced by some Wasserstein distance. In this case, the assumption on the initial datum can be slightly relaxed to requiring that $\|\theta_0\|_{L^1} \sim 1$. Because negative Sobolev norms are more popular in the mathematical mixing community, we will not further elaborate on this observation.

It is interesting to note that the time scale $T_{\kappa,\infty} = \log 1/\kappa$ agrees with the one identified by Batchelor that is needed to transfer energy from order one frequencies to those around the Batchelor scale, cf. [3]. In a certain sense, our findings in this paper confirm Batchelor’s identification. Moreover, we will see in Theorems 2 and 3 below that $T_{\kappa,s}$ is precisely the time scale that becomes characteristic at later times.

At later times, mixing is essentially governed by the decay of the concentration intensity. This decay can be measured in terms of any Lebesgue or (positive or negative) Sobolev norm on $\mathbb{T}^d$. Consider, for instance, the energy equality in the $L^2$ setting,
\[ \|\theta^\kappa(t)\|_{L^2}^2 + 2\kappa \int_0^t \|\nabla \theta^\kappa\|_{L^2}^2 \, dt = \|\theta_0\|_{L^2}^2. \] (11)

Because filamentation creates large concentration gradients, the energy equality suggests that dissipation must happen at much smaller time scales than in the purely diffusive situation. This effect is commonly referred to as enhanced dissipation [12, 38, 21, 14]. Heuristically, diffusive dissipation rates are determined, roughly, by the largest relevant frequencies, and are thus governed by the Batchelor scale.

It has been recently proved in [14] that, for a certain class of flows, namely Lipschitz continuous universal mixers that mix at a rate $e^{-\frac{At}{s}}$ for some $s \in (1, \infty]$, dissipation rates increase up to
\[ \|\theta^\kappa(t)\|_{L^2} \lesssim e^{-C/\sqrt{\log \frac{\kappa}{\kappa'}}}, \] (12)

for some $C > 0$ and any $t > 0$, compared to $O(e^{-C\kappa t})$ dissipation in the purely diffusive setting. Not unexpected, the rate of exponential decay in (12) falls down
to zero in the limit of vanishing diffusivity $\kappa \to 0$, which is consistent with the conservation of Lebesgue norms that is known to hold in the non-diffusive case \cite{[11]}

Our second result shows that exponential decay rates indeed have to be diffusivity dependent.

**Theorem 2.** Suppose that $u \in L^s(\mathbb{R}^d; W^{1,p}(\mathbb{T}^d))$ for some $p \in (1, \infty]$ and $s \in [1, \infty]$, satisfying $\|\nabla u\|_{L^1} \leq 1$. Let $q \in [1, \infty]$ be such that $\frac{1}{p} + \frac{1}{q} = 1$ and let $r \in [1, \infty]$ be such that $r \geq q$ if $\frac{1}{p} + \frac{1}{q} = 1$. Let $\theta_0$ be a mean-zero initial configuration satisfying $\|\theta_0\|_{L^1} \sim 1$ and $\|
abla \theta_0\|_{L^1} \sim 1$. If there is a positive constant $D$ such that

$$\|\theta^*(t)\|_{L^r} \lesssim e^{-Dt},$$

for any $t > 0$, then

$$D \lesssim \begin{cases} \log \frac{1}{s} & \text{if } s > 1, \\ \kappa & \text{if } s = 1. \end{cases} \tag{14}$$

The theorem implies that dissipation in fluid mixing with incompressible vector fields in the class $L^s(\mathbb{R}^d; W^{1,p}(\mathbb{T}^d))$ cannot proceed faster than with an exponential rate of order $\log^{-\frac{s}{s-1}} \frac{1}{\kappa}$ if $s > 1$. Moreover, by comparing the logarithmic dependence on the diffusivity in (12) with (14), we see that the results are optimal possibly up to a power $\alpha \in [1, 2]$ on the exponent $\log^{-\frac{s}{s-1}} \frac{1}{\kappa}$. Finally, in the case $s = 1$, a significantly enhanced dissipation — in the sense that the exponent shows a better $\kappa$-dependence compared to the purely diffusive setting — is a priori excluded. In view of (3), which tells us that $L^1(\mathbb{R}^d; W^{1,p}(\mathbb{T}^d))$ vector fields are unable to mix non-diffusive mixtures perfectly, such a behavior is not at all unexpected.

Around the same time a first version of the present paper was distributed by the author, Bruè and Nguyen uploaded a paper on the arXiv, that contains (among other results on diffusive mixing) estimates that are similar to (but slightly weaker than) those in Theorem 2 cf. \cite{[11]}. Indeed, in this work, the case $s = \infty$ and $p > 2$ is treated and $D$ is bounded by $\log^{-\frac{1}{p-1}} \frac{1}{\kappa}$.

The result in Theorem 2 has to be distinguished from the lower bound

$$e^{-C(\log^{-\frac{s}{s-1}} \frac{1}{\kappa})t} \lesssim \|\theta^*(t)\|_{L^r}, \tag{15}$$

that cannot be inferred from (13) and vice versa. To the best of the author’s knowledge, absolute lower dissipation bounds are known only in the Lipschitz setting $p = \infty$ and feature double exponential bounds \cite{[29]} (that are, however, consistent with (14) as pointed out in \cite{[11]}). We also refer to \cite{[6, 15, 13]} and the literature therein for bounds and sharpness results in the case of shear and circular flows.

It should be remarked that our scaling assumption on the initial data excludes certain non-generic scenarios of arbitrarily fast dissipating systems. One such example can be constructed as follows. Suppose that the initial datum $\theta_0$ is concentrated in Fourier space around a single wave number $k_0$, and consider a velocity field that aligns with the level sets of this function. Then $u \cdot \nabla \theta^* = 0$ during the evolution, and thus $\|\theta^*(t)\|_{L^2} \sim e^{-tCk_0^2\kappa} \|\theta_0\|_{L^2}$ for some $C > 0$ and any $t > 0$. The dissipation rate can thus be chosen arbitrarily large by letting $k_0 \gg 1$. However, we have $\|\nabla \theta_0\|_{L^2} \sim k_0 \|\theta_0\|_{L^2}$, and thus, such settings are discarded in Theorem 2.
We conclude the discussion with a comment on a possible extension of Theorem 2 to the stochastic turbulence studies performed recently in [4, 5].

Remark 1. By a small variation of its proof for the $s = \infty$ case, the statement in Theorem 2 can be extended to velocity fields with gradient bounds of the type

$$\int_0^t \|\nabla u\|_{L^2} \, dt \lesssim 1 + t.$$  

Such estimates (or some probabilistic analogous thereof) were obtained for the stochastically forced Navier–Stokes equations studied in [4, 5]. In particular, it seems that the enhanced dissipation estimates with $D \sim \log^{-1} \frac{1}{\kappa}$ obtained in Remark 1.4 of [4] are optimal. It is, however, unclear to the author, how the results of the present paper relate to the uniform-in-$\kappa$ rates computed in Theorem 1.3 of the same paper.

Because bounds in $\dot{H}^{-1}$ are stronger than enhanced dissipation bounds, the result in Theorem 2 can be transferred into an estimate on the rates of late-stage exponential mixing.

Theorem 3. Suppose that $u \in L^s(\mathbb{R}^d; W^{1,p}(\mathbb{T}^d))$ for some $p \in (1, \infty]$ and $s \in [1, \infty]$, satisfying $\|\nabla u\|_{L^p L^q} \leq 1$. Let $q \in [1, \infty]$ be such that $\frac{1}{p} + \frac{1}{q} \leq 1$ and $< 1$ if $p < 2$. Let $\theta_0$ be a mean-zero initial configuration satisfying $\|\theta_0\|_{L^1} \sim \|\theta_0\|_{L^q} \sim 1$ and $\|\nabla \theta_0\|_{L^1} \sim 1$. If there is a positive constant $D$ such that

$$\|\theta_\kappa(t)\|_{\dot{H}^{-1}} \lesssim \sqrt{\frac{\kappa}{D}} e^{-Dt},$$  

for any $t \gg \frac{1}{D}$, then

$$D \lesssim \begin{cases} \log^{-1} \frac{1}{\kappa} & \text{if } s > 1, \\ \kappa & \text{if } s = 1. \end{cases}$$  

Comparing Theorem 3 with (3), we observe an increase of the mixing rates for $s \in (1, \infty)$ and a limitation by diffusion for $s = \infty$. The latter was previously reported in the numerical study [29]. In the computational results obtained in there, a $\kappa$-dependent exponent could, however, not be identified. This phenomenon should be re-investigated in future numerical simulations.

Apart from the estimates established in Theorems 2 and 3 we also want to give a qualitative argument supporting the observation that exponential $\kappa$-independent lower bounds on mixing or on enhanced dissipation in the form (13) or (16) with $D \sim 1$ cannot be expected. Indeed, a very simple calculation reveals that $\kappa$-independent enhanced dissipation rates are related to the phenomenon of anomalous diffusion, that cannot occur in the setting of this paper. We recall that a system features anomalous dissipation if there exists a universal constant $\delta \in (0, \frac{1}{2})$ and a universal time scale $t_0 > 0$ such that

$$\delta \|\theta_0\|_{L^2}^2 \leq \kappa \int_0^t \|\nabla \theta_\kappa\|_{L^2}^2 \, dt.$$  

(18)
for any $t \geq t_0$, uniformly in $\kappa$. This, however, is excluded by DiPerna and Lions’s theory \[18\] because, otherwise, we could not infer (11) from (11) in the limit of vanishing diffusivity. The situation is different, for instance, for Hölder flows as constructed in \[19\]. The following lemma provides a relation between anomalous diffusion and dissipation enhancement at a uniform rate, which the author believes is not unknown in the community. It is included here for the non-expert readers.

**Lemma 1.** Suppose that there exist two constants $\Lambda > 0$ and $D > 0$ independent of $\kappa$ such that

$$\|\theta^\kappa(t)\|_{L^2} \leq \Lambda e^{-Dt}\|\theta_0\|_{L^2}$$

for any $t \geq 0$. Then there exist two constants $\delta > 0$ and $t_0 > 0$ independent of $\kappa$ such that (18) holds.

The converse is true provided that (18) holds uniformly in $\theta_0$.

This lemma also holds true in the $L^r$ instead of the $L^2$ setting with appropriate modifications (also in (18)).

We conclude this introduction with a comment. The prefactor occurring in (16) is a logarithmically corrected Batchelor scale $\ell_B$, which is defined in the literature by $\sqrt{\kappa}$ if $\|\nabla u\|_{L^2} \approx 1$. Phenomenologically, the Batchelor scale is the length scale at which flow-induced filamentation and diffusion balance. It is thus of the order of the width of the smallest filament that can exist due to straining before being damped by diffusion. If the bounds in Theorems 2 and 3 are attained by some configuration, it holds that

$$\ell_B \sim \frac{\|\theta^\kappa(t)\|_{H^{-1}}}{\|\theta^\kappa(t)\|_{L^2}} \sim \sqrt{\kappa \log \frac{1}{\kappa}}$$

because the $H^{-1}$ norm has the dimensions of length.

The occurrence of the Batchelor scale as a prefactor in large-time mixing rates is not unexpected, since, heuristically, $\|\theta^\kappa\|_{H^{-1}} \sim \ell_B$ and $\|\theta^\kappa\|_{L^2} \sim 1$ in the moment of the transition from advection-dominated to diffusion-dominated mixing. What is surprising to the author is the logarithmic correction in (20).

The seeming necessity of redefining the Batchelor scale for flows with (generalized) enstrophy constraints $\|\nabla u\|_{L^p L^q} \leq 1$ can already be seen on the level of the enhanced dissipation estimate. Heuristically, the large time dissipation rate is determined by the order of the smallest relevant wave number $k_B$, that is, $\|\theta^\kappa(t)\|_{L^2} \lesssim e^{-\kappa k_B^2 t}$. Our lower bound (14) suggests that $k_B \sim 1/\sqrt{\kappa \log \frac{1}{\kappa}}$. Now, if the system is stirred up to the Batchelor length $\ell_B$, it must hold that $k_B \sim \ell_B^{-1}$, leading to (20).

The author likes to stress that (20) is derived under the assumption that the bounds derived in Theorem 2 and 3 are indeed attained in some situations.

To summarize, our results show, at least in the case of Lipschitz regular flows, that early stage mixing rates are dictated by those which are relevant in the non-diffusive setting, Theorem 1. Moreover, the advection-dominated mixing stage lasts until a time scale of order $T_{\kappa,s}$ at which, according to Theorem 3, a crossover happens towards diffusion-dominated mixing at the associated exponential mixing rates. In view of Theorem 2, the enhanced dissipation occurs on the same time scale as the decay of the $H^{-1}$ mix-norm.
The remainder of this article is devoted to the proofs.

2 Proofs

2.1 Derivation of Theorem 3 from Theorem 2 and proof of Lemma 1

The following lemma shows that Theorem 2 implies Theorem 3.

Lemma 2. Suppose that \( \frac{1}{p} + \frac{1}{q} < 1 \) if \( p < 2 \). Let \( \theta_0 \) be a mean-zero initial configuration satisfying \( \| \theta_0 \|_{L^q} \lesssim 1 \). If there exists a positive constant \( D \) such that

\[
\| \theta^\kappa(t) \|_{H^{-1}} \lesssim \sqrt{\frac{\kappa}{D}} e^{-Dt}
\]

for any \( t \gg \frac{1}{D} \), then there exists a constant \( \alpha \in (0, 1) \) dependent only on \( p \) and \( q \) such that

\[
\| \theta^\kappa(t) \|_{L^p'} \lesssim e^{-\alpha Dt},
\]

for any \( t > 0 \).

In the statement of the lemma, \( p \) and \( p' \) are the Hölder dual exponents, \( \frac{1}{p} + \frac{1}{p'} = 1 \).

Proof. The statement for \( p = p' = 2 \) relies on the Lebesgue interpolation \( \| \theta^\kappa \|_{L^2}^2 \leq \| \nabla \theta^\kappa \|_{L^2} \| \theta^\kappa \|_{H^{-1}} \). Indeed, integrating this inequality in time and using the Cauchy–Schwarz inequality and the decay assumption on the \( H^{-1} \) norm in (21), we obtain

\[
\int_t^{t+\frac{1}{D}} \| \theta^\kappa \|_{L^2}^2 dt \leq \left( \int_t^{t+\frac{1}{D}} \| \nabla \theta^\kappa \|_{L^2}^2 dt \int_t^{t+\frac{1}{D}} \| \theta^\kappa \|_{H^{-1}}^2 dt \right)^{1/2}
\]

\[
\lesssim \frac{1}{D} e^{-Dt} \left( \int_t^{t+\frac{1}{D}} \| \nabla \theta^\kappa \|_{L^2}^2 dt \right)^{1/2}.
\]

Thanks to the energy estimate (11) and, hence, the monotonicity of the \( L^2 \) norm, the latter turns into

\[
\| \theta^\kappa(t + \frac{1}{D}) \|_{L^2}^2 \lesssim e^{-Dt} \| \theta_0 \|_{L^2} \leq e^{-Dt} \| \theta_0 \|_{L^q} \lesssim e^{-Dt},
\]

which yields (22) with \( \alpha = \frac{1}{2} \) for large times \( t \gg \frac{1}{D} \). For small times \( t \lesssim \frac{1}{D} \), this estimate holds true trivially since \( e^{-\frac{Dt}{2}} \sim 1 \).

If \( p \neq 2 \), we have to make use of the identity

\[
\| \theta^\kappa(t) \|_{L^q}^q + \kappa q(q-1) \int_s^t \int_{T^d} |\theta^\kappa|^{q-2} |\nabla \theta^\kappa|^2 \, dx \, dt = \| \theta^\kappa(s) \|_{L^q}^q,
\]

for any \( s < t \), which can be easily checked by a straightforward calculation.
If \( p < 2 \), then \( q > p' \) by assumption. We may thus interpolate \( \| \theta^\kappa \|_{L^p} \leq \| \theta^\kappa \|_{L^q} \| \theta^\kappa \|_{L^2}^{-\gamma} \) for some \( \gamma \in (0, 1) \), use the bound on the \( L^q \) norm in (23) and reduce the statement from the \( L^2 \) estimate.

Finally, if \( p > 2 \), we may without loss of generality assume that \( q = p' \). We use the interpolation \( \| \theta^\kappa \|_{L^q} \lesssim \| \nabla \theta^\kappa \|_{L^q} \| \theta^\kappa \|_{L^2} \), and bound \( \| \theta^\kappa \|_{H^{-1}} \leq \| \theta^\kappa \|_{L^2} \) because we are on a finite domain. Furthermore, using the Hölder inequality, we observe that

\[
\| \nabla \theta^\kappa \|_{L^q} \leq \left( \int_{\mathbb{T}^d} |\theta^\kappa|^{q-2} |\nabla \theta^\kappa|^2 \, dx \right)^{\frac{1}{2}} \| \theta^\kappa \|_{L^2} \sqrt{2}.
\]

Therefore,

\[
\| \theta^\kappa \|_{L^2}^{2q} \leq \left( \int_{\mathbb{T}^d} |\theta^\kappa|^{q-2} |\nabla \theta^\kappa|^2 \, dx \right)^{\frac{1}{2}} \| \theta^\kappa \|_{H^{-1}}.
\]

From here, the argument proceeds analogously to the \( p = 2 \) case, using (23) instead of (11).

We also establish a relation between anomalous diffusion and uniform mixing rates.

**Proof of Lemma 7.** We suppose first that \( u \) is dissipation enhancing with rate \( D \) and constant \( \Lambda \). Then, plugging (19) into the energy equality (11) yields

\[
\| \theta_0 \|_{L^2}^2 \leq \Lambda e^{-Dt} \| \theta_0 \|_{L^2}^2 + 2\kappa \int_0^t \| \nabla \theta^\kappa \|_{L^2}^2 \, dt,
\]

for any \( t > 0 \). Now, choosing \( t_0 > 0 \) large enough so that, for instance, \( \Lambda e^{-Dt} \leq \frac{1}{2} \) for any \( t \geq t_0 \) yields (18) with \( \delta = \frac{1}{4} \).

Now, if \( u \) generates anomalous diffusion with constant \( \delta \) and time scale \( t_0 \), we let \( n \in \mathbb{N} \) and deduce from the energy equality (11), the anomalous diffusion estimate (18), and iteration that

\[
\| \theta^\kappa(nt_0) \|_{L^2}^2 = \| \theta^\kappa((n-1)t_0) \|_{L^2}^2 - 2\kappa \int_{(n-1)t_0}^{nt_0} \| \nabla \theta^\kappa \|_{L^2}^2 \, dt
\leq (1 - 2\delta)^n \| \theta^\kappa((n-1)t_0) \|_{L^2}^2
\leq (1 - 2\delta)^n \| \theta_0 \|_{L^2}^2.
\]

Because \( (1 - 2\delta)^n \leq e^{-2\delta n} \), we infer for \( t \in [(n-1)t_0, nt_0] \) with \( n > 1 \) that

\[
\| \theta^\kappa(t) \|_{L^2}^2 \leq \| \theta^\kappa((n-1)t) \|_{L^2}^2 \leq e^{-2\delta n} \| \theta_0 \|_{L^2}^2 \leq e^{-2Dt} \| \theta_0 \|_{L^2}^2,
\]

where we have set \( D = \delta/t_0 \). For \( n = 1 \), this estimate is a trivial consequence of the energy equality (11). This proves (19). ■
2.2 Proof of Theorem 2

The proof of Theorem 2 does not make use of the transport equation (2). We may thus consider \( \theta \) in what follows as an arbitrary function.

We will make use of the following Kantorovich–Rubinstein distance with logarithmic cost function, which was introduced in \([35]\) in order to derive stability estimates for transport equations with Sobolev coefficients. For \( \delta > 0 \) and any mean zero function \( \theta \) on \( \mathbb{T}^d \), we define

\[
D_\delta(\theta) = \inf_{\pi \in \Pi(\theta^+, \theta^-)} \int \int \log \left( \frac{|x - y|}{\delta} + 1 \right) d\pi(x, y),
\]

where \( \theta^+ \) and \( \theta^- \) denote the positive and negative parts of \( \theta \), respectively, and \( \Pi(\theta^+, \theta^-) \) is the set of transport plans \( \pi : \mathbb{T}^d \times \mathbb{T}^d \to \mathbb{R}_+ \) with marginals \( \theta^+ \) and \( \theta^- \), i.e.,

\[
\int \varphi(x) + \psi(y) d\pi(x, y) = \int \varphi \theta^+ dx + \int \psi \theta^- dx,
\]

for all continuous functions \( \varphi \) and \( \psi \) on the torus. We remark that the Kantorovich–Rubinstein distance is finite only if \( \theta \) has zero mean, because then, both \( \theta^+ \) and \( \theta^- \) have the same total mass,

\[
\int \theta^+ dx = \int \theta^- dx.
\]

Our subsequent proofs will not use many properties of Kantorovich–Rubinstein distances, as some of the key estimates, above all the following lemma, can be taken from the existing literature. Yet, we refer the interested reader to \([37]\) for a comprehensive introduction into the theory of optimal transportation.

The rate of change of \( D_\delta \) under solutions to advection-diffusion equations has been investigated in \([31, 36]\), see also \([7, 34, 35]\) for related estimates in the purely advective case.

**Lemma 3** \((31, 36)\). Let \( \theta^\kappa \) be a mean-zero solution to the advection-diffusion equation \((1)\). Then \( D_\delta(\theta^\kappa) \) is absolutely continuous and it holds

\[
\frac{d}{dt} D_\delta(\theta^\kappa) \lesssim \| \nabla u \|_{L^p} \| \theta^\kappa \|_{L^{p'}} + \frac{\kappa}{\delta} \| \nabla \theta^\kappa \|_{L^1},
\]

(24)

where \( \frac{1}{p} + \frac{1}{p'} = 1 \).

Apart from its applications for mixing that we elaborate in the following, this estimate can be used to quantify the (weak) convergence in the vanishing diffusivity limit \( \kappa \to 0 \). In this context, \( \delta \) can be interpreted as the order of convergence. This observation has been exploited in order to bound the approximation error due to numerical diffusion generated by the upwind finite volume scheme for continuity equations in \([32, 33]\).

Our next result is a lower bound on the Kantorovich–Rubinstein distance in terms of the \( L^1 \) norms of \( \theta \) and its gradient.
Lemma 4. Let \( \theta \) be a mean zero function in \( W^{1,1}(\mathbb{T}^d) \). Then there exists a constant \( C > 0 \) such that
\[
D_\delta(\theta) \gtrsim \log \left( \frac{\|\theta\|_{L^1} + 1}{\delta C \|\nabla \theta\|_{L^1}} \right) \|\theta\|_{L^1}. \tag{25}
\]

The statement of the lemma is a consequence of an interpolation inequality between Kantorovich–Rubinstein distances with logarithmic cost function and the Sobolev norm, a variation of which was proved previously in \([7, 31, 34]\). It is a generalization of the endpoint Kantorovich–Sobolev inequality
\[
1 \lesssim \log^{-1} \left( \|\nabla \theta\|_{L^1}^{-1} + 1 \right) \inf_{\pi \in \Pi(\theta^+, \theta^-)} \iint \log (|x - y| + 1) \, d\pi(x, y)
\]
for probability distributions, see \([24]\) for standard Wasserstein versions.

**Proof.** We recall that by duality, it holds that
\[
\|\theta\|_{L^1} = \sup_{\|\psi\|_{L^\infty} \leq 1} \int \theta \psi \, dx. \tag{26}
\]
We now pick \( \psi \) arbitrary with \( \|\psi\|_{L^\infty} \leq 1 \) and denote by subscript \( R \) the convolution with a standard mollifier of scale \( R \). We then split
\[
\int \theta \psi \, dx = \int (\theta - \theta_R) \psi \, dx + \int \theta \psi_R \, dx, \tag{27}
\]
where we have used symmetry properties of the mollifier to shift the subscript from \( \theta \) to \( \psi \).

For the first term, we use the fact that \( \|\theta - \theta_R\|_{L^1} \lesssim R \|\nabla \theta\|_{L^1} \), so that
\[
\int (\theta - \theta_R) \psi \, dx \lesssim R \|\nabla \theta\|_{L^1}. \tag{28}
\]
For the second one, we introduce a second auxiliary length scale \( r \) and write
\[
\int \theta \psi_R \, dx
\]
\[
= \int (\theta^+ - \theta^-) \psi_R \, dx
\]
\[
= \iint (\psi_R(x) - \psi_R(y)) \, d\pi(x, y)
\]
\[
= \iint_{|x-y| \leq r} (\psi_R(x) - \psi_R(y)) \, d\pi(x, y) + \iint_{|x-y| > r} (\psi_R(x) - \psi_R(y)) \, d\pi(x, y),
\]
where \( \pi \in \Pi(\theta^+, \theta^-) \) is an arbitrary transport plan and we have used its marginal conditions in the second equality. On the one hand, because \( \psi_R \) is Lipschitz and \( \|\nabla \psi_R\|_{L^\infty} \lesssim \frac{1}{R} \|\psi\|_{L^\infty} \lesssim \frac{1}{R} \), we have that
\[
\iint_{|x-y| \leq r} (\psi_R(x) - \psi_R(y)) \, d\pi(x, y) \lesssim r \|\nabla \psi_R\|_{L^\infty} \iint d\pi(x, y) \lesssim \frac{r}{R} \|\theta\|_{L^1}.
\]
On the other hand, using $\|\psi_R\|_{L^\infty} \leq \|\psi\|_{L^\infty} \leq 1$, the monotonicity of the logarithm and setting $c(z) = \log \left( \frac{z}{\delta} + 1 \right)$, we estimate

$$\iint_{|x-y|>r} (\psi_R(x) - \psi_R(y)) \, d\pi(x,y) \leq \frac{2\|\psi_R\|_{L^\infty}}{c(r)} \iint c(|x-y|) \, d\pi(x,y) \lesssim \frac{1}{c(r)} \iint c(|x-y|) \, d\pi(x,y).$$

Combining the previous estimates and optimizing in $\pi$ on the right-hand side, we conclude that

$$\int \theta \psi_R \, dx \lesssim \frac{r}{R} \|\theta\|_{L^1} + \frac{1}{c(r)} D_\delta(\theta).$$

Plugging this estimate and (28) into the decomposition (27), we arrive at

$$\int \theta \psi \, dx \lesssim R \|\nabla \theta\|_{L^1} + \frac{r}{R} \|\theta\|_{L^1} + \frac{1}{c(r)} D_\delta(\theta),$$

for any $\psi$ such that $\|\psi\|_{L^\infty} \leq 1$. Maximizing in $\psi$ on the left-hand side and choosing $R \gg r$, we deduce that

$$\|\theta\|_{L^1} \lesssim r \|\nabla \theta\|_{L^1} + \frac{1}{c(r)} D_\delta(\theta),$$

and thus, the result follows upon choosing $r \ll \frac{\|\theta\|_{L^1}}{\|\nabla \theta\|_{L^1}}$. ■

We are now in the position to prove our bound on the dissipation rate.

**Proof of Theorem 2.** We may without loss of generality assume that $D < 1$. Moreover, by using Jensen’s inequality or an interpolation argument as in the proof of Lemma 2 we may restrict our attention to the case $r = q$.

We denote by $\delta t$ the dissipation time scale, i.e., $\delta t = \frac{1}{D}$, and set $t^n = n \delta t$. Integrating (24) over $[t^n, t^{n+1}]$ yields

$$\left| D_\delta(\theta^\kappa(t^{n+1})) - D_\delta(\theta^\kappa(t^n)) \right| \lesssim \int_{t^n}^{t^{n+1}} \|\nabla u\|_{L^p} \|\theta^\kappa\|_{L^q} \, dt + \frac{\kappa}{\delta} \int_{t^n}^{t^{n+1}} \|\nabla \theta^\kappa\|_{L^1} \, dt.$$

If $q \geq 2$, we use Jensen’s inequality and the energy estimate (11) to bound the gradient term on the right-hand side,

$$\frac{\kappa}{\delta} \int_{t^n}^{t^{n+1}} \|\nabla \theta^\kappa\|_{L^1} \, dt \leq \frac{\sqrt{\kappa \delta t}}{\delta} \left( \kappa \int_{t^n}^{t^{n+1}} \|\nabla \theta^\kappa\|_{L^2}^2 \, dt \right)^{\frac{1}{2}} \leq \frac{\sqrt{\kappa \delta t}}{\delta} \|\theta^\kappa(t^n)\|_{L^2} \leq \frac{\sqrt{\kappa \delta t}}{\delta} \|\theta^\kappa(t^n)\|_{L^4}.$$
Otherwise, if \( q \leq 2 \), we use the generalized energy equality
\[
\| \theta^\kappa(t_n+1) \|^q_{L^q} + \kappa q(q-1) \int_{t_n}^{t_{n+1}} \int_{\mathbb{T}^d} |\theta^\kappa|^{q-2} |\nabla \theta^\kappa|^2 \, dx \, dt = \| \theta^\kappa(t_n) \|^q_{L^q},
\]
which is derived via a standard computation, and estimate via interpolation and Jensen's inequality
\[
\frac{\kappa}{\delta} \int_{t_n}^{t_{n+1}} \| \nabla \theta^\kappa \|_{L^1} \, dt \leq \frac{\kappa}{\delta} \int_{t_n}^{t_{n+1}} \left( \int_{\mathbb{T}^d} |\theta^\kappa|^{q-2} |\nabla \theta^\kappa|^2 \, dx \right)^{\frac{1}{2}} \| \theta^\kappa \|_{L^{2-2q}}^{\frac{2-q}{q}} \, dt
\leq \sqrt{\kappa \delta t} \left( \kappa \int_{t_n}^{t_{n+1}} \int_{\mathbb{T}^d} |\theta^\kappa|^{q-2} |\nabla \theta^\kappa|^2 \, dx \, dt \right)^{\frac{1}{2}} \| \theta^\kappa(t_n) \|_{L^q}^{\frac{2-q}{q}}
\leq \frac{\sqrt{\kappa \delta t}}{\delta} \| \theta^\kappa(t_n) \|_{L^q}.
\]
In either case, we have
\[
|D_\delta(\theta^\kappa(t_{n+1})) - D_\delta(\theta^\kappa(t_n))| \lesssim \int_{t_n}^{t_{n+1}} \| \nabla u \|_{L^p} \| \theta^\kappa \|_{L^q} \, dt + \frac{\sqrt{\kappa \delta t}}{\delta} \| \theta^\kappa(t_n) \|_{L^2}.
\]
Invoking our assumption on the energy decay in (13), the monotonicity of the energy norm (23), and recalling that \( \delta t = \frac{1}{D} \) and thus \( DT_n = n \), the right-hand side can be further estimated,
\[
|D_\delta(\theta^\kappa(t_{n+1})) - D_\delta(\theta^\kappa(t_n))| \lesssim e^{-n} \int_{t_n}^{t_{n+1}} \| \nabla u \|_{L^p} \, dt + \frac{1}{\delta} \sqrt{\frac{\kappa}{D}} e^{-n}.
\]
Summing over \( n \) and invoking the triangle and Hölder's inequalities and the imposed bound on the velocity gradient yields
\[
D_\delta(\theta_0) \lesssim D_\delta(\theta^\kappa(t_N)) + \sum_{n=0}^{N} \left( e^{-n} \int_{t_n}^{t_{n+1}} \| \nabla u \|_{L^p} \, dt \right) + \frac{1}{\delta} \sqrt{\frac{\kappa}{D}}
\lesssim D_\delta(\theta^\kappa(t_N)) + \frac{1}{D^{1/2}} + \frac{1}{\delta} \sqrt{\frac{\kappa}{D}}.
\]
We now use the lower bound on our Kantorovich–Rubinstein distance (25) and the assumption on the initial datum to estimate the left-hand side from below. It holds that
\[
\log \left( \frac{1}{C\delta} + 1 \right) \lesssim D_\delta(\theta_0),
\]
for some \( C > 0 \). This constant can be chosen larger than 1 without restrictions, and thus,
\[
\log \left( \frac{1}{C\delta} + 1 \right) \geq \log \left( \frac{1}{\delta} + 1 \right) - \log C \geq \log \left( \frac{1}{\delta} + 1 \right),
\]
13
if $\delta$ is sufficiently small, which we will ensure later. On the other hand, because $|x - y| \leq 1$ on the torus, we have the following brutal estimate on the Kantorovich–Rubinstein distance

$$D_\delta(\theta^\kappa) \leq \log \left( \frac{1}{\delta} + 1 \right) \int \int d\pi(x, y) \lesssim \log \left( \frac{1}{\delta} + 1 \right) \|\theta^\kappa\|_{L^1},$$

and thus by (13), and the fact that $Dt^N = N$, (29) becomes

$$\log \left( \frac{1}{\delta} + 1 \right) \lesssim \log \left( \frac{1}{\delta} + 1 \right) e^{-N} + \frac{1}{D^{s-1}} + \frac{1}{\delta} \sqrt{\frac{\kappa}{D}}.$$

Because $N$ was arbitrary, the first term on the right-hand side can be dropped. We thus arrive at

$$\log \frac{1}{\delta} \lesssim \frac{1}{D^{s-1}} + \frac{1}{\delta} \sqrt{\frac{\kappa}{D}}.$$

If $s = 1$, choosing $\delta$ small but independent of $\kappa$ gives $D \lesssim \kappa$ as desired. For $s > 1$, the choice $\delta = \sqrt{\kappa}$ yields

$$\log \frac{1}{\kappa} \lesssim \frac{1}{D^{s-1}} + \frac{1}{D^{s-\frac{1}{2}}},$$

and thus, recalling that we assumed $D < 1$, we find $D \lesssim \log^{-\frac{1}{s-1}} \frac{1}{\kappa}$ if $s \geq 2$ and $D \lesssim \log^{-2} \frac{1}{\kappa}$ if $s < 2$. In the latter case we can do better. Indeed, choosing $\delta = \kappa^{\frac{1}{2}} \frac{1}{D^{\frac{1}{2} - \frac{1}{s}}}$, which is now guaranteed to be small if $\kappa$ is sufficiently small by this first (suboptimal) estimate on $D$, we find

$$\log \frac{1}{\kappa} \lesssim \log \frac{1}{\delta} \lesssim \frac{1}{D^{s-1}},$$

where in the first inequality we have used the preliminary bound on $D$ again. The claimed estimate for $s \in (1, 2)$ follows. 

\[ \blacksquare \]

2.3 Proof of Theorem

We finally turn to the proof of Theorem. It is based on the following estimate between solutions to the transport equation (2) and those to the advection-diffusion equation (3).

\textbf{Lemma 5.} Suppose that $\|\nabla u\|_{L^1_t L^\infty_x} \leq 1$ and $\|\theta_0\|_{L^2} \lesssim 1$. For any $\alpha \in (0, 1)$ and $s \in (1, \infty]$, it holds that

$$\|\theta(t) - \theta^\kappa(t)\|_{\dot{H}^{-1}} \lesssim \kappa^{1-\alpha},$$

for any $t \ll \log^{\frac{1}{s-1}} \frac{1}{\kappa}$. Moreover, if $s = 1$, it holds that

$$\|\theta(t) - \theta^\kappa(t)\|_{\dot{H}^{-1}} \lesssim \kappa t,$$

for any $t > 0$. 

14
Proof. We let $\varphi$ and $\varphi^\kappa$ denote the potentials associated with $\theta$ and $\theta^\kappa$, i.e., they solve $-\Delta \varphi = \theta$ and $-\Delta \varphi^\kappa = \theta^\kappa$, so that $\|\theta^\kappa - \theta\|_{\dot{H}^{-1}} = \|\nabla (\varphi^\kappa - \varphi)\|_{L^2}$. Differentiation of this quantity and multiple integration by parts then yield

$$\frac{1}{2} \frac{d}{dt} \|\nabla (\varphi^\kappa - \varphi)\|_{L^2}^2 = \int (\varphi^\kappa - \varphi) \partial_t (\theta^\kappa - \theta) \, dx$$

$$= \int (\varphi^\kappa - \varphi) u \cdot \nabla \Delta (\varphi^\kappa - \varphi) \, dx + \kappa \int (\varphi^\kappa - \varphi) \Delta \theta^\kappa \, dx$$

$$= \int \nabla (\varphi^\kappa - \varphi) \cdot \nabla u \nabla (\varphi^\kappa - \varphi) \, dx - \kappa \int \nabla (\varphi^\kappa - \varphi) \cdot \nabla \theta^\kappa \, dx.$$

We use the fact that $u$ is Lipschitz continuous with respect to the spatial variable and apply the Cauchy–Schwarz inequality to deduce that

$$\frac{d}{dt} \|\theta^\kappa - \theta\|_{\dot{H}^{-1}} \leq \|\nabla u\|_{L^\infty} \|\theta^\kappa - \theta\|_{\dot{H}^{-1}} + \kappa \|\nabla \theta^\kappa\|_{L^2}.$$

Applying the chain rule, we now observe that the latter implies the following control on a logarithm of the $H^{-1}$ norm,

$$\frac{d}{dt} \log \left( \frac{1}{\delta} \|\theta^\kappa - \theta\|_{\dot{H}^{-1}} + 1 \right) \leq \|\nabla u\|_{L^\infty} + \frac{\kappa}{\delta} \|\nabla \theta^\kappa\|_{L^2},$$

for any $\delta > 0$. In particular, integrating in time and using Jensen’s inequality, the imposed bound on the velocity gradient and the energy estimate (11) gives

$$\log \left( \frac{1}{\delta} \|\theta^\kappa(t) - \theta(t)\|_{\dot{H}^{-1}} + 1 \right) \leq T^{\frac{1}{2} - \frac{1}{s}} + \sqrt{\frac{\kappa T}{\delta}} \|\theta_0\|_{L^2},$$

for any $t \in [0, T]$ and any $T > 0$. The choice $\delta = \kappa^{\frac{s}{2}} T^{\frac{1}{2} - \frac{1}{s}}$ and the assumption on the initial configuration yield

$$\log \left( \frac{1}{\delta} \|\theta^\kappa(t) - \theta(t)\|_{\dot{H}^{-1}} + 1 \right) \lesssim T^{\frac{1}{2} - \frac{1}{s}},$$

and thus

$$\|\theta^\kappa(t) - \theta(t)\|_{\dot{H}^{-1}} \lesssim \kappa^{\frac{s}{2}} T^{\frac{1}{2} - \frac{1}{s}} \left( e^{CT^{\frac{1}{2} - \frac{1}{s}}} - 1 \right),$$

for some $C > 0$. From here, the statement follows immediately. $lacksquare$

We also provide an estimate on the maximal growth rate of gradients.

**Lemma 6.** Suppose that $\theta_0$ is a mean-zero initial configuration satisfying $\|\nabla \theta_0\|_{L^2} \lesssim 1$, and assume that $\|\nabla u\|_{L^s L^\infty} \leq 1$. Then

$$\|\nabla \theta(t)\|_{L^2} \lesssim e^{t^{\frac{1}{2} - \frac{1}{s}}},$$

(32)

for any $t > 0$. 

15
Notice that an estimate of this type fails to be true if the advecting velocity field is only Sobolev regular, see [23, 25, 8, 10, 2] for optimal regularity estimates and examples for loss of regularity.

Proof. The proof is straightforward. Estimating the advection equation (2) with respect to $x_i$, yields

$$\partial_t \partial_i \theta + u \cdot \nabla \partial_i \theta = \partial_i u \cdot \nabla \theta.$$ 

Testing with $\partial_i \theta$, summing over $i$, and using the incompressibility condition on $u$ implies the bound,

$$\frac{1}{2} \frac{d}{dt} \|\nabla \theta\|_{L^2}^2 \leq \|\nabla u\|_{L^\infty} \|\nabla \theta\|_{L^2}^2,$$

from which we deduce (32) by integration, Hölder’s inequality and the assumptions on $\theta_0$ and $u$. 

We conclude with the proof of Theorem 1.

Proof of Theorem 1. We only treat the case $s > 1$. The case $s = 1$ can be proved analogously.

We infer from the triangle inequality and (30) that

$$\|\|\theta(t)\|_{\dot{H}^{-1}} - \|\theta^c(t)\|_{\dot{H}^{-1}}\| \lesssim \kappa^\frac{1}{4},$$

for $t \ll \log \frac{s}{s-1} \frac{1}{\kappa}$. In particular, if $\theta$ obeys the exponential bound in (6), it holds that

$$\|\theta^c(t)\|_{\dot{H}^{-1}} \lesssim e^{-A t \frac{s}{s-1}} + \kappa^\frac{1}{4} \lesssim e^{-A t \frac{s}{s-1}},$$

which proves (7).

The derivation of (9) from (8) proceeds analogously.

In order to obtain (10), we write

$$0 \leq \|\theta^c - \theta\|_{L^2}^2 = \|\theta^c\|_{L^2}^2 - 2 \int \theta(\theta^c - \theta) \, dx - \|\theta\|_{L^2}^2,$$

and thus, using the energy conservation in (4) and the Cauchy–Schwarz inequality in Fourier space, we find

$$\|\theta_0\|_{L^2}^2 \leq \|\theta^c\|_{L^2}^2 + 2\|\nabla \theta\|_{L^2} \|\theta^c - \theta\|_{\dot{H}^{-1}}.$$

Thanks to the gradient bound in (32) and estimate (30), we have on the one hand that

$$\|\nabla \theta(t)\|_{L^2} \|\theta^c(t) - \theta(t)\|_{\dot{H}^{-1}} \lesssim \kappa^\frac{1}{4} e^{t \frac{s}{s-1}},$$

for any $t \ll \log \frac{s}{s-1} \frac{1}{\kappa}$. On the other hand, we have $\|\theta_0\|_{L^2} \gtrsim 1$ be assumption. Hence, for $t \ll \log \frac{s}{s-1} \frac{1}{\kappa}$, we find that

$$1 \lesssim \|\theta^c(t)\|_{L^2}^2,$$

which is what we had to prove. 

■
Acknowledgement

This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy EXC 2044 –390685587, Mathematics Münster: Dynamics–Geometry–Structure. The author thanks Hung Nguyen for sharing his preprint [11]. He moreover thanks Jacob Bedrossian for suggesting to investigate the stochastic setting, briefly discussed in Remark [1].

References

[1] Alberti, G., Crippa, G., and Mazzucato, A. L. Exponential self-similar mixing by incompressible flows. *J. Amer. Math. Soc.* 32, 2 (2019), 445–490.

[2] Alberti, G., Crippa, G., and Mazzucato, A. L. Loss of regularity for the continuity equation with non-Lipschitz velocity field. *Ann. PDE* 5, 1 (2019), Art. 9, 19.

[3] Batchelor, G. K. Small-scale variation of convected quantities like temperature in turbulent fluid. I. General discussion and the case of small conductivity. *J. Fluid Mech.* 5 (1959), 113–133.

[4] Bedrossian, J., Blumenthal, A., and Punshon-Smith, S. Almost-sure enhanced dissipation and uniform-in-diffusivity exponential mixing for advection-diffusion by stochastic navier-stokes, 2019.

[5] Bedrossian, J., Blumenthal, A., and Punshon-Smith, S. Almost-sure exponential mixing of passive scalars by the stochastic navier-stokes equations, 2019.

[6] Bedrossian, J., and Coti Zelati, M. Enhanced dissipation, hypoellipticity, and anomalous small noise inviscid limits in shear flows. *Arch. Ration. Mech. Anal.* 224, 3 (2017), 1161–1204.

[7] Brenier, Y., Otto, F., and Seis, C. Upper bounds on coarsening rates in demixing binary viscous liquids. *SIAM J. Math. Anal.* 43, 1 (2011), 114–134.

[8] Bresch, D., and Jabin, P.-E. Global existence of weak solutions for compressible Navier-Stokes equations: thermodynamically unstable pressure and anisotropic viscous stress tensor. *Ann. of Math. (2)* 188, 2 (2018), 577–684.

[9] Bressan, A. A lemma and a conjecture on the cost of rearrangements. *Rend. Sem. Mat. Univ. Padova* 110 (2003), 97–102.

[10] Bruè, E., and Nguyen, Q.-H. Sharp regularity estimates for solutions of the continuity equation drifted by sobolev vector fields, 2018.

[11] Bruè, E., and Nguyen, Q.-H. Advection diffusion equations with sobolev velocity field, 2020.
[12] Constantin, P., Kiselev, A., Ryzhik, L., and Zlatoš, A. Diffusion and mixing in fluid flow. Ann. of Math. (2) 168, 2 (2008), 643–674.

[13] Coti Zelati, M. Stable mixing estimates in the infinite péclet number limit, 2019.

[14] Coti Zelati, M., Delgadino, M. G., and Elgindi, T. M. On the relation between enhanced dissipation timescales and mixing rates. Comm. Pure Appl. Math. (online first) (2019).

[15] Coti Zelati, M., and Drivas, T. D. A stochastic approach to enhanced diffusion, 2019.

[16] Crippa, G., and De Lellis, C. Estimates and regularity results for the DiPerna-Lions flow. J. Reine Angew. Math. 616 (2008), 15–46.

[17] Depauw, N. Non unicité des solutions bornées pour un champ de vecteurs BV en dehors d’un hyperplan. C. R. Math. Acad. Sci. Paris 337, 4 (2003), 249–252.

[18] DiPerna, R. J., and Lions, P.-L. Ordinary differential equations, transport theory and Sobolev spaces. Invent. Math. 98, 3 (1989), 511–547.

[19] Drivas, T. D., Elgindi, T. M., Iyer, G., and Jeong, I.-J. Anomalous dissipation in passive scalar transport, 2019.

[20] Elgindi, T. M., and Zlatoš, A. Universal mixers in all dimensions. Adv. Math. 356 (2019), 106807, 33.

[21] Feng, Y., and Iyer, G. Dissipation enhancement by mixing. Nonlinearity 32, 5 (2019), 1810–1851.

[22] Iyer, G., Kiselev, A., and Xu, X. Lower bounds on the mix norm of passive scalars advected by incompressible enstrophy-constrained flows. Nonlinearity 27, 5 (2014), 973–985.

[23] Jabin, P.-E. Critical non-Sobolev regularity for continuity equations with rough velocity fields. J. Differential Equations 260, 5 (2016), 4739–4757.

[24] Ledoux, M. Sobolev-Kantorovich inequalities. Anal. Geom. Metr. Spaces 3, 1 (2015), 157–166.

[25] Léger, F. A new approach to bounds on mixing. Math. Models Methods Appl. Sci. 28, 5 (2018), 829–849.

[26] Lin, Z., Thiffeault, J.-L., and Doering, C. R. Optimal stirring strategies for passive scalar mixing. J. Fluid Mech. 675 (2011), 465–476.

[27] Lunasin, E., Lin, Z., Novikov, A., Mazzucato, A., and Doering, C. R. Optimal mixing and optimal stirring for fixed energy, fixed power, or fixed palenstrophy flows. J. Math. Phys. 53, 11 (2012), 115611, 15.
[28] Mathew, G., Mezić, I., and Petzold, L. A multiscale measure for mixing. *Phys. D* 211, 1-2 (2005), 23–46.

[29] Miles, C. J., and Doering, C. R. Diffusion-limited mixing by incompressible flows. *Nonlinearity* 31, 5 (2018), 2346–2350.

[30] Modena, S., and Székelyhidi, Jr., L. Non-uniqueness for the transport equation with Sobolev vector fields. *Ann. PDE* 4, 2 (2018), Art. 18, 38.

[31] Otto, F., Seis, C., and Slepčev, D. Crossover of the coarsening rates in demixing of binary viscous liquids. *Commun. Math. Sci.* 11, 2 (2013), 441–464.

[32] Schlichting, A., and Seis, C. Convergence rates for upwind schemes with rough coefficients. *SIAM J. Numer. Anal.* 55, 2 (2017), 812–840.

[33] Schlichting, A., and Seis, C. Analysis of the implicit upwind finite volume scheme with rough coefficients. *Numer. Math.* 139, 1 (2018), 155–186.

[34] Seis, C. Maximal mixing by incompressible fluid flows. *Nonlinearity* 26, 12 (2013), 3279–3289.

[35] Seis, C. A quantitative theory for the continuity equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 34, 7 (2017), 1837–1850.

[36] Seis, C. Optimal stability estimates for continuity equations. *Proc. Roy. Soc. Edinburgh Sect. A* 148, 6 (2018), 1279–1296.

[37] Villani, C. *Topics in optimal transportation*, vol. 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[38] Wei, D. Diffusion and mixing in fluid flow via the resolvent estimate, 2018.

[39] Yao, Y., and Zlatoš, A. Mixing and un-mixing by incompressible flows. *J. Eur. Math. Soc. (JEMS)* 19, 7 (2017), 1911–1948.